Springer Texts in Business and Economics

Giancarlo Gandolfo

International Trade Theory and Policy

With contributions by Federico Trionfetti

Second Edition

Springer

Springer Texts in Business and Economics

For further volumes:
http://www.springer.com/series/10099

Giancarlo Gandolfo

International Trade Theory and Policy

Second edition

With contributions by Federico Trionfetti

Springer

Giancarlo Gandolfo
Accademia Nazionale dei Lincei
Classe di Scienze Morali
Rome, Italy

ISBN 978-3-642-37313-8 ISBN 978-3-642-37314-5 (eBook)
DOI 10.1007/978-3-642-37314-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013943689

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the memory of my father
Edgardo Gandolfo

Preface

There is no lack of good international economics textbooks ranging from the elementary to the advanced, so that an additional drop in this ocean calls for an explanation. In the present writer's opinion, there seems still to be room for a textbook which can be used in both undergraduate and graduate courses and which contains a wide range of topics, including those usually omitted from other textbooks. These are the intentions behind the present book, which is an outcrop from undergraduate and graduate courses in international economics that the author has been holding at the Sapienza University of Rome and other universities from 1974 to 2010 and from his ongoing research work in this field.

Accordingly, the work is organised as two-books-in-one by distributing the material between text and appendices.

The treatment in the text is directed to undergraduate students and is mainly confined to graphic analysis and to some elementary algebra, but it is assumed that the reader will have a basic knowledge of microeconomics (so that the usual review material on production functions, indifference curves, etc. is omitted). Each chapter has a mathematical appendix, where (i) the topics treated in the text are examined at a level suitable for advanced undergraduate or first-year graduate students and (ii) generalisations and/or topics not treated in the text (including some at the frontiers of research, whose often obscure mathematical aspects are fully clarified) are formally examined.

The text is self-contained, and the appendices can be read independently of the text and can, therefore, also be used by students who already know 'graphic' international economics and want to learn something about its mathematical counterpart. Of course the connections between text and appendices are carefully indicated, so that the latter can be used as mathematical appendices by the student who has mastered the text and the text can be used as graphic and literary exposition of the results derived mathematically in the appendices by the student who has mastered these.

The book is mainly analytical, although reality is present through sections on the empirical verification of the main theories and through case studies and other empirical materials contained in appropriate boxes. However, by stressing the analytical aspects, the author hopes to give the student the tools for an understanding of facts and policies—tools that will survive the circumstances of the passing day.

This new edition has been thoroughly revised and enriched thanks to the contributions by Professor Federico Trionfetti of Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS, that bring the book up to date. He has contributed sections of Chaps. 4, 9, and the entire Chaps. 16 and 17, plus several minor revisions. He wishes to thank the participants in the Brixen Workshop and Summer School 2012 as well as the master's students of the Aix-Marseille School of Economics for precious comments.

* * *

I am grateful to the students from all over the world who have written me over the years to indicate unclear points and misprints of the previous editions and to Marianna Belloc, Nicola Cetorelli, Giuseppe De Arcangelis, Vivek H. Dehejia, Laura Sabani, and Francesca Sanna Randaccio, for their advice and comments. I am particularly indebted to Daniela Federici, who has made very useful suggestions as regards the new material, then checked it with painstaking care.

None of the persons mentioned have any responsibility for possible deficiencies that might remain.

Rome, Italy Giancarlo Gandolfo

Contents

List of Figures

List of Tables

List of Boxes

Part I
Introduction

Chapter 1
Introduction to International Trade Theory and Policy

1.1 International Economics as a Distinct Subject

While several specialistic fields of economics have been developed as distinct branches of general economic theory only in relatively recent times, the presence of a specific treatment of the theory of international economic transactions is an old and consolidated tradition in the economic literature. Various reasons can be advanced to explain the need for this specific treatment, but the main ones are the following.

The first is that factors of production are generally less mobile between countries than within a single country. Traditionally, this observation has been taken as a starting point for the development of a theory of international trade based on the extreme assumption of perfect national mobility and perfect international immobility of the factors of production, accompanied by the assumption of perfect mobility (both within and between countries) of the commodities produced, exception being made for possible restrictive measures on the part of governments.

The second is the fact that the mere presence of different countries as distinct political entities each with its own frontiers gives rise to a series of problems which do not occur in general economics, such as the levying of duties and other impediments to trade, the existence of different national currencies whose relative prices (the exchange rates) possibly vary through time, etc.

The References (Bhagwati et al. 1998; Caves et al. 2006; Feenstra and Taylor 2008; Jones and Neary 1984; Krugman et al. 2011; Salvatore 2010; Södersten and Reed 1994; Woodland 1982) at the end of this chapter are a list of recent and less recent textbooks where the nature of international economics is further elucidated.

The specialistic nature of international economics—a discipline of increasing importance given the increasing openness of the single national economic systems—does *not* mean that its methods and tools of analysis are different from those of general economic theory: on the contrary, international economics makes ample use of the methods and tools of microeconomics and macroeconomics, as we shall see presently.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_1,

As in any other discipline, also in international economics we can distinguish a *theoretical* and a *descriptive* part. The former is further divided into the *theory of international trade* and *international monetary economics*. All these distinctions are of a logical and pedagogical nature, but of course both the descriptive and the theoretical part, both the trade and the monetary branch, are necessary for an understanding of the international economic relations in the real world.

The descriptive part, as the name clearly shows, is concerned with the description of international economic transactions just as they happen and of the institutional context in which they take place: flows of goods and financial assets, international agreements, international organizations like the World Trade Organization and the European Union, and so forth.

The theoretical part tries to go beyond the phenomena to seek general principles and logical frameworks which can serve as a guide to the understanding of actual events (so as, possibly, to influence them through policy interventions). Like any economic theory, it uses for this purpose abstractions and models, often expressed in mathematical form. The theoretical part can be further divided, as we said above, into trade and monetary theory each containing aspects of both *positive* and *normative* economics; although these aspects are strictly intertwined in our discipline, they are usually presented separately for didactic convenience.

A few words are now in order on the distinction between international trade theory and international monetary theory.

The *theory of international trade* (which has an essentially microeconomic nature) deals with the causes, the structure and the volume of international trade (that is, which goods are exported, which are imported, and why, by each country, and what is their amount); with the gains from international trade and how these gains are distributed; with the determination of the relative prices of goods in the world economy; with international specialization; with the effects of tariffs, quotas and other impediments to trade; with the effects of international trade on the domestic structure of production and consumption; with the effects of domestic economic growth on international trade and vice versa; and so on. The distinctive feature of the theory of international trade is the assumption that trade takes place in the form of *barter* (or that money, if present, is only a veil having no influence on the underlying real variables but serving only as a reference unit, the *numéraire*). A by-no-means secondary consequence of this assumption is that the international accounts of any country vis-à-vis all the others always balance: that is, no balance-of-payments problem exists.

This part of international economics was once also called the *pure* theory of international trade, where the adjective "pure" was meant to distinguish it from *monetary* international economics.

International monetary theory (which is essentially of a macroeconomic nature) deals with the problems deriving from balance-of-payments disequilibria in a monetary economy, and in particular with the automatic adjustment mechanisms and the adjustment policies of the balance of payments; with the relationships between the balance of payments and other macroeconomic variables; with the various exchange-rate regimes; with the problems of international liquidity and other problems of the international monetary system; etc.

In this book we shall treat the theory of international trade. A companion volume treats international monetary theory, thus following the standard practice of international textbooks and courses.

One last word: in this work we shall be concerned mainly with the theoretical part (both positive and normative) of international economics, even if references to the real world will not be lacking. Thanks to the advances in econometrics and computer power, practically all theories of international trade have been subjected to a great number of empirical tests. As it would not be possible to consider all these tests, it was necessary to make occasionally arbitrary choices, though we feel that the most important empirical studies have been treated. In any case, where no treatment is given, we have referred the reader to the relevant empirical literature.

1.2 The Theory and Policy of International Trade: An Overview

The foundations of international trade theory are contained in three main models aimed at explaining the determinants of international trade and specialization:

1. The classical (Torrens-Ricardo) theory, according to which these determinants are to be found in technological differences between countries;
2. The Heckscher-Ohlin theory, which stresses the differences in factor endowments between different countries;
3. The neoclassical theory (which has had a longer gestation: traces can be found in J.S. Mill; A. Marshall takes it up again in depth, and numerous modern writers bring it to a high level of formal sophistication), according to which these determinants are to be found simultaneously in the differences between technologies, factor endowments, and tastes of different countries. The last element accounts for the possible presence of international trade, even if technologies and factor endowments were completely identical between countries.

From the chronological point of view, model (2) post-dates model (1), while model (3), as we said, has had a longer gestation and so has been developing in parallel to the others.

To avoid misunderstandings it must be stressed that the Heckscher-Ohlin theory is also neoclassical (in the sense in which the neoclassical vision is different from the classical one), as it accepts all the logical premises of, and follows the, neoclassical methodology. As a matter of fact the Heckscher-Ohlin model can be considered as a particular case of the neoclassical one in which internationally identical production functions and tastes are assumed. This loss in degree of generality is, according to some authors, the price that has to be paid if one wishes to obtain definite conclusions about the structure of the international trade of a country.

These models are treated in detail in Part II.

Part III is devoted to the *new explanations* of international trade. These are the theories which drop either one or both of the two fundamental assumptions of the traditional theory (perfect competition and product homogeneity), and analyse international trade in a context of imperfect competition and/or product differentiation.

Part IV deals with the problems of commercial policy, including the debate between free trade and protectionism. The *new protectionism*, whereby protection is based on non-tariff instruments and comes about through administrative procedures or lobbying activities, is examined in depth. *Strategic trade policy*, which assumes the presence of interaction between the firms involved in international trade (when the action taken by any one firm may have significant effects on other firms) will also be examined in this part.

Part V deals with the relations between *international trade and growth*, first in a comparative-static and then in a dynamic context. We shall examine both the traditional view and the new models based on the interaction between *endogenous growth* and the new trade theories.

Part VI treats a topic strictly related to international trade, namely *globalization*, examined both in its relation to the *new economic geography* and in its relation to wage inequality between nations.

1.3 Small and Large Open Economies

We shall use both *one-country* and *two-country* models. With the expression *one-country* or *small-country* model (also called SOE, *small open economy*) we refer to a model in which the rest of the world is taken as exogenous, in the sense that what happens in the country under consideration (call it country 1) is assumed to have a negligible influence (since this country is small relative to the rest of the world) on the rest-of-the-world variables (in particular, the terms of trade). This means that these variables can be taken as *exogenous* in the model.

With the expression *two-country* or *large country* model we refer to a model in which the effect on the rest-of-the-world's variables of country 1's actions cannot be neglected, so that the rest of the world has to be explicitly included in the analysis (as country 2). It follows that, through the channels of exports and imports of goods and services, and capital movements, the economic events taking place in a country have repercussions on the other country, and vice versa.

Two-country models may seen more realistic, as in the real world inter-country repercussions do take place. However, in such models the various countries making up the rest of the world are assumedly aggregated into a single whole (country 2), which is not necessarily more realistic. In fact, if the world is made up of n interdependent countries which interact more and more with one another (*globalization* is the fashionable word for this increasing interdependence and interaction), dealing with it as if it were a two-country world is not necessarily better than using the SOE assumption as a first approximation. These problems can be overcome by the construction of n-country models, which will be examined in the relevant Appendixes, given their degree of mathematical difficulty.

References

Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (chap. 1).

Caves, R. E., Frankel, J. A., & Jones, R. W. (2006). *World trade and payments: An introduction* (chap. 1).

Feenstra, R. C., & Taylor, A. M. (2008). *International economics* (Pt. I).

Jones, R. W., & Neary, J. P. (1984). *The positive theory of international trade*.

Krugman, P. R., Obstfeld, M., & Melitz, M. J. (2011). *International economics: Theory and policy* (chap. 1).

Salvatore, D. (2010). *International economics: Trade and finance* (chap. 1).

Södersten, B., & Reed, G. (1994). *International economics* (chap. 1).

Woodland, A. D. (1982). *International trade and resource allocation* (chap. 1).

Practically all international economics textbooks contain a treatment to show why it exists as a distinct subject and to illustrate its subdivisions. We therefore mention only a few; others might have served equally well. Ample references concerning the various theories will be given in the following chapters.

Throughout this book, the end-of-chapter references will be indicated only by name(s), date, title. Complete information as to publisher, place of publication, etc., is contained in the Bibliography at the end of this volume.

Part II
Foundations of Trade Theory

Chapter 2
The Classical (Ricardo-Torrens) Theory of Comparative Costs

2.1 Comparative Costs (Advantages) and International Trade

The classical theory of international trade is usually attributed to David Ricardo, who treated it in Chap. 7 of his *Principles* (Ricardo, 1817). But it is possible to find earlier statements of this theory in the work of Robert Torrens (1815): the reader interested in problems of historical priority should consult Viner (1937) and Chipman (1965a).

As far as the theory itself is concerned, we begin by observing that it affirms that the crucial variable explaining the existence and pattern of international trade is technology. A difference in *comparative costs* of production—the necessary condition for international exchange to occur—does, in fact, reflect a difference in the techniques of production. The theory also aims at showing that trade is beneficial to all participating countries.

If we simplify to the utmost, we can assume that there are two countries (England and Portugal in the famous example of Ricardo's), two commodities (cloth and wine), that all factors of production can be reduced to a single one, labour,[1] and that in both countries the production of the commodities is carried out according to fixed technical coefficients: as a consequence, the unit cost of production of each commodity (expressed in terms of labour) is constant.

It is clear that if one country is superior to the other in one line of production (where the superiority is measured by a lower unit cost) and inferior in the other line, the basis exists for a fruitful international exchange, as earlier writers, for example

[1] This is based on the classical labour theory of value. It is outside the scope of the present treatment to enter into the controversies concerning this theory, so that we shall simply observe that the validity of the classical theory of international trade is not based on the validity of the labour theory of value, as it is sufficient for unit costs of production to be measurable by a common unit across countries and to be constant.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_2,

Table 2.1 Example of absolute advantage

	Unit costs of production in terms of labour	
Commodities	In England	In Portugal
Cloth	4	6
Wine	8	3

Adam Smith, had already shown. The simple example in Table 2.1 is sufficient to make the point; the reader should bear in mind that here as in the subsequent examples, the cost of transport is assumed to be absent, as its presence would complicate the treatment without altering the substance of the theory. As we see, the unit cost of manufacturing cloth is lower in England than in Portugal while the opposite is true for wine production. It is therefore advantageous for England to specialize in the production of cloth and to exchange it for Portuguese wine, and for Portugal to specialize in the production of wine and to exchange it for British cloth. Suppose, for example, that the (international) *terms of trade* (i.e., the ratio according to which the two commodities are exchanged for each other between the two countries, or international relative price) equals one, that is, international exchange takes place on the basis of one unit of wine for one unit of cloth. Then England with 4 units of labour (the cost of one unit of cloth) obtains one unit of wine, which otherwise—if produced internally—would have required 8 units of labour. Similarly Portugal with 3 units of labour (the cost of one unit of wine) obtains one unit of cloth, which otherwise—if produced internally—would have required 6 units of labour.

In this example we have reasoned in terms of *absolute* costs, as one country has an absolute advantage in the production of one commodity and the other country has an absolute advantage in the production of the other. That in such a situation international trade will take place and benefit all participating countries is obvious. Less so is the fact that international trade may equally well take place even if one country is superior to the other in the production of *both* commodities. The great contribution of the Ricardian theory was to show the conditions under which even in this case international trade is possible (and beneficial to both countries).

Now, this theory affirms that the necessary condition for international trade is, in any case, that a difference in *comparative costs* exists. Comparative cost can be defined in *two* ways: as the ratio between the (absolute) unit costs of the two commodities in the same country, or as the ratio between the (absolute) unit costs of the *same* commodity in the two countries. Following common practice, we shall adopt the former, but they are totally equivalent.

In fact, if we denote the unit costs of production of a good in the two countries by a_1, a_2 (where the letter refers to the good and the numerical subscript to the country: this notation will be constantly followed throughout the book) and the unit costs of the other good by b_1, b_2, then

$$(a_1/b_1 = a_2/b_2) \Longleftrightarrow (b_1/a_1 = b_2/a_2) \Longleftrightarrow (a_1/a_2 = b_1/b_2) \Longleftrightarrow (a_2/a_1 = b_2/b_1),$$

Table 2.2 Example of comparative advantage

	Unit costs of production in terms of labour	
Commodities	In England	In Portugal
Cloth	4	6
Wine	8	10

and similarly

$$(a_1/b_1 \gtrless a_2/b_2) \Longleftrightarrow (a_1/a_2 \gtrless b_1/b_2) \Longleftrightarrow (b_2/a_2 \gtrless b_1/a_1) \Longleftrightarrow (b_2/b_1 \gtrless a_2/a_1).$$

It therefore makes no difference whether the comparison is made between a_1/b_1 and a_2/b_2 or between a_1/a_2 and b_1/b_2, and so on.

The basic proposition of the theory under examination is that *the condition for international trade to take place is the existence of a difference between the comparative costs*. This is, however, a necessary condition only; the sufficient condition is that the international terms of trade lie *between* the comparative costs without being equal to either. When both conditions are met, it will be beneficial to each country to specialize in the production of the commodity in which it has the relatively greater advantage (or the relatively smaller disadvantage). Let us consider the following example (Table 2.2).

As England is superior to Portugal in the production of both commodities, it might seem that there is no scope for international trade, but this is not so. Comparative costs are $4/8 = 0.5$ and $6/10 = 0.6$ in England and Portugal respectively. England also has a relatively greater advantage (a *comparative advantage*) in the production of cloth: its unit cost, in fact, is lower in England than in Portugal by 33.3 % (2/6), while the unit cost of wine is lower in the former than in the latter country by 20 % (2/10). It can similarly be seen that Portugal has a relatively smaller disadvantage in the production of wine: its unit cost, in fact, is higher in Portugal than in England by 25 % (2/8), while the unit cost of cloth is higher in Portugal than in England by 50 % (2/4).

Therefore—provided that the terms of trade are greater than 0.5 and smaller than 0.6—British cloth will be exchanged for Portuguese wine to the benefit of both countries. Let us take an arbitrary admissible value of the terms of trade, say 0.55 (that is, international exchange takes place at the terms of 0.55 units of wine per one unit of cloth). In England, on the basis of the existing technology, one unit of cloth exchanges for 0.5 units of wine: 0.5 is, in fact, the comparative cost, and, according to the classical theory, the relative prices of goods, that is their exchange ratios, are determined by costs. For one unit of cloth, England can obtain, by way of international trade, 0.55 units of wine, more than the amount obtainable internally. Similarly in Portugal, to obtain one unit of cloth, 0.6 units of wine (0.6 is Portugal's comparative cost) are necessary, while by way of international trade only 0.55 units of wine are required. It is obvious that international trade is beneficial to both countries.

It is possible to arrive at the same conclusion by reasoning in terms of production costs. England with 4 units of labour (the cost of one unit of cloth) obtains, on the international market, 0.55 units of wine which, if produced internally, would have required $0.55 \times 8 = 4.4$ units of labour. Similarly Portugal with 5.5 units of labour (the cost of 0.55 units of wine, given by 0.55×10) obtains one unit of cloth, which would have required 6 units of labour if produced internally.

It can easily be shown that the terms of trade must be strictly located between the two comparative costs. If, in fact, the terms of trade were equal to either comparative cost, the concerned country would have no interest in trading, since the internal price ratio (given by the comparative cost) would be equal to the international one (the terms of trade). This would mean that the country in question would obtain the other commodity by way of trade at the same cost as it could be got internally. Assume, for example, that the terms of trade are 0.5, equal to the British comparative cost. Then England would obtain, on the international market, with 4 units of labour (the cost of one unit of cloth) 0.5 units of wine, which would have required $0.5 \times 8 = 4$ units of labour if produced internally. In other words, by exchanging cloth for wine on the international market England would obtain exactly the same amount of wine obtainable internally (0.5 units of wine per one unit of cloth): there is, then, no reason for engaging in international trade. It can similarly be seen that, if the terms of trade were 0.6, there would be no reason for Portugal to engage in international trade at all. We leave it to the reader to check, as an exercise, that if the terms of trade were to fall outside the interval between the comparative costs (that is, in our example, if they were smaller than 0.5 or greater than 0.6) then, by engaging in international trade, one of the two countries would suffer a loss.

2.2 Alternative Graphic Representations

We can now show two simple diagrams to represent the theory of comparative costs. Let x denote (the amount of) cloth and y (the amount of) wine and consider country 1. With any given quantity of labour L_1 it is possible to obtain an amount of cloth

$$x = \frac{1}{a_1} L_1,$$

where a_1 (see Sect. 2.1) is the unit cost of producing cloth—a constant because of the assumption of fixed technical coefficients.

Likewise, with the same amount of labour it is possible to obtain

$$y = \frac{1}{b_1} L_1$$

of wine.

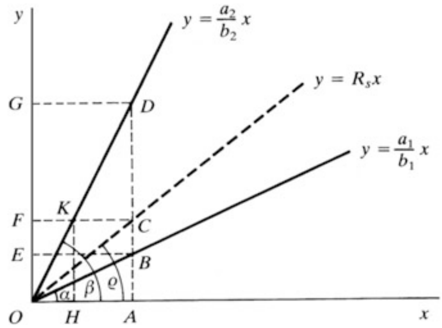


Fig. 2.1 Graphic representation of comparative costs

If we divide y by x we get

$$\frac{y}{x} = \frac{\frac{1}{b_1}L_1}{\frac{1}{a_1}L_1} = \frac{a_1}{b_1},$$

whence

$$y = \frac{a_1}{b_1}x. \tag{2.1}$$

We could have arrived at the same result by recalling that a_1/b_1 is the comparative cost, which (see Sect. 2.1) expresses the exchange ratio of the two commodities.

In an analogous way we get, for country 2, the relation

$$y = \frac{a_2}{b_2}x. \tag{2.2}$$

Equations (2.1) and (2.2) are represented in Fig. 2.1 as two straight lines starting from the origin. The elementary properties of straight lines tell us that $a_1/b_1 = \tan\alpha$ and $a_2/b_2 = \tan\beta$, that is, comparative costs are given by the slopes of the straight lines.

As the two lines do not coincide, there is a difference between the comparative costs: in fact, if these were equal ($a_1/b_1 = a_2/b_2$), the two lines would coincide. In this kind of diagram, therefore, the necessary condition for international trade is represented by the non-coincidence of the two lines.

Also the terms of trade can be represented as the slope of a straight line. In fact, if we denote these by R_s, then

$$\frac{y}{x} = R_s,$$

whence

$$y = R_s x, \tag{2.3}$$

which is a straight line through the origin with slope R_s. In Fig. 2.1 we have assumed that the sufficient condition for international trade is met, namely that line (2.3) falls strictly between lines (2.1) and (2.2); this amounts to saying that, having assumed $a_1/b_1 < a_2/b_2$, the inequality

$$\frac{a_1}{b_1} < R_s < \frac{a_2}{b_2} \tag{2.4}$$

holds. Of course, if $a_1/b_1 > a_2/b_2$, then the condition would be $a_1/b_1 > R_s > a_2/b_2$.

Inequality (2.4) is the same as

$$\tan\alpha < \tan\varrho < \tan\beta, \tag{2.5}$$

which has an obvious graphic interpretation. If this condition is satisfied, international trade will take place, and it will be profitable for country 1 to specialize in the production of x and for country 2 to specialize in the production of y. In terms of the diagram, in fact, the propositions so far examined are equivalent to saying (a) that the country whose line representing its comparative cost lies between the line representing the terms of trade and the horizontal axis will find it profitable to specialize in the production of (and in any case to export) the good measured on this axis, and (b) that the country whose comparative-cost line lies between the terms-of-trade line and the vertical axis will find it profitable to specialize in the production of (and in any case to export) the good measured on this axis.

To show this, let us suppose that, given the terms of trade R_s, a quantity OA of x is exchanged for OF of y. It is easy to see that the amount OA is exported by country 1 (and so imported by country 2) while the amount OF is exported by country 2 (and so imported by country 1). The proof is straightforward, and in the course of this proof we shall also have occasion to show a measure of the gains from trade accruing to each country. Now, at the domestic price ratio, country 1 would have obtained $OE = AB$ of y for OA of x, whilst it can obtain $OF = AC$ by way of international trade. It is therefore profitable for country 1 to engage in international trade following the pattern just described (that is, to export x and to import y). The gains from trade accruing to this country can be measured, for example, in terms of y: they are given by segment BC, namely by the additional quantity of y that country 1 obtains in exchange for the same quantity of x. Let us consider country 2 which, at the domestic price ratio, would have had to give up $OG = AD$ of y to obtain OA of x, whilst it has to give up $OF = AC$ by way of international trade. It is therefore profitable to country 2 to engage in international trade with the pattern just described, and the benefit accruing to this country, measured in terms of y, is given by segment DC.

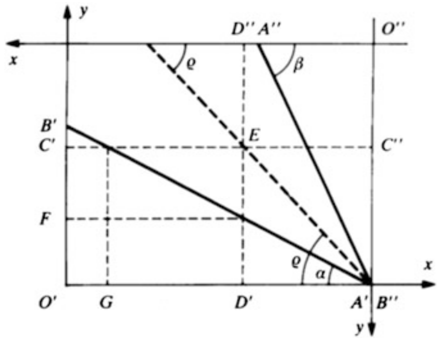


Fig. 2.2 Transformation curve and comparative costs

The gains from trade can also be measured in terms of x, but the measures are equivalent as can be shown by transforming them into each other by using the internal price ratio of the country concerned. For example, country 2 by trading OF of y on the international market obtains $OA = FC$ of x instead of $OH = FK$: the benefit in terms of x is, therefore, KC. But if we consider the right-angled triangle KCD we obtain $DC = KC \cdot \tan C\hat{K}D = KC \cdot \tan\beta$, where $\tan C\hat{K}D = \tan\beta =$ comparative cost or domestic exchange ratio of the two goods in country 2.

An alternative diagram of the theory of comparative costs is based on the concept of *transformation curve* (or *production-possibility frontier*) studied in microeconomic theory (see also below, Sect. 3.1). In our simplified model, in which there is only one factor of production and the technical coefficients are fixed, the transformation curve is linear (the general case will be treated in Sect. 3.1). It is in fact given, for country 1, by the equation

$$a_1x + b_1y = \overline{L}_1, \tag{2.6}$$

where $\overline{L}_1$ is the total amount of labour existing in country 1. Equation (2.6) is the equation of a monotonically decreasing straight line in the (x, y) plane, since we can write it as

$$y = -\frac{a_1}{b_1}x + \frac{\overline{L}_1}{b_1}. \tag{2.7}$$

In absolute value, the slope of this line equals the comparative cost in country 1. Comparative cost and marginal rate of transformation (or opportunity cost: see Sect. 3.1) are therefore one and the same thing.

In a similar way, we obtain the transformation curve of country 2. Consider then Fig. 2.2, where we have brought together the transformation curves of the two countries.

The line $A'B'$ is the transformation curve of country 1, i.e. the diagram of (2.7); in absolute value, $\tan\alpha$ equals the comparative cost of country 1. The line $A''B''$ is the transformation curve of country 2, rotated anticlockwise by 180° and placed so that point B'' coincides with point A'; it goes without saying that $O''B''$ and $O'B'$ are parallel. The absolute value of $\tan\beta$ equals the comparative cost in country 2.

Let us take an arbitrary admissible value of the terms of trade, say $\tan\varrho$, and assume that international trade occurs at point E, whose coordinates are the quantities exchanged. Country 1 specializes completely in the production of commodity x, of which it produces the amount $O'A'$; of this, a part is consumed domestically ($O'D'$), whilst the remaining part ($D'A'$) is exported in exchange for the quantity $O'C' = ED' = C''B''$ of commodity y. Note that, since the terms of trade are measured by $\tan\varrho$, and since (by considering the right-angled triangle $ED'A'$) we have $ED' = D'A' \cdot \tan\varrho$, it follows that by giving $D'A'$ of x, ED' of y can be obtained, and vice versa. This means that the trade balance is necessarily in equilibrium. In fact, balance-of-trade equilibrium, or value of exports = value of imports, requires

$$p_x D'A' = p_y ED'$$

or

$$\frac{p_x}{p_y} D'A' = ED', \tag{2.8}$$

which is indeed true, since commodities are exchanged at a relative price (p_x/p_y) given by the terms of trade, namely $p_x/p_y = \tan\varrho$.

Similarly, country 2 completely specializes in y and produces the amount $O''B''$ of this commodity, consuming $O''C''$ domestically and exporting $C''B''$ in exchange for $O''D'' = D'A'$ of commodity x. This result (complete specialization in both countries) is the normal outcome of trade in the Ricardian model. This may not be the outcome when one country (say country 1) is small with respect to the other, so that this country's production of x is not sufficient to fully satisfy, in addition to its own domestic demand, also the demand for this commodity by country 2. In such a case country 2 will not specialize completely in commodity y and will continue to produce both y and x.

As can be seen, point E lies *beyond* both transformation curves, and so it represents a basket of goods that neither country could have obtained in autarky. Consider, for example, country 1. In autarky, together with $O'D'$ of x this country could have obtained $O'F$ of y (less than the amount $O'C'$ that it obtains through international trade). The gains from trade accruing to this country can be measured, in terms of y, by $C'F$ (in terms of x they are measured by GD'). The gains from trade accruing to country 2 can be found in a similar way.

It is also obvious from the diagram that the closer the terms-of-trade line is to a country's transformation curve, the smaller that country's share of the gains; this share drops to zero when the terms-of-trade line coincides with that country's transformation curve (and all the gains go to the other country). This is an alternative way of showing the result already demonstrated in the previous treatment.

2.3 A Modern Interpretation in Terms of Optimization

The theory of comparative costs has been taken up again by modern scholars in terms of optimization. The general treatment will be given in Sects. 18.1 and 18.2; here we shall limit ourselves to a reformulation in these terms of the simple problem treated in the previous section.

We recall from that treatment that the benefits from international trade can be seen as an increase in the quantity of goods, and so in the real income (output) which can be obtained from the given amount of labour (by assumption, equal to the total amount available). It follows that the optimum can be interpreted as the maximization of real income given a certain input of labour; such an optimum, however, can be seen either from the point of view of the single country or from the point of view of the world as a whole (consisting, in our simple model, of two countries only).

2.3.1 *Maximization of Real Income in Each Country*

Let us begin by examining the optimum as the maximization of real national income in each country separately considered. Let p_x and p_y be the absolute prices (expressed in terms of some external unit of measurement, for example, gold). The generic value of monetary national income is $Y = p_x x + p_y y$, where x and y are the outputs of the two goods. If we divide Y by the price of either good, for example by p_y, we obtain real national income Y_R measured in terms of y.

Since, as we shall see presently, the relative price in the problem is given, the result would not change if we measured real income in terms of good x. On the other hand, since p_x and p_y are given, we could just as well consider Y, which would then be national income at constant prices. Thus there is no loss of generality by considering good x as the *numéraire* (unit of measurement).

We thus have the following two problems of constrained maximization:

$$\max Y_{1R} = \left(p_x/p_y\right) x_1 + y_1 \quad \text{sub } a_1 x_1 + b_1 y_1 \le \overline{L}_1, \quad x_1 \ge 0, \quad y_1 \ge 0, \tag{2.9}$$

and

$$\max Y_{2R} = \left(p_x/p_y\right) x_2 + y_2 \quad \text{sub } a_2 x_2 + b_2 y_2 \le \overline{L}_2, \quad x_2 \ge 0, \quad y_2 \ge 0, \tag{2.10}$$

where for each country the constraints are the respective transformation curve (the $\le$ sign means that, in principle, all points internal to the curve are also admissible) and the non-negativity of the outputs.

The exchange ratio or relative price of the two goods, p_x/p_y, is to be taken as given, determined on the international market (in the same way in which, in Sect. 2.2, we considered the terms of trade as exogenously given). In fact, owing to

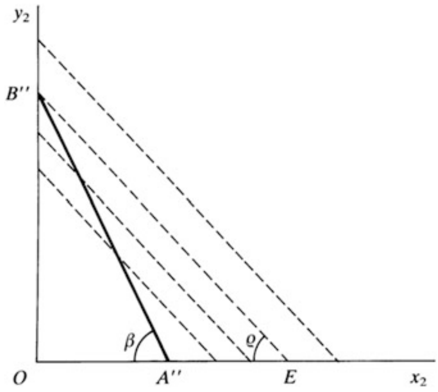


Fig. 2.3 Transformation curve and maximization of real income

the assumptions of perfect competition and of absence of transport costs, with free trade the domestic price ratio must necessarily be equal to the international terms of trade.

The data are then completed by assumption (2.4) of Sect. 2.2.

With these premises, problems (2.9) and (2.10)—which are linear programming problems of the simplest sort—can find an easy graphic solution. In fact, the function to be maximized can be represented by a family of parallel straight lines with a negative slope, each of which represents the locus of all combinations of x and y yielding the same real income (a *budget line* or, as we prefer to call it, an *isoincome line*: this terminology has the same derivation as isocost, isoquant, etc.); furthermore, the farther any such line is from the origin, the higher the corresponding real income. As a matter of fact, from the equation $Y_R = \left(p_x/p_y\right)x + y$ we get

$$y = -\left(p_x/p_y\right)x + Y_R, \tag{2.11}$$

which, if we consider Y_R as a parameter, defines a family of straight lines with the properties stated.

The graphic solution of our problem then consists in finding the highest isoincome attainable without going beyond the transformation curve of the country concerned, and remaining in the first quadrant (non-negativity constraints). If we consider, for example, country 2, we can draw Fig. 2.3, where $\tan\varrho$ = international relative price (terms of trade) and $\tan\beta$ = *marginal rate of transformation* = a_2/b_2; given the assumptions, $\tan\varrho < \tan\beta$.

It can easily be seen that, given the constraint, the highest isoincome attainable is $B''E$; consequently, the constrained-optimum point is B''. Country 2 thus maximizes its real national income by specializing entirely in the production of good y.

In a similar way it can be shown that country 1 maximizes its real national income by specializing entirely in the production of good x.

The reader will remember that complete specialization is indeed the outcome of the theory of comparative costs. This theory therefore implies the maximization of the real national income of each country separately considered.

2.3.2 Maximization of Real World Income

The same problem of maximizing real income can be formulated from the point of view of the world as a whole. Real world income in terms of good y is

$$Y_{RM} = (p_x/p_y)(x_1 + x_2) + (y_1 + y_2) = (p_x/p_y)x_M + y_M, \tag{2.12}$$

where x_M and y_M are the quantities of the two goods globally produced in our two-country world. In order to proceed in the same way as before, it is necessary to determine the world transformation curve.

The world transformation curve is defined as that curve which—for the world as a whole and within the limits of total existing resources—gives the maximum producible quantity of y for any given quantity of x to be produced, and vice versa. This transformation curve must, therefore, be derived from a maximization procedure. Let us note that, in general, any transformation curve is the outcome of a maximization procedure and is, therefore, a locus of points sharing the property of efficiency in production. In the case of a single country and fixed technical coefficients the procedure is trivial: given for example the quantity x_1, the labour required to produce it is $x_1 a_1$. As the total amount of labour is $\overline{L}_1$, we are left with $\overline{L}_1 - x_1 a_1$ to produce y, the maximum output of which is $y_1 = (\overline{L}_1 - x_1 a_1)/b_1$, which is Eq. (2.7) already examined in Sect. 2.2.

Also at the world level the derivation of the world transformation curve is a fairly simple matter, thanks to the assumption of fixed technical coefficients.

With reference to Fig. 2.4, let us begin by determining the extreme points (intercepts): these are A and B. Segment OA represents the maximum possible output of x, obtained when all world resources are employed to produce this good. It is obvious that this segment is the sum of segments $O'A'$ and $O''A''$ in Fig. 2.2; algebraically we have $OA = \overline{L}_1/a_1 + \overline{L}_2/a_2$. Similarly the maximum world output of y turns out to be $OB = O'B' + O''B'' = \overline{L}_1/b_1 + \overline{L}_2/b_2$.

To find the other points of the world transformation curve, let us suppose we start from point A and forgo one unit of good x: a certain amount of labour will then become available for employment in the production of good y. As we are reasoning at world level, we must determine—on the basis of technology—which country it is better to perform these operations in, so as to optimize the result, that is to obtain the maximum amount of y_M for the one unit of x_M we have forgone.

Now, if we forgo one unit of x in country 1, we free an amount of labour equal to a_1 which, if employed in that country to produce y, will allow an increase in the output of y equal to a_1/b_1 (that is, obviously, country 1's marginal rate of transformation). If we carry out the same operations in country 2, we get a_2/b_2 more

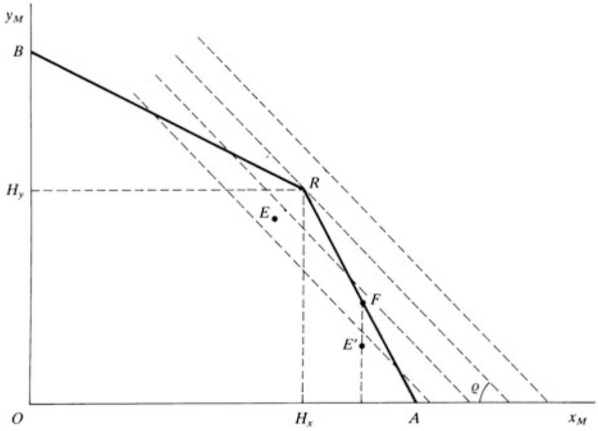


Fig. 2.4 World transformation curve and maximization of real world income

of y for one unit less of x. As we have assumed (see above) that $a_1/b_1 < a_2/b_2$, the operations under consideration are better carried out in country 2, and since the marginal rate of transformation is constant, this continues to hold for further decreases in x_M.

Therefore, starting from *A*, the best course of action is that country 1 continues to produce only good x, whilst the world output of y will be maximized by "transforming" x into y in country 2, according to this country's transformation curve.

We shall therefore move along segment *AR*, whose slope equals that of country 2's transformation curve: actually, this segment is nothing more than the transformation curve of country 2 drawn with reference to the auxiliary origin H_x.

When it arrives at point *R*, country 2 will produce exclusively good y, whilst country 1 will still be entirely specialized in the production of good x: this point corresponds to the Ricardian situation and is therefore called the *Ricardo point* by Dorfman, Samuelson, and Solow (1958, p. 35). From this point, further reductions in x_M and increases in y_M can only take place in country 1, along its transformation curve (this is *RB*, with reference to the auxiliary origin H_y), whilst country 2 will produce exclusively good y, as shown above.

The world transformation curve is thus the kinked curve *ARB*. The reader might like to check that the same curve would be obtained by starting from point *B*.

If we now draw the isoincome lines representing real world income as defined in Eq. (2.12), we obtain a family of straight lines with the usual properties. The highest isoincome attainable is the one passing through the Ricardo point: it is therefore demonstrated that the solution found by the theory of comparative costs implies the maximization of real world income.

The above treatment also enables one to give an answer to the objections of Pareto (1906) and successive authors to the theory of comparative costs. According to Pareto, it is possible for international trade to give rise to a worse situation than the autarkic one, for example when the quantity of a good increases but the quantity of the other decreases. If we interpret this criticism in terms of Fig. 2.4, we see that the coordinates of point R represent greater quantities of both goods with respect to, say, point E (inside the transformation curve), but not with respect to all internal points. At point E', for example, the quantity of x is greater, but that of y is smaller, than at point R. In a case like this it is not possible, according to Pareto, to establish whether one point is preferable to the other without introducing utility, and when this is done, it may well be that point E' will yield a greater utility than point R. It is however possible to rebut Pareto's criticism without having to introduce assumptions on the utility function. In fact, the efficiency properties of the world transformation curve allow us to state that, for any internal point, it is possible to find a point on the frontier which denotes a better situation (in the example above, the latter is point F, where the quantity of x is the same as, but the quantity of y is greater than, at point E'). Therefore international trade will always be preferable to autarky provided that it gives rise to points on the world transformation curve; this will indeed be the case for any admissible terms of trade.

2.4 Generalizations

In Sects. 2.1–2.3 we have considered the simple case of international trade concerning two goods and two countries. In this section we first examine the extension of the Ricardian theory to n countries trading two goods and then the general case of n countries and m goods. Further treatment of the classical theory is contained in Allen (1965), Bhagwati et al. (1998), Chacholiades (1978), Edgeworth (1894), Graham (1923), Haberler (1936), Hartwick (1979), Jones (1961), McKenzie (1954a, 1954b, 1955), Ricardo (1817), Whitin (1953). In the Appendix, Sect. 18.3, we study the generalization to a continuum of goods. Before moving to these generalizations we mention other advancements in research concerning the sources of the differences in comparative costs between countries. One traditional source, probably the most direct, is the technology in the strict sense of the engineering aspects of the production process. But other sources are definitely to be considered. As a matter of fact, anything that contributes to determining the unit cost of production is a potential source of comparative cost/advantage. Among such sources one may list the quality of institutions, of commercial laws, of infrastructures, the features of the labour market, the effectiveness of law enforcement and cultural traits of economic agents. For developments in these directions see, e.g., Cuñat and Melitz (2007), Levchenko (2007), Nunn (2007), Costinot (2009), Belloc and Bowles (2013), and Belloc (2006) for a review of the role of institutions.

2.4.1 Two Goods and n Countries

A necessary condition for international trade to take place when there are n countries is that at least two of these have different comparative costs, for it is self-evident that, if all had the same comparative cost, there would be no incentive to engage in international trade, exactly as in the two-country case. Once this condition is satisfied, it is not very relevant whether all countries have different comparative costs or whether there exist subsets of countries with the same comparative cost; to simplify the treatment, we shall adopt the former assumption. No loss of generality is involved in assuming that the countries can be ordered in such a way that

$$\frac{a_1}{b_1} < \frac{a_2}{b_2} < \ldots < \frac{a_n}{b_n}. \tag{2.13}$$

Now, once the necessary condition is met, the sufficient condition is that the terms of trade are strictly included between the two extreme comparative costs,

$$\frac{a_1}{b_1} < R_s < \frac{a_n}{b_n}. \tag{2.14}$$

A new complication should be noted: even if (2.14) is satisfied, R_s may happen to coincide with some intermediate comparative cost. In this case, the country concerned will not participate in international trade, which will involve the remaining $n-1$ countries. In any case we shall find a certain number of countries with a comparative cost lower than R_s while the remaining ones will have a comparative cost higher than R_s, namely

$$\frac{a_1}{b_1} < \ldots < \frac{a_i}{b_i} \leq R_s \leq \frac{a_{i+1}}{b_{i+1}} < \ldots < \frac{a_n}{b_n}, \tag{2.15}$$

where $i = 2, 3, \ldots, n-1$ denotes any country other than the first and the last. If the equality sign holds in the weak inequality $a_i/b_i \leq R_s$, then country i will not engage in international trade.

Once condition (2.15) is satisfied, international trade will take place between the countries with a comparative cost lower than R_s, on the one hand, and the countries with a comparative cost higher than R_s, on the other. The former group of countries will specialize entirely in the production of x, whilst the latter will specialize entirely in the production of y: therefore, x will be exported by the former to the latter group, and vice versa for y.

This result can be given a simple graphic interpretation in terms of the world transformation curve. When there are n countries, a world transformation curve can be constructed by way of the same procedures explained in the case of two countries: starting, for example, from the point where the world produces exclusively good x, the best course of action will be to "transform" good x into good y along country

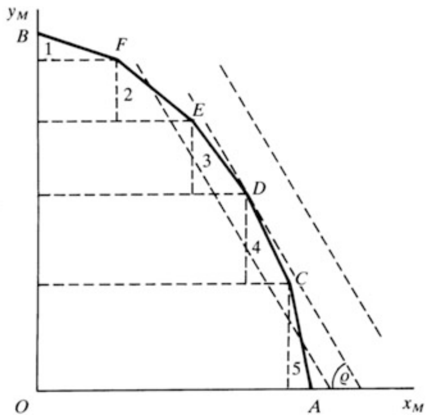


Fig. 2.5 The world transformation curve with n countries

n's transformation curve, then along country $(n-1)$'s and so on (the reasoning is altogether similar to that employed in Sect. 2.3).

If we assume, for example, that there are five countries, we get Fig. 2.5, where the numbers denote the transformation curves of the various countries stacked one on top of the other in the usual manner. In the diagram, given for example the terms of trade measured by $\tan\varrho$, the maximization of real world income $Y_{RM} = (p_x/p_y)(x_1 + x_2 + \ldots + x_5) + (y_1 + y_2 + \ldots + y_5) = (p_x/p_y)x_M + y_M$ is obtained at point D, so that countries 1,2,3 specialize entirely in the production of good x, and countries 4,5 in the production of good y. It is in fact easy to see that $a_1/b_1 < a_2/b_2 < a_3/b_3 < R_s < a_4/b_4 < a_5/b_5$.

In the particular case in which $a_i/b_i = R_s$, the isorevenue line will be tangent to a facet of the polygonal curve $ACDEFB$ (the facet corresponding to country i's transformation curve) and the solution will be indeterminate. In such a case, as we said, country i will not participate in international trade and will produce the same output combination as before, when no international trade existed: this will enable us to determine the precise point on the facet under consideration. The result is that country i will not necessarily specialize, whilst all the remaining countries will, as explained above.

2.4.2 m Goods and n Countries

Let us begin by examining the case of m goods and two countries. For this purpose, it is expedient to adopt the alternative definition of comparative cost (see Sect. 2.1), namely the ratio between the absolute unit costs of the same good in the two countries. Without loss of generality, we can order the comparative costs in an increasing manner (namely in order of diminishing country 1 comparative advantage), that is

$$\frac{a_2}{a_1} > \frac{b_2}{b_1} > \frac{c_2}{c_1} > \ldots > \frac{m_2}{m_1}. \tag{2.16}$$

For motives that will become clear further on, it is expedient to introduce the ratio between the two countries' unit money wage rates, both expressed in a common monetary unit, say gold (as the exchange rate is assumed to be perfectly rigid, it can be set at one without loss of generality). Let this ratio be $\omega = w_1/w_2$.

It can then be shown that the condition for international trade to take place is that ω is strictly included between the two extreme comparative costs, i.e.

$$\frac{a_2}{a_1} > \omega > \frac{m_2}{m_1}. \tag{2.17}$$

It can also be shown that all goods with a comparative cost lower than ω will be exported by country 2, which will specialize entirely in their production, whilst all goods having a comparative cost higher than ω will be exported by country 1, which will specialize entirely in their production. In the particular case in which there is a good having a comparative cost exactly equal to ω, this good will, in general, be produced by both countries and will not be internationally traded.

To prove these statements, we begin by observing that, given the money wage rates w_1 and w_2, the (monetary) unit cost of production and so the (monetary) price of the various goods in the two countries, before international trade is opened, will be

$$\begin{array}{ll} p_{A_1} = w_1 a_1, & p_{A_2} = w_2 a_2, \\ p_{B_1} = w_1 b_1, & p_{B_2} = w_2 b_2, \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ p_{M_1} = w_1 m_1, & p_{M_2} = w_2 m_2. \end{array} \tag{2.18}$$

Now, given the assumptions of free trade, perfect competition and no transport costs, each good will be bought where it costs least. Therefore if—for example—we have $p_{C_1} < p_{C_2}$, country 2 will buy good C from country 1 (which will become an exporter of this good) instead of producing it internally, and vice versa. Furthermore, since, in the pure of theory of international trade, imports must be paid for by exports, each country must be able to export some good. It is now obvious that, if it were

$$\omega \geq \frac{a_2}{a_1}, \tag{2.19}$$

country 2 would produce all goods at a lower price than country 1, which could not then engage in international trade, being unable to export anything. In fact, since $\omega = w_1/w_2$ by definition, from Eq. (2.19) we get

$$1 \geq \frac{a_2 w_2}{a_1 w_1}, \tag{2.20}$$

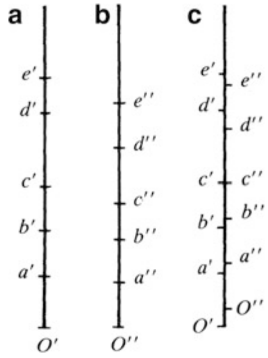


Fig. 2.6 Exchange of more than two goods between two countries

whence, given Eqs. (2.18),

$$p_{A_1} \geq p_{A_2}, \tag{2.21}$$

so that country 1 produces good A at a price higher than (or at most equal to) country 2. Now, account being taken of Eq. (2.16), if (2.19) holds, it will also be true that ω is higher than all other comparative costs and so, by similar reasoning, that the price of $B,C,\ldots$ is higher in country 1. There is, therefore, no scope for international trade.

In a similar way it can be proved that if $\omega \leq m_2/m_1$, country 2 produces good M at a price higher than (or at most equal to) country 1 etc., so that, also in this case, there can be no international trade.

If, on the contrary, inequality (2.17) holds, by considering the left-hand side of it we get

$$p_{A_1} < p_{A_2}, \tag{2.22}$$

whilst by considering the right-hand side we have

$$p_{M_2} < p_{M_1}, \tag{2.23}$$

so that there exists at least one good (A) which country 1 can export and at least one good (M) exportable by country 2.

If we now indicate by the subscript Ω a generic good and by θ the corresponding technical coefficient, it can easily be seen that $\theta_2/\theta_1 < \omega$ is equivalent to $p_{\Omega_2} < p_{\Omega_1}$ (good Ω will be exported by country 2), whilst $\theta_2/\theta_1 > \omega$ is equivalent to $p_{\Omega_2} > p_{\Omega_1}$ (good Ω will be exported by country 1). This demonstrates the second part of the proposition.

In conclusion, given ω, we can divide all goods into two groups: one comprising the goods exported from country 1 to country 2 (these are the goods having a comparative cost lower than ω) and the other comprising the goods exported from country 2 to country 1 (those with a comparative cost higher than ω).

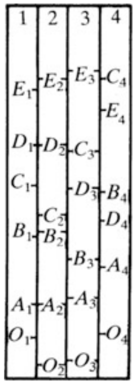


Fig. 2.7 Exchange of more than two goods among more than two countries

This treatment is amenable to a simple graphic representation, provided by Edgeworth. In Fig. 2.6a start from origin O' and draw segments representing the *logarithms* of the technical coefficients (unit costs in terms of labour) of the various goods in country 1, that is, $O'a' = \log a_1$, $O'b' = \log b_1$ and so on up to good E (we have considered only five goods, but they can be of any number). Similarly, in Fig. 2.6b draw segments representing the logarithms of the technical coefficients in country 2 ($O''a'' = \log a_2$, etc.).

Then put the two diagrams together in Fig. 2.6c *in such a way that the distance between the two origins represents the logarithm of the parameter* ω, that is $O'O'' = \log \omega$, stipulating that O'' will be above O' if $\omega_2 > \omega_1$ and so $\omega > 1$ (whence $\log \omega > 0$), and below it in the opposite case. Once the figure has been drawn, we can immediately check whether (2.17) is met and determine the point where the succession of goods is divided between those exported by country 1 and those exported by country 2. In fact, if we consider the inequality $a_1/a_2 < \omega$ and take the logarithms, we get

$$\log a_1 < \log a_2 + \log \omega, \tag{2.24}$$

the graphic counterpart of which is

$$O'a' < O''a'' + O'O'', \tag{2.25}$$

which is certainly satisfied since a' is below a''. It follows that the relative position of the various points in Fig. 2.6c will immediately tell us the division of the goods in the two groups: good A and good B will be exported by country 1; good C (for which $c_1/c_2 = \omega$) will not be traded internationally; goods D and E will be exported by country 2.

Edgeworth's ingenious diagram was extended by Viner to any number of countries, thus enabling us to examine the exchange of n goods among m countries graphically. In Fig. 2.7, adapted from Viner (1937, p. 465), we consider five commodities and four countries; the diagram is drawn accordingto the same principles

Table 2.3 Pattern of trade of five goods among four countries

	Country 1	Country 2	Country 3	Country 4
Exports	A	C	B	D,E
Imports	B,C,D,E	B,D,E	A,C,D,E	A,B,C

as Fig. 2.6 and the distances between the origins represent the relative money wage rates of the various countries. From an inspection of the figure the pattern of trade immediately results (see Table 2.3). Note, finally, that country 2 may either export, import, or not trade in commodity *A* as this commodity is on the margin of trade for that country.

2.5 The Problem of the Determination of the Terms of Trade

In the previous treatment we have determined the limits within which the terms of trade must lie, but—as the reader may have noticed—we have not specified how, and at what value, the terms of trade themselves are determined within these limits.

As a matter of fact, it is a generally accepted opinion that the Ricardian theory of comparative costs as such is incapable of determining the terms of trade and only determines the limits within which they must lie. This would constitute a serious limitation to this theory seen as a model aimed at the explanation of international trade, for any such model ought to explain not only the causes and pattern of trade, but also the terms of trade. The limitation, on the contrary, would be almost irrelevant if one believes that the Ricardian theory must be seen from the normative, rather than the positive, point of view. According to Bhagwati (1964, p. 4), for example, the Ricardian theory is more plausibly seen "as a highly simplified model which was intended to be, and served as, an eminently successful instrument for demonstrating the welfare proposition that trade is beneficial" rather than "as a serious attempt at isolating the crucial variables which can be used to 'explain' the pattern of trade". In our opinion, both elements are present in the theory under consideration, and we have treated it in this sense in the present chapter.

In order to solve the problem of the determination of the terms of trade—the accepted opinion goes on—it is necessary to introduce the demand side in addition to the productive side focused on by the original formulation of the theory of comparative costs.

The first precise reasoning in this sense was J.S. Mill's equation of international demand, according to which the terms of trade are determined so as to equate the value of exports and the value of imports. As Mill (1848, chap. XVIII, sect. 4, pp. 592–593) writes,

> The law which we have now illustrated, may be appropriately named, the Equation of International Demand. It may be concisely stated as follows. The produce of a country exchanges for the produce of other countries, at such values as are required in order that the whole of her exports may exactly pay for the whole of her imports. This law of International

> Values is but an extension of the more general law of Value, which we called the Equation of Supply and Demand. We have seen that the value of a commodity always so adjusts itself as to bring the demand to the exact level of the supply. But all trade, either between nations or individuals, is an interchange of commodities, in which the things that they respectively have to sell constitute also their means of purchase: the supply brought by the one constitutes his demand for what is brought by the other. So that supply and demand are but another expression for reciprocal demand: and to say that value will adjust itself so as to equalize demand with supply, is in fact to say that it will adjust itself so as to equalize the demand on one side with the demand on the other.

We find here, in a nutshell, the elements that were to be taken up again and further developed by Alfred Marshall in his theory of international reciprocal demand curves, leading to the neoclassical theory of international trade, that will be treated in the next chapter. In fact, from the point of view of the history of economic thought, J.S. Mill cannot be considered entirely as a member of the classical school, as in his writings many elements are present which later were to characterize the neoclassical school.

Actually, there is no dearth of attempts (for surveys of the earlier literature see Viner, 1937; Chipman, 1965a; Takayama, 1972, chap. 5) at introducing demand in the theory of comparative costs, leaving all its other hypotheses unaltered. We shall examine in the Appendix (see Sect. 18.3) an elaboration of the Ricardian model (with a continuum of goods and the presence of demand functions) due to Dornbusch, Fischer, and Samuelson (1977) .

We must at this point ask ourselves what is the validity of the received opinion. It obviously leads to considering the classical theory of comparative costs, enriched by the introduction of demand functions, as a particular case of the neoclassical theory, which would occur when one assumed fixed-coefficient production functions. This has been challenged by those who maintain that such a view would misrepresent the classical theory, whose vision is completely different from the neoclassical one.

In particular, Negishi (1982) maintains that, contrary to the received opinion, the *original* Ricardian theory is perfectly able to determine the terms of trade without having recourse to demand factors, but by using solely cost-price relations. This would be possible, according to Negishi (p. 200), by making use of “the classical theory of wages, the rate of profit, and the role of exporters and importers, which have been missing in the standard interpretation of the classical theory of international trade”. For an examination of this interesting thesis, we refer the reader to the Appendix, Sect. 18.4.

References

Allen, W. R. (Ed.). (1965). *International trade theory: Hume to Ohlin*.

Belloc, M. (2006). Institutions and international trade: A reconsideration of comparative advantage.

Belloc, M., & Bowles, S. (2013). Cultural-institutional persistence under autarchy, international trade, and factor mobility.

Bhagwati, J. N. (1964). The pure theory of international trade: A survey.

Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (chaps. 2, 3, 4).
Chacholiades, M. (1978). *International trade theory and policy* (Pt. II).
Chipman, J. S. (1965a). A survey of the theory of international trade, Part 1, the classical theory.
Costinot, A. (2009). On the origins of comparative advantage.
Cuñat, A., & Melitz, M. (2007). Volatility, labour market flexibility, and the pattern of comparative advantage.
Dorfman, R., Samuelson, P. A., & Solow, R. M. (1958). *Linear programming and economic Analysis.*
Dornbusch, R., Fischer, S., & Samuelson, P. A. (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods.
Edgeworth, F. Y. (1894). The theory of international values, I, II, III, pp. 53 ff. of the 1925 reprint.
Graham, F. D. (1923). The theory of international values reexamined.
Haberler, G. (1936). *The theory of international trade* (chap. X).
Hartwick, J. M. (1979). Distribution of world income in the Ricardo-Mill model of international trade.
Jones, R. W. (1961). Comparative advantage and the theory of tariffs: A multi-country multi-commodity model.
Levchenko, A. A. (2007). Institutional quality and international trade.
McKenzie, L. (1954a). On equilibrium in Graham's model of world trade and other competitive systems.
McKenzie, L. (1954b). Specialization and efficiency in world production.
McKenzie, L. (1955). Specialization in production and the production possibility locus.
Mill, J. S. (1848). *Principles of political economy*, (chaps. XVII, XVIII) (pages refer to the 1961 reprint).
Negishi, T. (1982). The labour theory of value in the Ricardian theory of international trade.
Nunn, N. (2007). Relation-specificity, incomplete contracts and the pattern of trade.
Pareto, V. (1906). *Manuale d'economia politica* (chap. VII, sects. 43ff).
Ricardo, D. (1817). *On the principles of political economy and taxation* (chap. VII).
Takayama, A. (1972). *International trade: An approach to the theory* (Pt. II).
Torrens, R. (1815). *An essay on the external corn trade.*
Viner, J. (1937). *Studies in the theory of international trade* (chap. VIII).
Whitin, T. M. (1953). Classical theory, Graham's theory, and linear programming.

Chapter 3
The Neoclassical Theory of International Trade

Before discussing the neoclassical model of international trade (Sect. 3.3 and following), it is advisable to recall from microeconomics some widely-used diagrammatic tools (Sect. 3.2) and to show how the general equilibrium of production and consumption is determined in a simple closed economy (Sect. 3.2), where two final goods (A and B) are produced by the full employment of two *primary* factors of production (K and L). The problems deriving from the presence of *produced* factors of production will be tackled in Sects. 6.4, 6.4.1 and 14.1.

The given data are:

(a) The total amounts of the two factors existing in the economy;
(b) The distribution of these among the members of the economy, namely the amounts of K and L owned by each member;
(c) The tastes of consumers;
(d) The state of technology, represented by well-behaved aggregate production functions (a "well-behaved" production function shows constant returns to scale and has positive but decreasing marginal productivities: see Sect. 19.1.3).

Perfect competition obtains in all markets (commodities and factors).

3.1 The Transformation Curve and the Box Diagram

The tools that we wish to recall are the Haberler-Viner-Lerner-Leontief *product transformation curve* (otherwise known as the *production-possibility curve* or *production-possibility frontier*) and the Edgeworth-Bowley *box diagram* (originally intended to derive the contract curve between two consumers and applied to production problems by Lerner and Stolper-Samuelson); for a detailed analysis of "who was the first" and references, see Savosnick (1958).

The product-transformation curve (henceforth called the *transformation curve*) represents the maximum amount of one commodity obtainable for any given amount

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_3,

of the other. This requires that the given fixed amounts of productive factors are optimally allocated between the two commodities in accordance with certain marginal productivity conditions which are easily found by using the *box diagram*.

3.1.1 The Box Diagram

In Fig. 3.1, the length of the sides of the box represents the total amounts of the two productive factors existing in the economy, respectively $O_AG = O_BH$ of labour and $O_AH = O_BG$ of capital.

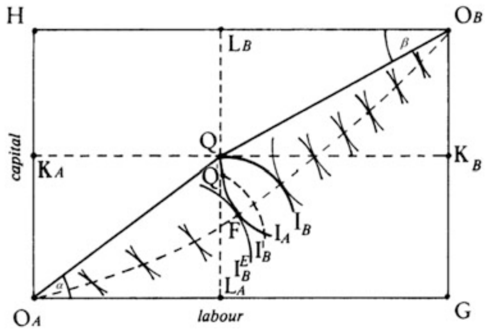
The isoquants concerning the production of commodity A are drawn with reference to the origin O_A, and the isoquants concerning the production of commodity B are drawn with reference to the origin O_B and so appear upside down. In fact, the box can be considered as obtained by first drawing the isoquant maps for the two commodities in the usual way—with the proviso that the lengths of the axes are equal—and then turning one of the two upside down so that the extremes of the axes (points H and G) coincide. Both isoquant maps have the usual properties.

Let us now find the condition of efficiency in production, also called Pareto optimality in the producing sectors. By efficiency in the producing sectors we mean a situation in which—on the assumption of full employment of all factors of production—these factors are allocated between the two commodities in such a way that, given the output of one commodity, the output of the other is maximized. An equivalent definition is that it is not possible, by reallocating the given fixed amounts of productive factors, to increase the output of one commodity without decreasing that of the other. It is clear that if instead it is possible, by means of such a reallocation, to increase the output of one commodity while keeping the output of the other constant, then the situation is inefficient. It can be proved graphically that the condition for efficiency is that the A isoquants and the B isoquants are tangent (for simplicity's sake we neglect possible corner solutions), namely that the marginal rates of technical substitution (MRTS) are equal in the two productive sectors.

For this purpose, consider for example point Q in Fig. 3.1. This point lies at the intersection of isoquant I_A (concerning the production of commodity A) with isoquant I_B (concerning the production of commodity B). The allocation of the productive factors can be read by drawing the coordinates of Q on the sides of the box, which gives O_AL_A of labour to (the production of) commodity A and $L_AG = O_BL_B$ to commodity B, and similarly O_AK_A and $K_AH = O_BK_B$ of capital to commodities A and B respectively. If we connect point Q to the origins by means of two straight line segments, we can read the factor intensities as the slopes of these segments: for example, $\tan\alpha = O_AK_A/O_AL_A$ is the capital/labour ratio in the A sector and $\tan\beta = O_BK_B/O_BL_B$ is the capital/labour ratio in the B sector.

Point Q is not efficient: in fact, by reallocating the productive factors it is possible to move for example to Q' on I'_B while still remaining on I_A; point Q' gives

Fig. 3.1 The box diagram and the efficiency locus



a greater output of commodity B (isoquant I'_B being farther from the origin O_B than I_B, represents a greater output). Continuing in this manner we arrive at the point of tangency F, which corresponds to the highest B isoquant $\left(I_B^E\right)$ achievable, given I_A, namely at F we get the maximum output of B given the output of A.

Further increases in B can be obtained only at the expense of a reduction in A, therefore F is an efficient point. The (optimal) allocation of the productive factors between the two sectors and the corresponding factor intensities can be read as shown above with reference to the non-optimal point Q.

The locus of all such points of tangency is the efficiency locus we are looking for; it is also (improperly) called *contract curve* (this was the original Edgeworth-Bowley denomination, but with reference to consumers' exchange).

3.1.2 *The Transformation Curve and Its Properties*

The passage from the efficiency locus to the transformation curve is simple: it is sufficient to transfer the indexes attached to each couple of tangent isoquants (these indexes are numbers representing quantities of the two commodities) to the coordinate axes in the *(A,B)* plane. In this way (Fig. 3.2) we obtain a diagram showing the maximum amount of B obtainable for any given amount of A, namely the transformation curve. Since the maximization procedure is perfectly symmetric, the efficiency locus and the transformation curve are the same if we maximize the output of A for any given amount of B. An alternative procedure for deriving the transformation curve from the box diagram is represented in Fig. 3.3 (Savosnick, 1958), which is similar to Fig. 3.1 except that now the right-hand vertical side of the box is used to measure the output of commodity A and the lower horizontal side is used to measure the output of commodity B. For simplicity's sake we assume constant returns to scale in both commodities. In this case, as we know from the properties of production functions homogeneous of the first degree (see Sect. 19.1.3), an isoquant which intersects a straight line through the origin twice as far away as another isoquant will represent twice as large an output. If we take

Fig. 3.2 The transformation curve

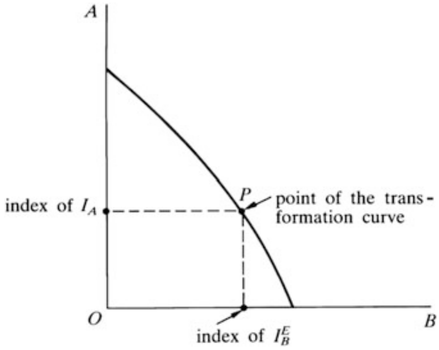
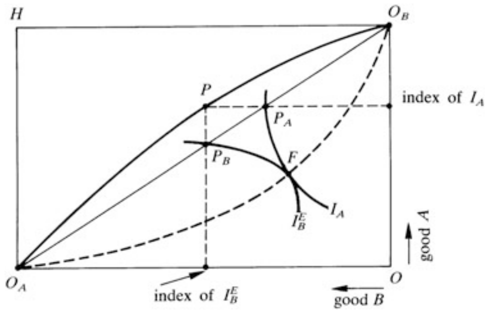


Fig. 3.3 The transformation curve derived from the box diagram



the diagonal of the box (straight line $O_A O_B$) as such a straight line, we can use it to project the outputs on the output axes; these projections will correspond exactly to the relationship just mentioned, and each output axis will have a uniform scale.

Consider for example point F. The isoquant I_A intersects the diagonal at point P_A, whose projection on the A axis gives the index of I_A; similarly, isoquant I_B^E intersects the diagonal at point P_B, whose projection on the B axis gives the index of I_B^E. Therefore point P, which has these projections as coordinates, is a point of the transformation curve. In this way we obtain the transformation curve $O_A P O_B$ which is the same—apart from scale factors and position—as the curve in Fig. 3.2. This construction also illustrates the one-to-one correspondence between points on the efficiency locus and the transformation curve: to every point on this curve representing an output combination there corresponds a point on the efficiency locus representing an input combination, and vice versa.

With both production functions exhibiting constant returns to scale, the efficiency locus must lie on one side of the diagonal of the box diagram and can never cross it, although locus and diagonal may coincide. In fact, when a point of the efficiency locus lies on the diagonal, then the whole efficiency locus coincides with the diagonal itself. This follows from the fact that with constant returns to scale the marginal rate of technical substitution is constant along a straight line through the origin. Therefore, if the MRTS of an A isoquant is equal to the MRTS of a B isoquant at a point on the diagonal of the box, then these MRTS remain the same

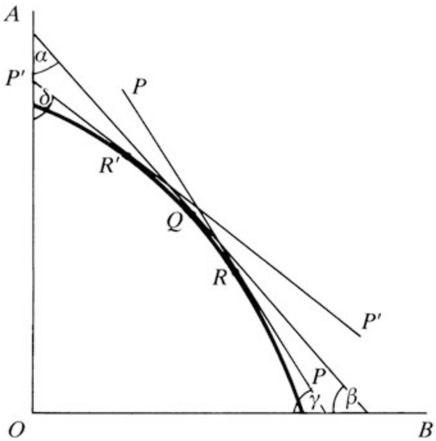


Fig. 3.4 Marginal rate of transformation, opportunity cost, and relative prices

along the diagonal, and if they are equal at one point they must be equal everywhere (in such a special case the capital-labour ratio is the same for both commodities). On the other hand, if the MRTS of an A isoquant and the MRTS of a B isoquant are different at some point on the diagonal, then they must be different at all other points on the diagonal.

The transformation curve can in principle be either convex or concave or both, but with constant returns to scale in both sectors, it will always be strictly concave to the origin if we exclude the particular case of identical capital-labour ratios, just dealt with (in which case it will be linear). This follows immediately from the graphic construction given in Fig. 3.3 and from the property that the efficiency locus must lie on one side of the diagonal of the box diagram (note that if it were all above the diagonal instead than below it, we would measure the output of commodity A on the left-hand vertical side of the box and the output of commodity B on the upper horizontal side, and would obtain a transformation curve concave to the origin now given by point H). Other simple graphic proofs can be found, for example, in Chacholiades (1978, pp. 107–109) and in Findlay (1970, pp. 26–29). In Sect. 19.1.2 we give a general mathematical proof in which we also consider the properties of transformation curves derived from production functions which do not possess constant returns to scale.

The (absolute value of the) slope of the transformation curve (for example tan β in Fig. 3.4) is called the *marginal rate of transformation* or (marginal) *opportunity cost* of B in terms of A, namely the amount of A that the economy has to give up to obtain an additional unit of B. It should be noted that this notion of opportunity cost has general validity, independently of the theoretical frame of reference. For example, in the Ricardian theory treated in Chap. 2, it is possible to identify the opportunity cost with the comparative cost.

It goes without saying that the opportunity cost of A in terms of B is measured by the (absolute value of the) slope of the transformation curve with reference to the A axis, namely $\tan\alpha$ in Fig. 3.4. Note that, since the transformation curve is derived from an optimizing procedure, the amount of A (or of B) given up is the *minimum* possible under the given technical knowledge. The concavity of the transformation curve implies that its slope increases as we move along it to the right, i.e. the opportunity cost of B increases as more of it is produced.

A fundamental proposition is that, under competitive conditions, the economy will always operate on the transformation curve, at a point where the marginal rate of transformation equals the price ratio or relative price of the two commodities p_B/p_A.

To prove the first part of the proposition it is sufficient to show that pure competition will bring producers onto the efficiency locus. Cost minimization requires that the MRTS in each sector is equated to the factor-price ratio, and since with perfect factor mobility the price of a factor is the same everywhere, it follows that the MRTS is the same in both sectors, which is the condition of efficiency.

To prove the second part of the proposition it suffices to show that profit maximization requires equality between opportunity cost and commodity-price ratio. Suppose, for example, that the economy is at point Q while the relative price is indicated by the slope of the line *PP*. This means that the opportunity cost of producing more B is lower than its price, $\tan\beta < \tan\gamma$ (note that this comparison makes sense because both the opportunity cost and the relative price under consideration are dimensionally homogeneous, being measured in terms of commodity A as *numéraire* or unit of measurement). It follows that producers can increase their profits by increasing the output of B. Only at R are the opportunity cost and the relative price equal and profits maximized. Similarly, if the relative price were given by the slope of the line $P'P'$, the opportunity cost of producing more A ($\tan\alpha$) would be lower than the relative price of A (p_A/p_B is measured by $\tan\delta$), and producers would maximize their profits by moving to point R'.

Another illuminating way of proving the equality between the marginal rate of transformation and the commodity price ratio is to pass through marginal costs. Suppose that we move slightly to the right on the transformation curve, thus increasing the output of commodity B and decreasing that of commodity A. If we consider a small displacement, an amount $\mathrm{d}K$ of capital and $\mathrm{d}L$ of labour will be transferred from sector A to sector B, and the additional cost in producing B is $\mathrm{d}C_B$, where of course

$$\mathrm{d}C_B = p_K \mathrm{d}K + p_L \mathrm{d}L. \tag{3.1}$$

Since we are moving on the transformation curve and therefore along the efficiency locus, the prices of productive factors must be equal in both sectors. Therefore the additional cost in producing B must be equal to the reduction of the cost in producing A, namely

$$\mathrm{d}C_B = -\mathrm{d}C_A. \tag{3.2}$$

The marginal costs of the two commodities are defined as

$$MC_A = \mathrm{d}C_A/\mathrm{d}A, \quad MC_B = \mathrm{d}C_B/\mathrm{d}B. \tag{3.3}$$

From this and the previous relation we obtain $MC_B = -\mathrm{d}C_A/\mathrm{d}B$, and if we compute the ratio of the marginal costs we get

$$MC_B/MC_A = (-\mathrm{d}C_A/\mathrm{d}B)/(\mathrm{d}C_A/\mathrm{d}A) = -\mathrm{d}A/\mathrm{d}B. \tag{3.4}$$

Now, $-\mathrm{d}A/\mathrm{d}B$ is the negative of the slope of the transformation curve (measured with respect to the B axis), namely (since this slope is negative) its absolute value, and it thus measures the marginal rate of transformation. Therefore Eq. (3.4) states that the marginal rate of transformation must be equal to the ratio of the marginal cost of B to the marginal cost of A. This is a general proposition, which is important in itself. To conclude our proof it is sufficient to recall that under competitive conditions in the output markets the price of a commodity equals its marginal cost, $MC_A = p_A$ and $MC_B = p_B$, so that we can rewrite (3.4) as

$$-\mathrm{d}A/\mathrm{d}B = p_B/p_A, \tag{3.5}$$

which was to be proved.

In Sect. 19.1.1 we give rigorous proofs of the results arrived at intuitively here.

3.2 General Equilibrium in a Simple Closed Economy

3.2.1 The Supply Curves

The first step is to derive from the transformation curve the supply curves of the two commodities as a function of the price ratio or relative price, p_B/p_A. With reference to Fig. 3.5a, suppose that p_B/p_A is equal to $\tan\alpha$: the optimum point on the transformation curve is then H, where the marginal rate of transformation and the relative price are equal. Therefore, quantities OA' of commodity A and OB' of commodity B will be supplied when the relative price is $\tan\alpha$. Similarly, quantities OA_E of commodity A and OB_E of commodity B will be supplied when p_B/p_A is $\tan\beta$. In short, a unique productive combination will correspond to every admissible price ratio.

In Fig. 3.5b we measure the price ratio on the vertical axis and the quantities of the two commodities on the horizontal axis: increasing quantities of A are measured from O to the right and increasing quantities of B from O to the left. Let $OP = \tan\alpha$: to this relative price, therefore, quantities OA' of commodity A and OB' of commodity B will correspond, which are equal to the coordinates of

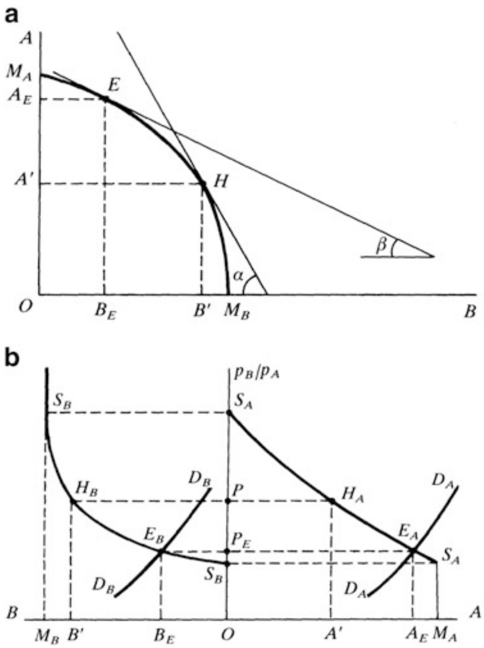


Fig. 3.5 Transformation curve, supply curves, and determination of general equilibrium in a closed economy

point H in Fig. 3.5a. Thus we obtain points H_A and H_B in Fig. 3.5b: since OA' is the quantity of commodity A supplied when the relative price is OP, point H_A will belong to the supply curve of commodity A; similarly, point H_B will belong to the supply curve of commodity B.

Continuing in this manner we obtain the supply curve of commodity A, S_AS_A, and the supply curve of commodity B, S_BS_B, which are general equilibrium supply curves. Both are increasing with respect to the appropriate relative price: S_BS_B is increasing with respect to p_B/p_A (the relative price of commodity B in terms of commodity A), and S_AS_A is increasing with respect to p_A/p_B. But, since S_AS_A is also drawn as a function p_B/p_A in Fig. 3.5b, it will be monotonically decreasing because p_B/p_A decreases as p_A/p_B increases.

The curve S_AS_A meets the vertical axis at a point which corresponds to that price ratio which causes the optimum point on the transformation curve to coincide with point M_B (see Fig. 3.5a), where the quantity of commodity A is zero and, correspondingly the quantity of commodity B is at its physically possible maximum, namely all the existing productive factors are employed in the production of B (for simplicity's sake we assume that the transformation curve's slope is neither infinite at M_B nor zero at M_A). This is denoted by the vertical stretch of the S_BS_B curve corresponding to OM_B in Fig. 3.5b, to show that it is impossible to produce more of commodity B than this amount.

Similarly, the $S_B S_B$ curve meets the vertical axis at a point which corresponds to that price ratio which causes the optimum point on the transformation curve to coincide with point M_A (see Fig. 3.5a), where the quantity of commodity B is zero and, correspondingly the quantity of commodity A is at its physically possible maximum OM_A.

3.2.2 The Demand Curves

The second step is to derive the demand curves of the two commodities as a function of p_B/p_A. As we have shown in Sect. 3.1, a point on the efficiency locus in the box diagram corresponds to each point on the transformation curve (namely to each price ratio), and so the marginal productivities of capital and labour are determined, for these productivities depend only on the factor ratios when the production functions are homogeneous. We recall from microeconomics that, in competitive equilibrium, the real rewards of the productive factors coincide with their marginal productivities; therefore—since the distribution of these factors is given, as assumed in point (b), Sect. 3.1—the real income of each individual is determined. The fact that a precise real income of each individual corresponds to each given price ratio means that—unlike in partial equilibrium analysis—individual real income cannot be assumed constant as relative prices change (see below). Now, given relative prices and income, each individual, by means of the well-known maximization of a utility index subject to the budget constraint, will determine the quantities of commodity A and of commodity B demanded. Summing these quantities for all individuals, we obtain the overall demands for A and B. If we repeat this procedure for all possible ratios p_B/p_A we obtain the market demand curves for goods A and B as functions of p_B/p_A.

It should be emphasized that these demand curves are different from the usual Marshallian or partial equilibrium demand curves, which express the quantity demanded of a good as a function of its (relative) price, and are obtained on the *ceteris paribus* assumption, namely that everything else—*including* (individual) *income*—is equal. On the contrary, in our derivation *income changes as* p_B/p_A *changes*: in fact, when p_B/p_A is different, we are at a different point on the transformation curve and so at a different point on the efficiency locus in the box diagram; therefore the marginal productivities of the factors will be different and, consequently, each individual's real income will be different. In other words, the demand curves we are dealing with are *general equilibrium demand curves*, which depend on real income as well as on relative prices; but, since real income depends on relative prices alone as shown above, we can express these demand curves as functions of relative prices alone.

For simplicity's sake we assume that these demand curves are decreasing with respect to the appropriate relative price (a rigorous treatment of this topic will be given in Sect. 19.2.3), so that $D_A D_A$—which is decreasing with respect to p_A/p_B—is increasing with respect to p_B/p_A.

3.2.3 General Equilibrium and Walras' Law

The last step is to draw the demand and supply curves on the same diagram, as in Fig. 3.5b. The intersection between the demand and supply curves of a commodity determines the equilibrium point, respectively E_A and E_B for goods A and B; the corresponding equilibrium quantities are OA_E and OB_E, and the equilibrium price ratio is OP_E, equal to $\tan \beta$. The equilibrium point on the transformation curve (Fig. 3.5a) is E; therefore—as we explained above—the allocation of the productive factors between the two sectors is determined, from which the determination of the marginal productivities and hence of the factors' real rewards and of the distribution of income follow. The general equilibrium of the economy has been established.

Last but not least, an important point needs clarification: in Fig. 3.5b we have taken it for granted that the equilibrium price ratio is the same in both markets. This equality is fundamental, since if the two markets were to be in equilibrium at different relative prices, the model would be inconsistent. A simple proof, based on Walras' law, allows us to conclude that if one market is in equilibrium the other must also be in equilibrium, so that the equilibrium price ratio cannot be different in the two markets.

Let p_K and p_L indicate factor rewards, S_A and S_B the quantities of the two commodities supplied, K and L with a subscript A or B the quantities of the two factors allocated in the two sectors. Let us now recall that in each sector total factor rewards equal the value of output. This is true with constant returns to scale (first-degree homogeneous production functions: see Euler's theorem in Sect. 19.1.3), but is also true with any kind of production function provided that free entry and exit of competing firms obtain (see, for example, Mas-Colell, Whinston, & Green, 1995, sect. 10.F). Thus we have

$$p_K K_A + p_L L_A = p_A S_A,$$
$$p_K K_B + p_L L_B = p_B S_B,$$

from which

$$p_K (K_A + K_B) + p_L (L_A + L_B) = p_A S_A + p_B S_B. \tag{3.6}$$

The left-hand side of (3.6) is the total income of all the individuals in the economy (that they obtain by selling the services of the productive factors they own). Since in this model income is entirely spent in buying commodities A and B, we can write

$$p_K (K_A + K_B) + p_L (L_A + L_B) = p_A D_A + p_B D_B, \tag{3.7}$$

where D_A and D_B are the quantities demanded of the two commodities. Equation (3.7) is the aggregate budget constraint. From Eqs. (3.6) and (3.7) it follows that the right-hand sides must be equal, as the left-hand sides are equal. Therefore

$$p_A D_A + p_B D_B = p_A S_A + p_B S_B, \tag{3.8}$$

whence

$$p_A (D_A - S_A) + p_B (D_B - S_B) = 0, \tag{3.9}$$

which is true for any admissible value of p_A and p_B. The form (3.8) states that the sum of the values of the quantities demanded must equal the sum of the values of the quantities supplied; the form (3.9) states that the sum of the values of the excess demands must be equal to zero. This relationship, whichever the form used, is known as *Walras' law*. In general, given n markets linked by a (budget) constraint, Walras' law implies that if $n-1$ markets are in equilibrium, the nth must also be in equilibrium. In our case there are only two markets, so that if one is in equilibrium the other must also be: for example, if $D_A = S_A$ then Eq. (3.9) implies $D_B = S_B$, and vice versa.

3.3 General Equilibrium in Open Economies and International Trade

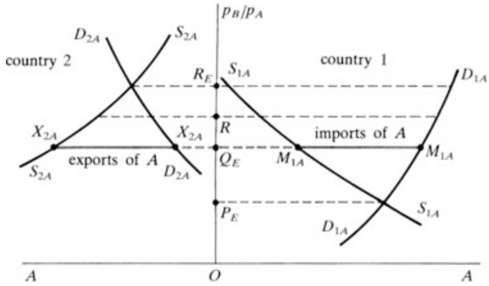
In this section we extend the previous analysis to the international economy. In addition to those already made, we make the assumptions that only two countries exist, country 1 (the home country) and country 2 (the rest of the world), that transport costs are absent (these will be considered in Chap. 6) and that perfect competition prevails in international markets. Both countries use the same factors, which are internationally immobile, and produce the same goods.

In the absence of international trade, both countries will be in a situation of equilibrium similar to that described in Fig. 3.5b. But, as factor endowments, technology, and tastes are different in each country, it is very unlikely that the equilibrium price ratio will be the same in both. If this were so, there would be no scope for international trade. Let us then assume that the closed economy equilibrium price ratios are different in the two countries; without loss of generality we can assume that this ratio is greater in country 2 than in country 1, as shown in the back-to-back diagram drawn in Fig. 3.6. This diagram was introduced by Cunynghame (1904) and Barone (1908), but in a partial equilibrium framework: see Viner (1937, pp. 589–590).

To avoid confusion with Fig. 3.5b, we stress that in Fig. 3.6 the demand and supply curves refer to the *same commodity in the two countries*: in the right-hand part there are the demand and supply curves for commodity A in country 1, and in the left-hand part there are the demand and supply curves for the same commodity in country 2. As assumed above, the closed-economy equilibrium price-ratio in country 2 (OR_E) is greater than in country 1 (OP_E).

It can be easily shown that when trade is opened up, commercial relations are possible only if the international price ratio or *terms of trade* lies somewhere between the two internal equilibrium price ratios. We first observe that with free trade, perfect competition and no transport costs, the same commodity must have

Fig. 3.6 Determination of international equilibrium



the same price everywhere (the law of one price), so that the international and the national price ratios are the same. Now, for terms of trade higher than OR_E, both countries would demand commodity A internationally, because in both of them there would be an excess demand for this commodity, and no equilibrium would be possible. Similarly, for terms of trade lower than OP_E, both countries would supply commodity A internationally, because in both of them there would be an excess supply of this commodity. Therefore, only intermediate terms of trade are to be considered, since between OP_E and OR_E country 1 will demand, and country 2 will supply, commodity A.

International equilibrium will be established at a point where the excess demand for good A by country 1 (country 1's demand for imports) is exactly matched by the excess supply of the same commodity by country 2 (country 2's supply of exports). This point is shown in Fig. 3.6 at the terms of trade OQ_E, where $M_{1A}M_{1A} = X_{2A}X_{2A}$. It can be shown that this equilibrium is stable under the usual dynamic behaviour assumption, i.e., that price varies according to excess demand.

Suppose, for example, that we are at point R_E, where country 2 is in internal equilibrium and so will not demand or supply anything abroad. On the contrary, country 1 will have an excess demand for commodity A measured by the horizontal distance between the $D_{1A}D_{1A}$ and $S_{1A}S_{1A}$ curves in correspondence to OR_E. According to Walras' law—see Eq. (3.9)—an excess supply of commodity B will also be present in country 1. Therefore this country will supply commodity B (the exportable commodity) and demand commodity A (the importable commodity) on the international market. But, since there is no demand for B nor supply of A coming from country 2, on the international market there will be an excess supply of B and an excess demand for A. As a consequence the international relative price of B with respect to A will decrease, for example to OR.

When the terms of trade is OR, in country 1 there is still an excess demand for commodity A (and so an excess supply for commodity B) though smaller than before, whereas in country 2 an excess supply of A (and so an excess demand for B) has appeared. But it is easy to see that the excess demand for A by country 1 is greater than the excess supply of it by country 2, so that on the international market an excess demand for A (and thus an excess supply of B) will still be present. A further decrease in the terms of trade will occur, and this process will go on until point Q_E is reached, where the excess demand for good A by country 1 is exactly matched by the excess supply of the same good by country 2. On the international market for good A, there is equilibrium between demand and supply at the terms of trade OQ_E (and, as we shall see presently, the international market for good B will also be in equilibrium). Country 1 will import an amount $M_{1A}M_{1A}$ of commodity A, exactly equal to the amount $X_{2A}X_{2A}$ of the same commodity exported by country 2; conversely (see below) country 1 will export, and country 2 will import, commodity B.

We could have arrived at the same point Q_E by starting from a lower price ratio, for example OP_E (internal equilibrium in country 1; excess supply of A and excess demand for B in country 2 and hence on the international market; increase in the relative price of B, etc.).

In Fig. 3.6, the position of the supply and demand curves for A in each country depends, as we know from Sect. 3.2, on factor endowments, technology, and tastes existing in the country. These are the elements that determine, *ceteris paribus*, the relative position of the two sides of the diagram under consideration and, therefore, which commodity will be imported and which exported. In fact, if the above elements were such that OR_E were lower than OP_E, then it would be country 1 which would export, and country 2 which would import, commodity A. This proves the following important conclusion: *in the neoclassical model of international trade, the existence of commercial relations, the pattern and the volume of trade, and the terms of trade, are jointly determined in a general equilibrium setting by factor endowments, technology,and tastes, none of which can be in general said to be an exclusive or predominant causal agent.*

We have stated above that the terms of trade which equate demand and supply in the international market for commodity A must necessarily equate it in the other market. This is a consequence of Walras' law extended to the international economy. In each country, the total value of demands equals the total value of supplies as stated in Eq. (3.8), and if we let the subscripts 1 and 2 refer to countries 1 and 2 respectively, we have

$$\begin{aligned} p_A D_{1A} + p_B D_{1B} &= p_A S_{1A} + p_B S_{1B}, \\ p_A D_{2A} + p_B D_{2B} &= p_A S_{2A} + p_B S_{2B}. \end{aligned} \tag{3.10}$$

By addition we obtain

$$p_A \left(D_{1A} + D_{2A}\right) + p_B \left(D_{1B} + D_{2B}\right) = p_A \left(S_{1A} + S_{2A}\right) + p_B \left(S_{1B} + S_{2B}\right), \tag{3.11}$$

namely the total value of world demands equals the total value of world supplies. This equation can also be written as

$$p_A \left[\left(D_{1A} - S_{1A}\right) + \left(D_{2A} - S_{2A}\right)\right] + p_B \left[\left(D_{1B} - S_{1B}\right) + \left(D_{2B} - S_{2B}\right)\right] = 0, \tag{3.12}$$

or

$$p_A\left[(D_{1A}+D_{2A})-(S_{1A}+S_{2A})\right]+p_B\left[(D_{1B}+D_{2B})-(S_{1B}+S_{2B})\right]=0, \tag{3.13}$$

namely the sum of the values of world's excess demands must equal zero for any admissible value of p_A and p_B.

Suppose now that, at a particular price ratio, the international market for commodity A is in equilibrium, i.e.

$$D_{1A}+D_{2A}=S_{1A}+S_{2A}; \tag{3.14}$$

then it follows from Eq. (3.11) that

$$D_{1B}+D_{2B}=S_{1B}+S_{2B}, \tag{3.15}$$

namely that the international market for commodity B is also in equilibrium. From (3.14) and (3.15) it also follows that

$$\begin{aligned} D_{1A}-S_{1A} &= S_{2A}-D_{2A}, \\ S_{1B}-D_{1B} &= D_{2B}-S_{2B}, \end{aligned} \tag{3.16}$$

which state that excess demand for good A by country 1 (country 1's demand for imports) is equal to excess supply of the same good by country 2 (country 2's supply of exports) and that country l's supply of exports of good B is equal to country 2's demand for imports of the same good.

It is also worth pointing out that conditions (3.10) imply that no country can be a net importer or exporter of *both* commodities. In fact, if we rewrite these conditions as

$$\begin{aligned} p_A\left(D_{1A}-S_{1A}\right) &= p_B\left(S_{1B}-D_{1B}\right), \\ p_A\left(D_{2A}-S_{2A}\right) &= p_B\left(S_{2B}-D_{2B}\right), \end{aligned} \tag{3.17}$$

we see that if $D_{1A} > S_{1A}$ (excess demand for commodity A by country 1, which thus imports this commodity), then $S_{1B} > D_{1B}$ (country 1 exports commodity B) and vice versa. This result is obvious if we think that in the barter model under consideration a country can obtain imports only by paying for them with exports. It should also be noticed that Eqs. (3.17) can be interpreted as the equality, for each country, between the value of its imports and the value of its exports when both are *evaluated at the given international prices*. Therefore, as is typical in the pure theory of international trade, the balance of trade always balances.

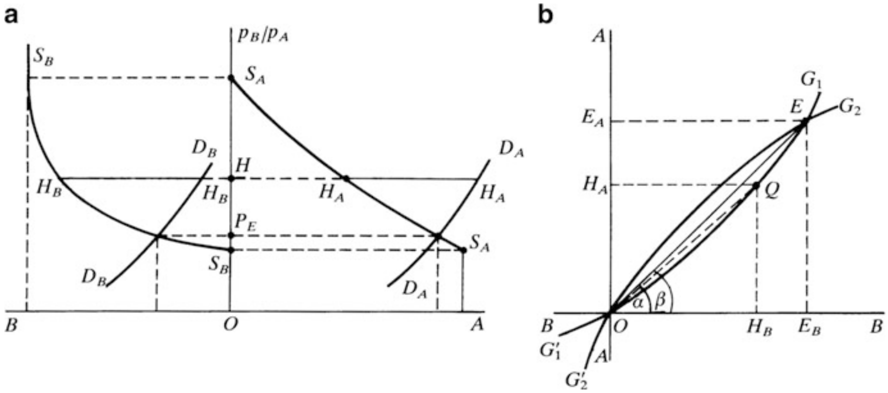


Fig. 3.7 Derivation of the offer curve

3.4 Offer Curves, International Equilibrium, and Stability

3.4.1 Derivation of the Offer Curve

An alternative way of determining international equilibrium is to use the Marshallian *reciprocal demand curves* (also called *offer curves* and *demand-and-supply curves*). The notion of reciprocal demand is already present in J.S. Mill, as we saw in Sect. 2.5, but the first complete treatment is to be found in Marshall (1879), who also introduced the graphic apparatus of the offer curves, though he did not show how they are derived from the underlying production and demand conditions.

The offer curve of a country can be defined as the locus of all points which represent the (maximum) quantity of the exported good that the country is willing to give in exchange for a given amount of the imported good (or, if we prefer, the (minimum) quantity of the imported good that the country is willing to accept in exchange for a given amount of the exported good). Equivalently, this curve indicates the various terms of trade at which the country is willing to trade.

There are several ways of obtaining a country's offer curve geometrically; one is Meade's (1952) ingenious geometric technique based on trade indifference curves and the transformation curve. However, the graphic apparatus developed in Sect. 3.2 allows a very simple derivation of the offer curve, and we shall use this instead of Meade's technique.

In Fig. 3.7a the same diagram contained in Fig. 3.5b is drawn. Let us consider an arbitrary price ratio, for example OH. At this relative price, country 1 has an excess demand for good A equal to $H_A H_A$ and an excess supply of good B equal to $H_B H_B$. This country, therefore, is willing to exchange $H_B H_B$ of B for $H_A H_A$ of A on the international market, namely it is willing to import an amount $H_A H_A$ of commodity A and to export, in exchange for this, an amount $H_B H_B$ of commodity B.

In Fig. 3.7b we draw the amounts of A and B just obtained, measuring the demand for imports on the vertical axis ($OH_A = H_A H_A$) and the supply of exports on the horizontal axis ($OH_B = H_B H_B$); we thus obtain point Q. The *terms of trade* in Fig. 3.7b are represented by OH_A/OH_B, we recall that p_B/p_A expresses the number of units of A for one unit of B, and the same thing is expressed by the ratio (OH_A/OH_B), namely by the slope of OQ, which is $\tan\alpha$; this is equal to OH in Fig. 3.7a.

If we let the price ratio take on all values from OP_E upwards, we obtain other points in a similar way, which give rise to the curve OG_1. For values of the price ratio lower than OP_E the export-import situation of country 1 will be reversed, because there will be an excess supply of commodity A and an excess demand for commodity B. If we adopt the convention of measuring the import demand for B by country 1 on the horizontal axis from O to the left, and the export supply of A by this same country on the vertical axis from O downwards, we obtain the branch OG_1' of the offer curve of country 1. If the price ratio is OP_E in country 1 there will be no excess demand or excess supply, therefore this country's offer curve will pass through the origin; the slope of the $G_1'OG_1$ curve measured at the origin is equal to the internal equilibrium price-ratio OP_E.

To sum up: every point of the OG_1 curve gives the demand by country 1 for imports of commodity A and the corresponding supply of exports of commodity B; every point of the OG_1' curve gives the supply by country 1 of exports of A and the corresponding demand for imports of B. The curve $G_1'OG_1$ is, therefore, the offer curve of country 1. Note that, since the domestic demand and supply curves have been obtained by an optimization procedure (as shown in Sect. 3.2), concerning both the demand and the supply, the excess demands and supplies which give rise to the offer curve, and therefore this curve, have an optimal nature.

In a similar way we can build the offer of country 2, $G_2'OG_2$. Given the assumption made in Fig. 3.6, when the price-ratio is lower than OR_E (which equals the slope at the origin of the $G_2'OG_2$ curve in Fig. 3.7b), country 2 has an excess supply of commodity A (and so an excess demand for commodity B). Then each point of the OG_2 curve gives the supply by country 2 of exports of A and its corresponding demand for imports of B.

This derivation of the offer curve shows the truth of Edgeworth's often quoted statement: "There is more than meets the eye in Professor Marshall's foreign trade curves. As it has been said by one who used this sort of curve, a movement along a supply-and-demand curve of international trade should be considered as attended with rearrangement of internal trade; as the movement of the hand of a clock corresponds to considerable unseen movements of the machinery" (Edgeworth, 1905, p. 70; p. 143 of the reprint. He was actually quoting himself: see Edgeworth, 1894, pp. 424–425; p. 32 of the reprint).

3.4.2 International Equilibrium and Stability

We saw above that no international trade is possible when the terms of trade are lower than OP_E or higher than OR_E, and this is reflected in the fact that in the third quadrant in Fig. 3.7b both countries are net suppliers or net demanders of the same commodity. The branches OG'_1 and OG'_2, therefore, are not relevant, and only the first quadrant has to be considered, where country 1 demands A and supplies B, and country 2 supplies A and demands B. The offer curves OG_1 and OG_2 intersect at point E, which is the equilibrium point: country 1 demands OE_A of commodity A, exactly equal to the amount of A supplied by country 2, and supplies OE_B of commodity B, exactly equal to the amount of B demanded by country 2. International trade will take place on the basis of OE_B of B (exported by country 1 and imported by country 2) for OE_A of A (imported by country 1 and exported by country 2); the equilibrium terms of trade are measured by $\tan\beta$ (slope of the ray OE), which is equal to OQ_E in Fig. 3.6.

The offer curves are widely used in international economics not only for determining international equilibrium but also for a number of other purposes, as we shall see in this and in the following chapters. It is therefore important to bear in mind that they are derived from the underlying production and demand conditions, as pointed out in Edgeworth's statement quoted above.

We now put the offer curves to use for examining the *stability* of the equilibrium point E when the adjustment process directly involves quantities rather than the terms of trade. It is well-known that to examine the stability of equilibrium we need *behaviour assumptions* concerning the reaction of the relevant variables to a disequilibrium situation. In Sect. 3.3 we examined the problem of stability by making the assumption that the variable which adjusts itself in the first instance is the terms of trade, reacting to excess demand and supply on the international market. In other words, the adjustment mechanism acted on the relative *price*, and quantities followed. Now—following Marshall (1879, 1923)—we make the assumption that the variables which adjust themselves in the first instance are the *quantities* of the two commodities. There are, however, at least two ways in which this adjustment may take place, namely there are at least two possible behaviour assumptions,[1] that we will now examine.

Behaviour Assumption I Consider any non-equilibrium point P. Owing to the competition between its traders, each country adjusts the quantity of its exports towards that quantity which it would offer at the terms of trade actually prevailing, if such terms remained fixed for all the time needed to complete the adjustment.

With reference to Fig. 3.8, assume that the initial non-equilibrium point is P. Now, OH_1 is the initial quantity of exports of country 1 and OH_2 is the initial

[1] See Kemp (1964, chap. 4), who attributes assumption II to Marshall, while leaving assumption I unnamed. Owing to the ambiguity of Marshall's statements (1879, 1923) on this topic, we believe that both assumptions are consistent with what he wrote. See also Samuelson (1947).

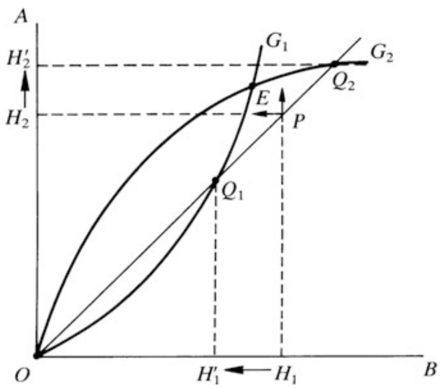


Fig. 3.8 Adjustment of quantities and stability of international equilibrium: Behaviour assumption I

quantity of exports of country 2; the terms of trade are measured by the slope of the straight line passing through P and O. Given these terms of trade the quantity of exports that country 1 wishes to supply is determined, by the very definition of an offer curve, by the abscissa of point Q_1. Therefore, country 1 is inclined to decrease its exports, and adjusts them from OH_1 towards OH_1'. By similar reasoning, it can be seen that country 2 tends to expand its exports, by adjusting them from OH_2 towards OH_2'. Thus point P moves in a direction to be found between the two arrows, tending to point E.

It is perhaps worth pointing out that this method of studying stability by means of pointed arrows representing the forces at play—a method now widely used in international economics as well as in other branches of economics—was first introduced by Marshall (1879) in order to study the stability of international equilibrium. It should however be stressed that the arrows do *not*, by themselves, make it possible to determine the actual trajectory of point P and even less to say whether this point will converge to the equilibrium point, or how. They are useful expository devices, but cannot replace a rigorous formal analysis (for further comments on arrow diagrams, see Gandolfo, 2009, chap. 19, sect. 19.3). This analysis is carried out in Sect. 19.4.2 for behaviour assumptions I and II, to the latter of which we now turn.

Behaviour Assumption II Consider any point P different from the equilibrium point. Each country adjusts its supply of exports towards that quantity of exports which it would offer if the current quantity of imports (corresponding to point P) remained fixed for the whole time needed to complete the adjustment.

In other words, each country moves towards the point on the respective offer curve corresponding to the prevailing quantity of the country's imports. With reference to Fig. 3.9, assume that the initial non-equilibrium point is P. Now, OY is the initial quantity of imports of country 1 and OX is the initial quantity of imports of country 2. The quantity of exports that country 1 wishes to offer in exchange

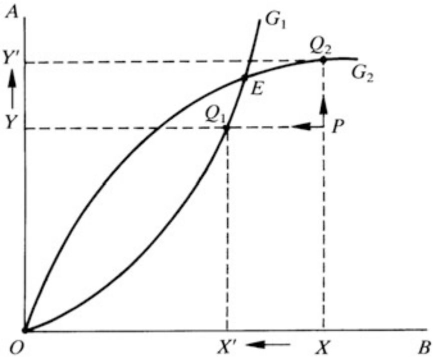


Fig. 3.9 Adjustment of quantities and stability of international equilibrium: Behaviour assumption II

for the current quantity of imports is OX'; consequently, this country adjusts its exports from the current quantity OX towards the desired quantity OX'. Similarly, it can be seen that country 2 adjusts its exports from the current quantity OY towards the desired quantity OY'. Thus point P moves in a direction to be found between the two arrows, tending to point E.

Thus we have seen that the equilibrium point E is stable according to both behaviour assumptions. But this has occurred because we have assumed that the offer curves have the "normal" form, i.e., they are both monotonically increasing and each one is concave to its import axis. But other shapes of the offer curves are admissible, so that cases may arise in which equilibrium is unstable according to both behaviour assumptions, as well as cases in which equilibrium is stable according to one assumption and unstable according to the other (Kemp, 1964, pp. 68–69).

It can be shown (see Sect. 19.4.2) that the local stability conditions can be expressed in terms of the elasticities of the offer curves. These elasticities can be defined in several ways (elasticity of imports with respect to exports, elasticity of exports with respect to imports, etc.). We follow Kemp (1964) in defining the elasticity of an offer curve as the proportional change in (the supply of) exports divided by the proportional change in (the demand for) imports. This implies that, when writing the offer curve as an explicit function, we choose to express (the supply of) exports as a function of (the demand for) imports instead of the other way round. This choice is consistent with the dynamic behaviour assumption just examined, where the variable which adjusts itself is the supply of exports. Formally, let $B^S = G_1(A^D)$ be the offer curve of country 1. The quantity B^S is country 1's supply of exports, which in turn is equal to the domestic excess supply, as shown in Sect. 3.4.1. In symbols, $B^S = S_1^B - D_1^B$. Similar observations hold for A^D, A^S, B^D.

The elasticity of the offer curve—for infinitesimal changes—is defined as

$$e_1 = \frac{\mathrm{d}B^S/B^S}{\mathrm{d}A^D/A^D} = \frac{\mathrm{d}B^S}{\mathrm{d}A^D} \cdot \frac{A^D}{B^S}, \tag{3.18}$$

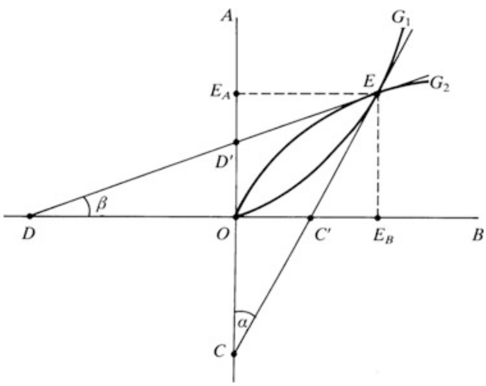


Fig. 3.10 Graphic measurement of offer curve elasticity

where dB^S/dA^D is the slope of the OG_1 curve referred to its import axis. Similarly, letting $A^S = G_2\,(B^D)$ be the offer curve of country 2, its elasticity is

$$e_2 = \frac{dA^S/A^S}{dB^D/B^D} = \frac{dA^S}{dB^D} \cdot \frac{B^D}{A^S}. \tag{3.19}$$

These elasticities can be measured graphically in a simple way. Consider for example point E in Fig. 3.10. The slope of the OG_1 curve with respect to its import axis is $\tan\alpha$. Now, $\tan\alpha = EE_A/E_AC = OE_B/E_AC$; note also that the angle $C'\hat{E}E_B$ is equal to α, so that $\tan\alpha = C'E_B/EE_B$ as well. Furthermore, $A^D = OE_A = EE_B$, and $B^S = OE_B = EE_A$. Therefore

$$e_1 = \frac{OE_B}{E_AC} \cdot \frac{OE_A}{OE_B} = \frac{C'E_B}{EE_B} \cdot \frac{EE_B}{OE_B},$$

from which

$$e_1 = \frac{OE_A}{E_AC} = \frac{C'E_B}{OE_B}. \tag{3.20}$$

In a similar way we obtain

$$e_2 = \frac{OE_B}{E_BD} = \frac{D'E_A}{OE_A}. \tag{3.21}$$

Equations (3.20) and (3.21) are simple and useful expressions for measuring the elasticities of the offer curves graphically. Note that if we defined these elasticities the other way round, their graphic measures would be the reciprocals of the expressions given in Eqs. (3.20) and (3.21).

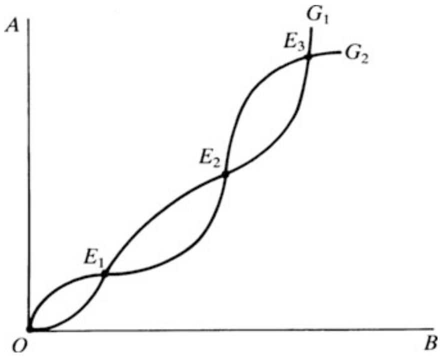


Fig. 3.11 Multiple equilibria and stability-instability

Going back to the stability conditions, it turns out (see Sect. 19.4.2) that necessary and sufficient conditions for local stability are

$$\frac{1-e_1e_2}{(1-e_1)\,(1-e_2)}>0 \tag{3.22}$$

if we adopt behaviour assumption I, and

$$1-e_1e_2>0 \tag{3.23}$$

if we adopt behaviour assumption II. If both elasticities are positive and smaller than 1, as they are in the cases examined so far, then both (3.22) and (3.23) are satisfied. But in abnormal cases anything may happen, for example contradictory results of the two behaviour assumptions, as already mentioned.

In conclusion, let us note that *multiple equilibria* may occur, as was demonstrated by Marshall (1879, 1923). One of the cases that were treated by him is shown in Fig. 3.11. According to Marshall, point E_2 is unstable, whereas points E_1 and E_3 are stable, thus respecting his proposition XIII (1879; p. 24 of the 1930 reprint) that in the case of multiple equilibria stable and unstable equilibria alternate. Although this proposition is not universally true, it holds in the case of Fig. 3.11, as can be seen either by using the graphic method of arrows or by applying conditions (3.22) and (3.23). It turns out that both e_1 and e_2 are greater than one at point E_2, whereas they are both smaller than one and positive at points E_1 and E_3. Therefore, neither (3.22) nor (3.23) is satisfied at point E_2, whereas both are satisfied at points E_1 and E_3, so that in this case Marshall's proposition holds independently of the behaviour assumption accepted.

3.5 Increasing Returns to Scale

In general, the presence of non-constant (decreasing or increasing) returns to scale has an effect on the curvature of the transformation curve. Since there seems to be a certain amount of imprecision in the literature when this effect is dealt with, we give a brief summary of the result (for proofs the reader is referred to Herberg (1969); see also Sect. 19.1.3).

In what follows, concavity and convexity are referred to the origin, different factor intensities in the two sectors are assumed, and it is also assumed that increasing (decreasing) returns to scale in a sector can be described by a homogeneous production function of degree higher (lower) than the first.

1. The transformation curve is strictly concave if both sectors have production functions with decreasing returns to scale or, more generally, if no sector produces with increasing returns.
2. Only slightly increasing returns in both sectors will make the transformation curve strictly convex near the coordinate axes and strictly concave somewhere in the intermediate range.
3. The transformation curve is strictly convex everywhere if, and only if, no sector has decreasing and at least one sufficiently strong increasing returns. The amount by which the degree(s) of homogeneity must exceed one is, ceteris paribus, the smaller the less the factor intensities of the commodities differ.
4. The transformation curve has at least one point of inflection if there are increasing returns in one sector and decreasing returns in the other. If the factor intensities happen to be equal in the two sectors, then:
5. Proposition 9.1 remains true if we exclude the case of constant returns in both sectors (in which case, as we know from Sect. 19.1, the transformation curve is linear).
6. The transformation curve is strictly convex if, and only if, one sector has increasing and the other no decreasing returns.

However, increasing returns to scale do not by any means only influence the shape of the transformation curve. As is well known, unlimited increasing returns to scale due to *internal* economies are incompatible with perfect competition; internal economies are however compatible with other market forms, for example monopoly (typical outcome of unlimited increasing returns) or monopolistic competition (these cases will be examined in Chap. 9). The compatibility of increasing returns with perfect competition is however preserved by the introduction of Marshallian *external* economies. On the other hand, when external economies are present, marginal social cost and marginal private cost are no longer the same. As a consequence, it is not certain that the economy produces on the transformation curve (the production point may lie inside this curve) and, even if it does, it is not certain that in equilibrium the price ratio will be equal to the marginal rate of transformation (for details of these problems, see Chipman, 1965b, pp. 736–749). We follow Meade (1952), Kemp (1964, 1969b) and others in assuming away these complications, namely we hypothesize that, notwithstanding the presence of increasing returns, the economy produces on the transformation curve at a point where the price ratio equals the marginal rate of transformation (sufficient conditions for this to be true are given by Kemp (1964, chap. 7, 1969b, chap. 8); for a treatment of the case in which the equality between price ratio and marginal rate of transformation no longer

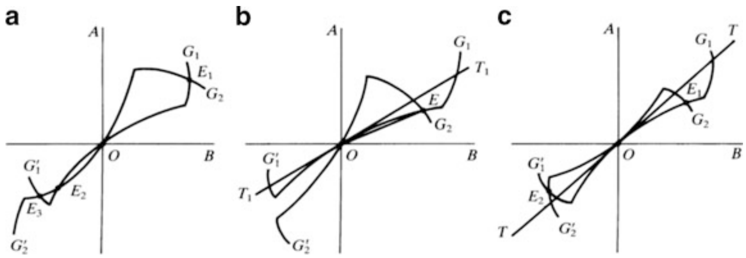


Fig. 3.12 Increasing returns to scale, offer curves, and international equilibrium

holds, see Chacholiades, 1978, chap. 7). We also assume that the transformation curve is strictly convex to the origin.

As a consequence of these simplifying assumptions, the formal analysis of increasing returns to scale does not differ from the analysis of constant returns to scale, for we only have to deal with the fact that the transformation curve is convex, instead of concave, to the origin.

The most interesting results concerning a trading world with increasing returns to scale in both countries are:

(a) In general there are multiple equilibria, and the direction of trade is not univocally determined;
(b) The equilibrium terms of trade may well lie outside the interval defined by the two closed-economy price ratios;
(c) Trade can take place even when the two closed-economy price ratios are equal.

These results can be easily obtained by using the offer curves. It turns out that, under increasing returns to scale, the offer curves have the shape shown in Fig. 3.12 (for their derivation see Chacholiades, 1978; Kemp, 1964, 1969b; Meade, 1952).

Figure 3.12a depicts a situation in which there are three equilibrium points: E_1, E_2, and E_3. Since, in the first quadrant, country 1 wishes to import commodity A and to export commodity B (and vice versa for country 2), whereas in the third quadrant the opposite is true, we see that the direction of trade is indeterminate. In other words, while in the case of constant returns to scale possible multiple equilibria do not alter the direction of trade, in the case under consideration a *normal* consequence of multiple equilibria is that of giving rise to *different* directions of trade. Therefore the direction of trade cannot be predicted on a priori grounds.

Figure 3.12b shows a case in which there is only one equilibrium point, and the equilibrium terms of trade (slope of the straight line segment OE) are lower than the autarkic price ratio in country 1 (the latter is measured, as in Sect. 3.4, by the slope at the origin of the $G_1'OG_1$ offer curve, namely by the slope of the straight line T_1T_1, which is tangent to $G_1'OG_1$ at the origin).

Finally, Fig. 3.12c depicts a situation in which the two autarkic price ratios coincide, for they are both equal to the slope of the straight line TT, which is the common tangent to both offer curves at the origin. Notwithstanding this, trade can and does take place, as shown by the two equilibrium points E_1 and E_2. We should like to underline this result, which shows that *increasing returns to scale can be a determinant of international trade*.

For a fuller treatment of increasing returns to scale in international trade the reader is referred to Kemp (1964, chaps. 8 and 12, sects. 7–8; 1969b, chaps. 8 and 11, sects. 7–8), Negishi (1972, chaps. 5 and 8), Chacholiades (1978, chap. 7), Helpman (1984b), Vanek (1962), Krauss (1979), Herberg et al. (1982) .

3.6 The Gains from Trade

We saw in the context of the classical theory that international trade is beneficial in so far as it enables a country to obtain a commodity at a lower cost than the domestic production cost or, alternatively, to obtain commodity bundles which were out of reach under autarky. A similar conclusion holds in neoclassical theory.

Consider for example Fig. 3.13 and suppose that the pre-trade closed-economy price ratio is represented by the slope of the straight line *PP*, whereas the terms of trade (post-trade open-economy price ratio) are represented by the slope of the straight line *RR*. Before trading started the country produced and consumed a commodity bundle given by the coordinates of point E. When trade is opened up, the country produces the commodity bundle given by the coordinates of point E' (*production* point). But it can now trade along the *RR* line, thus attaining previously unattainable points, outside its transformation curve. For example, it can move to point E'' (*consumption* point) by trading $H_B E'_B$ of commodity B (exportables) for $H_A E'_A$ of commodity A (importables); point E'' is clearly better (excluding inferior commodities) than the pre-trade point E because the amounts of both commodities are greater at E'' than at E. It can also be seen that—since we have assumed that A is the imported, and B the exported, commodity—the opportunity cost of A in terms of B is greater in the closed economy situation (slope of *PP* referred to the vertical axis) than in the open economy situation (where the additional amount of B that has to be given up to obtain an additional amount of A is measured by the appropriate terms of trade, namely by the slope of *RR* referred to the vertical axis).

But what if the post-trade situation is E'''? This point is undoubtedly outside the transformation curve, and thus it could not be reached before trade, but since with respect to E it contains a greater amount of commodity A and a smaller amount of commodity B, it cannot be considered unambiguously better than E. It is however easy to observe that the value of national income at E''' is in any case greater than at E. This is true whether national income is calculated at the closed-economy (pre-trade) prices or at the new (post-trade) prices. Let us first consider the closed-economy prices. The value of national income at E is given by the position of the equal income line (which we call *isoincome*) *PP*, while at E''' it is given by the position of the isoincome line (not shown in the diagram) parallel to *PP* and passing through E''', which is clearly more distant from the origin than *PP*. It follows that national income evaluated at the closed-economy prices is higher at E''' than at E.

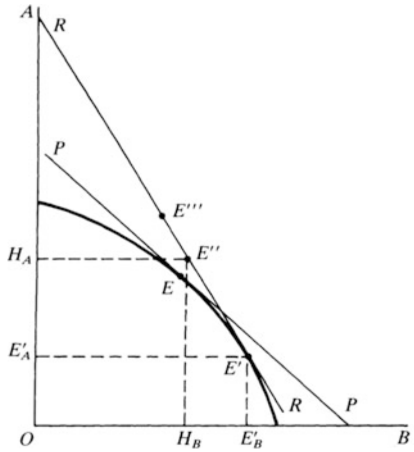


Fig. 3.13 The gains from trade

At the post-trade prices, the value of national income at E''' is given by the position of the isoincome line RR, while at E it is given by the isoincome line (not shown in the diagram) parallel to RR and passing through E, which is clearly nearer to the origin than RR (and hence represents a lower income).

It could also be observed that, since trade is *free* and not compulsory the fact that the country chooses point E''', instead of point E'', means that it prefers, in some sense, the former to the latter: we are in the presence of a sort of revealed preference.

The gains from trade can be given a more precise treatment if one is willing to accept the concept of *community* or *social indifference curves*. The problems raised by this concept are among the moot questions in welfare economics (see, for example, Mas-Colell et al., 1995, sect. 4.D; Chacholiades, 1978, chaps. 5 and 16). This notwithstanding, these curves are widely used in international economics and we do not depart from general practice by using them as a helpful expository device, though fully aware of their shortcomings.

In Fig. 3.14a the pre-trade (autarkic equilibrium) situation is depicted; social welfare is maximized at point E, where a social indifference curve is tangent to the transformation curve. In Fig. 3.14b, the terms-of-trade line RR is drawn: the highest indifference curve attainable is that which is tangent to this line, thus determining the consumption point E_C precisely, as well as the imported and the exported commodities and the amounts traded ($H_B E'_B$ of exports for $H_A E'_A$ of imports). The gains from trade are immediately visible, as the social indifference curve tangent at E_C is higher than the curve tangent at E, and so represents a better situation. Ideally, the gains from trade can be subdivided into a *consumption gain* and a *production gain*. The first is due to international exchange only, and can be seen by freezing the production point at the pre-trade point E. In this situation the country can trade along the $R'R'$ terms of trade line, parallel to the RR line; the optimum position is reached at point E'_C. Since the social indifference curve tangent at E'_C is higher than that tangent at E, there is a gain: the *consumption gain*.

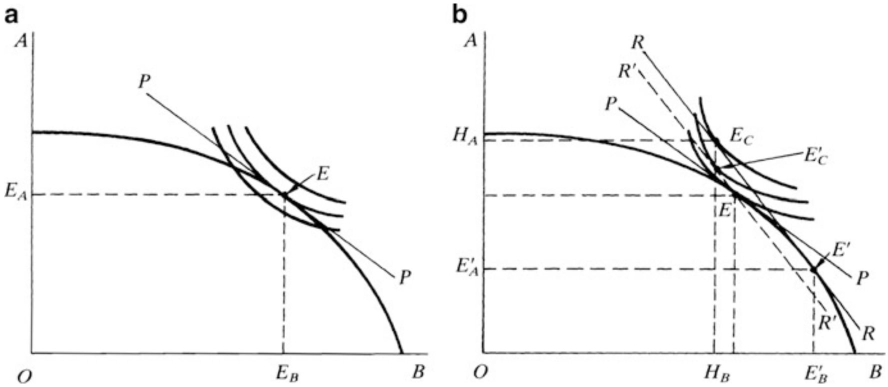


Fig. 3.14 Social indifference curves and the gains from trade: consumption and production gains

The *production gain* is due to specialization, since, as a consequence of the difference between the post-trade and the pre-trade commodity prices, the country changes its pattern of production and specializes (though incompletely) in the production of B, moving from the (now) inefficient production point E to the efficient one E'. This enables the country to reach a still higher indifference curve: the production gain is represented by the movement from E'_C to E_C.

We have stated that specialization is incomplete, as shown in the diagram. In fact, in the neoclassical theory—unlike the classical theory, where *complete* specialization was the necessary outcome of international trade—the specialization is normally incomplete (though complete specialization cannot be excluded: this occurs when the terms-of-trade line is tangent to the transformation curve at one of the points where this curve intersects the axes). The different results are due to the different assumptions concerning opportunity cost. Given a difference between the internal opportunity cost and the terms of trade, the productive combination will be modified in the direction of greater convenience. Now, if these modifications do not alter the opportunity cost (as in the classical theory: linear transformation curve), the inevitable outcome is complete specialization. On the contrary, when they bring about changes in the opportunity cost (as in the neoclassical theory), specialization will stop when opportunity cost becomes equal to the given terms of trade; this will normally occur at a point on the transformation curve somewhere between its two intercepts.

So far, we have considered one country only. What about our two-country world? It can be shown that trade is beneficial to *both* countries. In Fig. 3.15 we have drawn the transformation curves of the two countries together with the pre-trade and post-trade equilibria. The closed-economy equilibrium price ratio p_B/p_A is lower in country 1 (slope of P_1P_1 referred to the horizontal axis) than in country 2 (slope of P_2P_2): we are in a situation similar to that depicted in the back-to-back diagram (Fig. 3.6). The post-trade price ratio will lie between the two pre-trade ratios; country 1 will import commodity A and export commodity

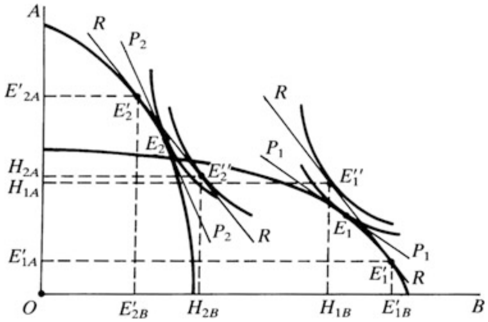


Fig. 3.15 Trade is beneficial to all countries

B, whereas the opposite will occur in country 2. This is shown in Fig. 3.15, where the slope of the terms-of-trade line *RR* measures the post-trade price ratio.Country 1 moves its production pattern from E_1 to E_1' (specializing in B), and country 2 moves its production pattern from E_2 to E_2' (specializing in A). Then country 1 exports $H_{1B}E_{1B}'$ of commodity B (equal to the quantity $E_{2B}'H_{2B}$ imported by country 2) and imports $E_{1A}'H_{1A}$ of commodity A (equal to the quantity $H_{2A}E_{2A}'$ exported by country 2).

As a consequence of these exchanges country 1's consumption point is at E_1'' (which lies on the highest social indifference curve of country 1 attainable given the terms-of-trade line RR) and similarly country 2's consumption point is at E_2'': as we see, both countries are on a higher indifference curve than in the pre-trade situation.

3.7 Generalizations

We have so far worked with the well-known $2 \times 2 \times 2$ model (two countries, two goods, two factors). But what happens when there are many countries, many commodities, and many factors? Among the first attempts to treat this problem formally is Yntema's (1932); 12 years later the problem was again tackled by Mosak (1944). Both of these, however, treated this topic *à la Walras*, namely by writing down equilibrium conditions and then counting equations and unknowns. The equilibrium conditions for the general problem can be written by making a straightforward extension of those holding in the $2 \times 2 \times 2$ model. In fact, application of the optimizing procedures to both the production side and the consumption side of each country makes it possible to derive the supply of and the demand for each commodity in each country as functions of relative prices only. Then world equilibrium requires that for each commodity world demand equals world supply, and, by summing the budget constraints, we find that, if all but one excess demands are equal to zero, then the last must also be.

But, as is well known, the mere counting of equations and unknowns is not a satisfactory procedure for proving the existence of an equilibrium, for in general the equality of the number of equations and of the number of unknowns is neither a necessary nor a sufficient condition for existence. An adequate proof must therefore rely on the same methods used in mathematical economics to prove the existence of general competitive equilibrium in a closed economy. Among the first modern proofs along these lines is Nikaidô's (1956, 1957); for further details, see Chipman (1965b, sect. 2.6).

There is, however, a price to be paid for this generality, because one must be content with knowing that an equilibrium exists (and with analysing its stability), without being able to find operational propositions allowing one to determine the structure and the volume of international trade, etc., in a simple way. On the other hand, the neoclassical theory can be used to yield simple predictions on the structure of international trade by restricting its generality. As a matter of fact, from the purely analytical point of view, the Heckscher-Ohlin theory (with all its corollaries, such as the factor price equalization theorem, etc.) can be considered as a particular case of the neoclassical theory: see Chap. 4.

The neoclassical theory can be generalized in several other directions, for example, by relaxing the assumption of fixed quantities of factors and introducing variable factor supplies, or by introducing transport costs, non-traded goods, specific factors etc. (see Chap. 6).

3.8 Duality Approach

Duality theory, which studies the dual relations between cost functions and production functions, between direct and indirect utility functions, etc. (for an introduction see Varian, 1992, chap. 6; a more advanced treatment is Diewert's, 1974, 1982) is being increasingly applied to microeconomics and to general equilibrium theory, as it enables us—among other things—to derive in a formally simpler way the comparative statics theorems originally deduced from maximizing behaviour.

Among the first applications of duality theory to international trade is the one by Jones (1965), who showed the dual nature of the Stolper-Samuelson (see below, Sect. 5.3) and Rybczynski (Sect. 5.4) theorems. Indeed, the whole pure theory of international trade can be rewritten by using duality theory: see, for example, Dixit and Norman (1980), Woodland (1982), and Sgro (1986).

However, much of the literature (especially as regards elementary and intermediate international economics textbooks, in some of which the duality approach is not even mentioned) is still based on the conventional approach. One reason may be that the conventional approach more easily lends itself to an intuitive verbal and graphical treatment and hence is more student-friendly. Another may be that the whole doctrinal body of international trade theory, from Ricardo to Heckscher-Ohlin and further, has been constructed and refined through the conventional approach.

Be it as it may, we have adhered to the conventional approach throughout the text, while treating the duality approach in the appendices (see Sect. 19.5 for the basic elements), where we also show in the appropriate places(e.g., Sects. 20.1–20.3, 21.2, 21.3, 22.3, 22.6) how certain formal results can be more easily derived using duality theory instead of the conventional approach.

3.9 Empirical Studies

Surprisingly enough, the neoclassical model of international trade in its general version has received little or no empirical attention, as practically all empirical studies have concentrated on the Heckscher-Ohlin model (see Sect. 4.6), that from the theoretical point of view can be considered as a particular case of the general neoclassical model.

A first step in the direction of filling this gap in the empirical literature was taken by Harrigan (1997), who specified a model of international specialization consistent with the neoclassical explanation. This model, where relative technology levels and factor supplies jointly determine international specialization, gives fairly good empirical results, so that "the neoclassical model comes out looking rather well" (Harrigan, 1997, p. 477).

For further considerations see Sect. 4.6.5.

References

Barone, E. (1908). *Principi di economia politica*.

Chacholiades, M. (1978). *International trade theory and policy* (Pt. III).

Chipman, J. S. (1965b). A survey of the theory of international trade: Part 2, the neoclassical theory.

Cunynghame, H. (1904). *A geometric political economy*.

Diewert, W. E. (1974). Applications of duality theory.

Diewert, W. E. (1982). Duality approaches to microeconomic theory.

Dixit, A. K., & Norman, V. (1980). *Theory of international trade: A dual, general equilibrium approach* (chaps. 1–3).

Edgeworth, F. Y. (1894). The theory of international values, I, II, III.

Edgeworth, F. Y. (1905). Review of H. Cunynghame's book "A Geometrical Political Economy".

Findlay, R. (1970). *Trade and specialization* (chap. 2).

Gandolfo, G. (2009). *Economic dynamics* (Pt. II, chap. 19, sect. 19.3; Pt. III, chap. 21, sect. 21.3.2).

Harrigan, J. (1997). Technology, factor supplies, and international specialization: Examining the neoclassical model.

Helpman, E. (1984b). The factor content of international trade.

Herberg, H. (1969). On the shape of the transformation curve in the case of homogeneous production functions.

Herberg, H., Kemp, M. C., & Tawada, M. (1982). Further implications of variable returns to scale.

Jones, R. W. (1965). The structure of simple general equilibrium models.

Kemp, M. C. (1964). *The pure theory of international trade* (Pt. II).

Kemp, M. C. (1969b). *The pure theory of international trade and investment* (Pt. I).

Krauss, M. B. (1979). *A geometric approach to international trade* (chap. II, 15–24 and 29–32).

Marshall, A. (1879). *The pure theory of foreign trade*.

Marshall, A. (1923). *Money, credit and commerce* (Appendix J).

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*.

Meade, J. E. (1952). *A geometry of international trade* (chaps. I–V).

Mosak, J. L. (1944). *General-equilibrium theory in international trade*.

Negishi, T. (1972). *General equilibrium theory and international trade* (chaps. 5 and 8).

Nikaidô, H. (1956). On the classical multilateral exchange problem.

Nikaidô, H. (1957). A supplementary note to "On the classical multilateral exchange problem".

Samuelson, P. A. (1947). *Foundations of economic analysis* (pp. 266–267).
Savosnick, K. M. (1958). The box diagram and the production possibility curve.
Sgro, P. M. (1986). *The theory of duality and international trade*.
Vanek, J. (1962). *International trade: Theory and economic policy*.
Varian, H. R. (1992). *Microeconomic analysis* (chap. 6).
Viner, J. (1937). *Studies in the theory of international trade*.
Woodland, A. D. (1982). *International trade and resource allocation*.
Yntema, T. O. (1932). *A mathematical reformulation of the general theory of international trade*.

Chapter 4
The Heckscher-Ohlin Model

4.1 Basic Assumptions and Their Meaning

We shall first examine the Heckscher-Ohlin theory (Heckscher, 1919; Ohlin, 1933) in its simplest version, that is a model in which there are two countries, two final goods and two primary factors of production. Extensions will be examined later on, in this chapter. Given the great contribution made by P. A. Samuelson to the refinement and diffusion of this theory, many authors call it the Heckscher-Ohlin-Samuelson model.

This model stresses the differences in factor endowments as the cause of international specialization and trade. More precisely, the key element in the theory is that countries are endowed with factors in different proportions. This, gives rise to different relative marginal cost of production and will make that each country exports the commodity which uses the country's more abundant factor more intensively (the Heckscher-Ohlin theorem).

In addition to the usual basic assumptions (no transport costs, free trade, perfect competition, international immobility of factors, presence of only two commodities and two factors) there are the following:

1. The production functions exhibit positive but decreasing returns to each factor (i.e., positive but decreasing marginal productivities) and constant returns to scale (i.e., first degree homogeneity). They are internationally identical, but, of course, different between the two goods, that is the production function of good A is the same in country 1 and country 2, and is different from that of good B (which is identical in the two countries).
2. The structure of demand, that is the proportions in which the two goods are consumed at any given relative price, is identical in both countries and independent of the level of income.
3. Factor-intensity reversals are excluded (see below).

The first assumption, which embodies the usual properties of well-behaved production functions, and excludes the presence of international technological

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_4,

differences, is self-evident. The difference between the production functions of the two goods is of course necessary, otherwise it would not be possible to speak of two *different* goods.

The second assumption implies that tastes are internationally identical *and* represented by utility functions such that the income elasticity of demand is constant and equal to one for each good. Utility functions having this property belong to the class of *homothetic* utility functions (see any microeconomics textbook). This assumption serves to exclude the possibility that, although tastes are internationally identical, the two goods are consumed in different proportions in the two countries because of possible differences in income levels.

It is then clear that the first two assumptions serve to exclude any difference between the countries as regards technology and demand, so that one can concentrate on the differences in factor endowments.

The third assumption is necessary to determine univocally the *relative factor intensities* of the two goods. In general, given two factors (capital K and labour L) and two commodities A and B, we say that a commodity (for example A) uses a factor more intensively or is more intensive in a factor (for example capital) relative to the other commodity if the (K/L) input ratio in the former commodity is greater than the (K/L) input ratio in the latter.

Now, if production of each good took place according to only one technique with fixed and constant technical coefficients (L-shaped isoquants), it would be an easy matter to determine the relative factor intensities once and for all. But since we are dealing with production functions with a continuum of techniques[1] (smoothly continuous isoquants), different techniques will be chosen—in accordance with the standard cost minimization procedure—for each good at different factor-price ratios. As already clarified in the previous chapter, we follow common practice in talking of the price of a factor in the sense of price of the services or rental for the services of the factor, or unit factor reward. This warning is to be considered as implicitly recalled throughout the rest of the book.

It follows that the classification of goods according to their factor intensities becomes ambiguous. To remove this ambiguity we add the requirement that the classification must remain the same for any (admissible) factor-price ratio, namely—in our example—that commodity A is more capital-intensive relative to commodity B if the (K/L) input ratio in the former commodity is greater than the (K/L) input ratio in the latter *for all factor-price ratios*.

Conversely, when *factor-intensity reversal(s)* occur, it is not possible to rank the commodities unambiguously for all factor-price ratios, that is, the classification changes according to the value of the factor-price ratio. For example, it may happen that A is more capital-intensive relative to B for a certain range of factor-price ratios, whilst B becomes more capital-intensive relative to A for another range of factor-price ratios: a factor-intensity reversal has occurred.

[1]The same problem would arise in the presence of many techniques, but limited in number, of the fixed-coefficients type, such as are dealt with by activity analysis.

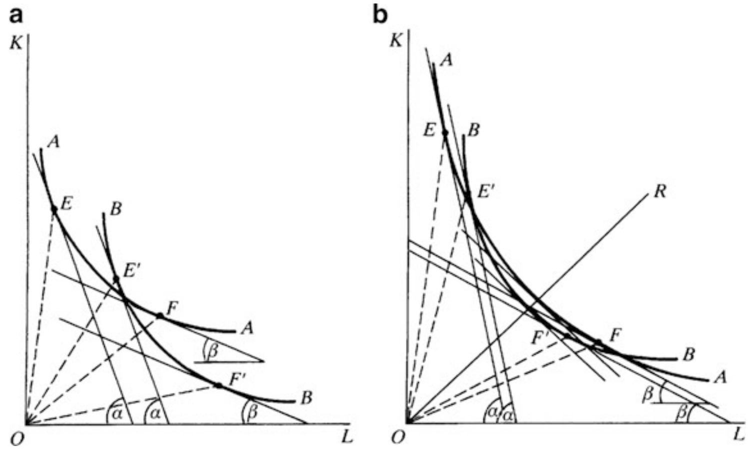


Fig. 4.1 Factor intensities: absence and presence of reversals

The condition which *excludes* factor-intensity reversals is that the representative isoquants of A and B, when drawn in the same diagram, do not cross more than once, as shown in Fig. 4.1a. Since with constant returns to scale all isoquants of the same production function have the same shape, the expansion path is linear and the input ratio, given the factor-price ratio, is the same for any output level, so that we can compare the representative isoquant of A with that of B, for example the unit isoquants. Note also that, owing to the assumption of internationally identical production functions, the following (Lerner-Pearce) diagram can refer equally well to either country.

Let us first consider Fig. 4.1a, where AA and BB indicate the unit isoquant of A and B respectively; these isoquants cross only once. If the factor-price ratio is, for example, equal (in absolute value) to $\tan\alpha$, then—by drawing a family of isocosts and following the usual cost minimization procedure (it goes without saying that the assumption of perfect domestic mobility of factors implies that the same factor-price ratio obtains in both industries)—we find the optimum input combinations: these are represented by point E in the A industry and by point E' in the B industry. The input ratios (K/L) in the two industries can be read off the diagram as the slopes of OE and OE' respectively, so that A is the capital-intensive commodity. At a different factor-price ratio, for example $\tan\beta$, the new optimum input combinations will be represented by points F and F' in the A and B industries respectively, so that A is, again, the capital-intensive commodity (slope of OF > slope of OF'). An examination of Fig. 4.1a will show that this property holds for each and all factor-price ratios: commodity A is, therefore, unambiguously capital-intensive relative to commodity B. It goes without saying that, in parallel, commodity B is unambiguously labour-intensive relative to commodity A.

Let us then consider Fig. 4.1b, where the isoquants intersect twice. When the factor-price ratio is equal to $\tan\alpha$, the optimum input combinations in the two

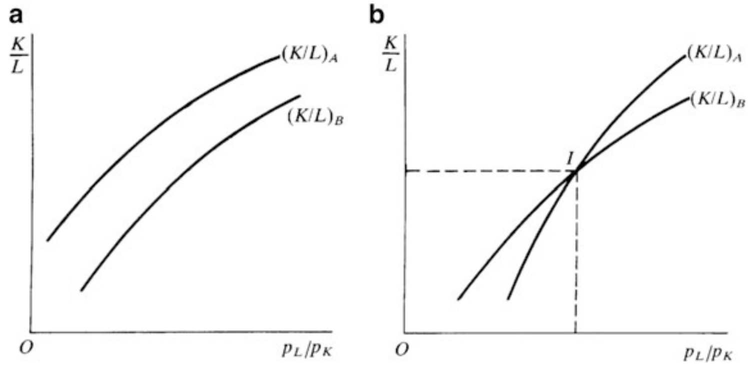


Fig. 4.2 Behaviour of the K/L ratio without and with factor-intensity reversal

industries are such that A is capital-intensive relative to B (slope of OE > slope of OE'), whilst the opposite is true when the factor-price ratio is equal to tan β (slope of OF' > slope of OF, so that now B is capital-intensive relative to A): a factor-intensity *reversal* has occurred. The reader can check as an exercise that such a reversal also occurs when one of the isoquants is internal to the other and they are tangent to each other at a point. This is not surprising, for a point of tangency is—loosely speaking—more similar to a multiple than to a single intersection. In mathematical terms, a point of tangency between two curves can be considered as the limit to which two (or more) intersection points tend when approaching indefinitely.

In Fig. 4.1b, the K/L ratio corresponding to which the reversal takes place is given by the slope of ray OR, along which the unit isoquant of A and the unit isoquant of B have the same slope, as can be seen from the fact that the two straight lines tangent to the isoquants along ray OR are parallel. A. P. Lerner (1952, p. 14) called this ray a "radiant of tangency", as all the A and B isoquants will have the same slope along it. It can be read off the diagram that the K/L ratio is greater in the A than in the B industry for all factor-price ratios such that the optimum input combinations lie above OR, and vice versa in the opposite case.The behaviour of the K/L ratio in the two sectors in the absence and presence of a factor intensity reversal is shown in Fig. 4.2. In all cases the K/L ratio is a monotonically increasing function of the factor-price ratio or relative price of factors (p_L/p_K), since producers will find it profitable to substitute capital for labour as the relative price of labour increases. This can be derived diagrammatically by considering the various points of tangency to the unit isoquant of isocosts with varying slope. But, whilst in the case of no factor-intensity reversals the two curves never intersect, in the case of a factor-intensity reversal they do.

In Fig. 4.2a, derived from Fig. 4.1a, the curve representing the K/L ratio in industry A—curve $(K/L)_A$—lies above the $(K/L)_B$ curve throughout: commodity A is always capital-intensive relative to commodity B.

In Fig. 4.2b, derived from Fig. 4.1b, the curves under consideration intersect in correspondence to the K/L ratio represented by the slope of OR, which in turn corresponds to the p_L/p_K ratio given by the common slope of the two isocosts tangent to the two unit isoquants along *OR*. To the left of the point of intersection I, that is for lower p_L/p_K ratios—corresponding to the part of Fig. 4.1b to the right of *OR*—commodity B is capital-intensive relative to A, whilst the opposite is true to the right of I (higher p_L/p_K ratios, corresponding to the part of Fig. 4.1b to the left of *OR*).

We have so far examined the case of a single reversal, corresponding to the fact that the A and B unit isoquants intersect twice, but it cannot be excluded that the unit isoquants intersect more than twice, giving rise to *more than one* factor-intensity reversal; in such a case, the two curves in Fig. 4.2 would intersect twice or more. In general, $n - 1$ factor-intensity reversals correspond to n intersections of the unit isoquants. The phenomenon of factor-intensity reversals is related to the elasticity of substitution between factors. In fact, the economic meaning of the circumstance that the isoquants cut twice is that the possibilities of factor substitution are different between the two sectors. Loosely speaking, the isoquants can cut twice when one is more curved (more convex to the origin) than the other, and the curvature of an isoquant is related to the elasticity of substitution (the more highly curved the isoquant is, the poorer substitutes the two factors of production are). This can be generalized to more than two intersections (see Sect. 20.1).

4.1.1 Relative Price of Goods and Relative Price of Factors

Although not immediately relevant, it is convenient to show now that, in the case of no factor-intensity reversal (also called the *strong* factor-intensity assumption), *a unique factor-price ratio corresponds to each commodity-price ratio*, *and vice versa*, i.e. there is a *one-to-one correspondence* between the relative price of goods and the relative price of factors.

Let us for example assume that the commodity-price ratio is $p_B/p_A = 4$, that is, four units of A exchange for one unit of B; in perfect competition, this implies that the production cost of one unit of B must be the same as that of four units of A. In fact, in the long run perfect competition leads to a situation in which the price of a commodity equals its production cost (see any microeconomics textbook). Since we have assumed $p_B/p_A = 4$ (the price of B is four times that of A) it follows that the production cost of one unit of B must be the same as that of four units of A.

In Fig. 4.3 (which, owing to the assumption of internationally identical production functions, equally applies to either country) we have drawn the isoquants $4A$ and $1B$. Since factor prices are equal in the two sectors and since the production cost of one units of B is the same as that of four units of A, it follows that the optimum (i.e., the minimum) isocost will be the same for $1B$ and $4A$. So we must find an isocost which is simultaneously tangent to isoquants $4A$ and $1B$; once found, (the absolute value of) its slope will give us the relative price of factors.

It can be clearly seen in Fig. 4.3a that only one such isocost (*CC*) exists in the case of a single intersection of the isoquants: therefore, a unique factor-price ratio corresponds to the given commodity-price ratio. It should be noted that the result does not change if we consider any couple of A and B isoquants standing in the ratio 4:1. For example, in Fig. 4.3a the unique isocost being simultaneously tangent to

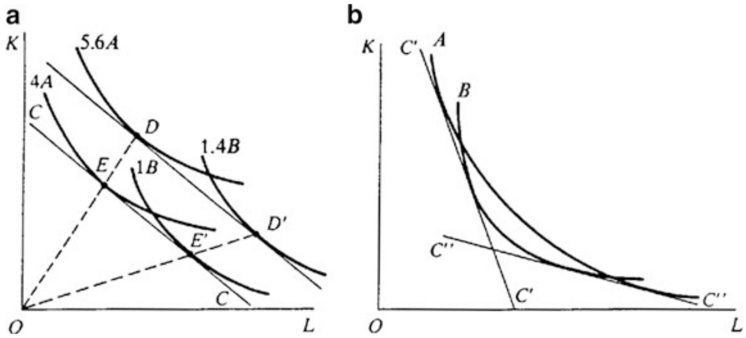


Fig. 4.3 Relative price of goods and relative price of factors

isoquants $5.6A$ and $1.4B$ (at D and D' respectively) is parallel to isocost CC (hence it represents the same factor-price ratio). This parallelism derives from the properties of homogeneous functions of the first degree. As we know (see Sect. 19.1.3) the isoquants of these functions have the same slope along any ray from the origin and, furthermore, their index is proportional to their distance from the origin (an isoquant twice as far from the origin represents a quantity twice as great). The space included between the two rays starting from the origin and passing through E and E' is called by Chipman (1966, p. 23) a "cone of diversification". Only one such cone exists in the absence of factor-intensity reversals; two or more of them will exist in the presence of reversals.

The correspondence between relative prices of factors and relative prices of goods is *one-to-one*, which means that a *unique* commodity-price ratio corresponds to each admissible factor-price ratio. In fact, the reasoning made above to pass from the relative price of goods to the relative price of factors can be inverted. Graphically, if we consider any family of isocosts with the same slope (for example that to which CC belongs), then each of them must necessarily determine a unique couple of isoquants simultaneously tangent to it and representing quantities of goods in the ratio of $4A$ to $1B$. On the contrary, in Fig. 4.3b, where the isoquants cut twice (that is, a factor intensity reversal is present, as explained above), there are two isocosts ($C'C'$ and $C''C''$) with the property of being simultaneously tangent to the isoquants $4A$ and $1B$: the factor-price ratio corresponding to the given commodity-price ratio is not unique.

We conclude this section by examining the behaviour of the relationship between the relative price of goods and the relative price of factors both with and without factor-intensity reversals. In the latter case such a relationship is monotonic, in the former it is not. Let us consider Fig. 4.4, which reproduces Fig. 4.3a, and assume that the relative price p_B/p_A shifts from 4 to 5, so that we must now find the isocost simultaneously tangent to the isoquants $5A$ and $1B$. As can be seen, a greater factor-price ratio p_L/p_K corresponds to the greater commodity-price ratio p_B/p_A, because $\tan\beta > \tan\alpha$. Since, as shown above, the correspondence is one-to-one, we can conclude that as the relative price of labour (p_L/p_K) increases, the relative

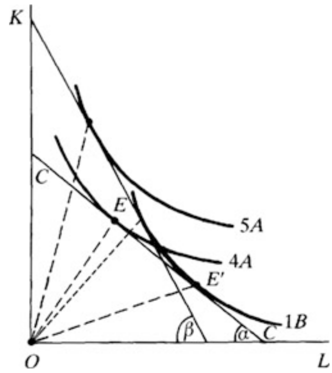


Fig. 4.4 Change in the factor-price ratio following a change in the commodity-price ratio

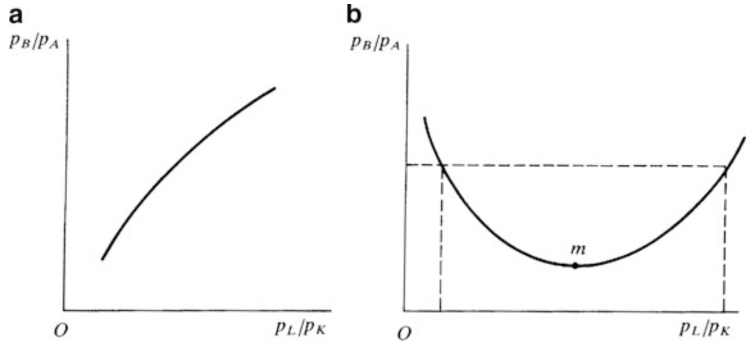


Fig. 4.5 Various relationships between relative price of factors and relative price of goods

price of commodity B (which is the labour-intensive commodity) increases. This relationship is drawn in Fig. 4.5a. We must note that it is monotonically *increasing* because we have assumed that B is the labour-intensive commodity; in the opposite case it would be monotonically *decreasing*; but in either case it is monotonic.

In the presence of factor-intensity reversals, the relationship under consideration, as we known, is no longer one-to-one, as two (or more, according to the number of reversals) factor-price ratios will correspond to any given commodity-price ratio. A case in which there is only one reversal is represented in Fig. 4.5b, where point m corresponds to the factor-price ratio at which the factor-intensity reversal occurs.

4.2 The Heckscher-Ohlin Theorem

The basic proposition of the Heckscher-Ohlin model is the following:

Theorem (Heckscher-Ohlin). *Each country exports the commodity which uses the country's more abundant factor more intensively.*

The concept of (relative) factor intensity has been clarified in Sect. 4.1; it is now the turn of the concept of (relative) *factor abundance*. The definition that immediately comes to mind is in *physical terms*: we say that a country (say country 1) is abundant in one factor (say capital) relative to the other, or that country 1 is relatively more endowed with capital than country 2, if the former country is endowed with more units of capital per unit of labour relative to the latter: $K_1/L_1 > K_2/L_2$, where K_1 is the total amount of capital available in country 1, etc.

An alternative definition is however possible, which makes use of the relative price of factors and is therefore called the *price definition*: country 1 is said to be capital abundant, relative to country 2, if capital is relatively cheaper (with respect to labour) in the former than in the latter country, at the (pre-trade) autarkic equilibrium, namely $p_{1K}/p_{1L} < p_{2K}/p_{2L}$, where p_{1K} is the price of capital in country 1, etc.

It is obvious that the physical definition reflects relative physical abundance, whilst the price definition reflects relative *economic* abundance. Since, thanks to the simplifying assumptions made at the beginning of Sect. 4.1, the Heckscher Ohlin theorem can be demonstrated with both the physical and the economic definition, we shall not claim the superiority of either one. Here we shall use the physical definition; the economic definition will be used in Sect. 4.5.1 where a brief discussion of the two definitions will also be given.

In the following treatment, we assume that commodity A is capital intensive relative to commodity B and that country 1 is capital abundant relative to country 2; it goes without saying that B is labour intensive relative to A and 2 labour abundant relative to 1. Thus we must prove that country 1 will export commodity A whilst country 2 will export commodity B.

The first step (a lemma) *in our proof is to show that—at the same commodity-price ratio—a country abundant in one factor has a production bias in favour of the commodity which uses that factor more intensively* namely, in our case, that country 1 has a production bias in favour of A whilst country 2 has a production bias in favour of B.

This can be shown by using the transformation curves or production-possibility frontiers (see Sect. 3.1) of the two countries; their relative position reflects the fact that country 1 is capital abundant relative to country 2 and that commodity A is capital intensive relative to commodity B (see Fig. 4.6). An alternative geometric treatment in terms of Edgeworth-Bowley boxes can be found in Lancaster (1957). It should be noted that it is not necessary for the two curves to intersect: what matters is that they have a *different* slope along any ray through the origin. If relative factor endowments were the same in both countries, then their transformation curves would have the *same* slope (that is, an identical opportunity cost) along any ray through the origin (in other words, they would be radial blow-ups of each other); similarly, the ratio of the outputs in the two sectors would be the same in both countries at any given common commodity-price ratio. In such a situation, given the assumption of identical structures of demand, there would be *no* scope for international trade.

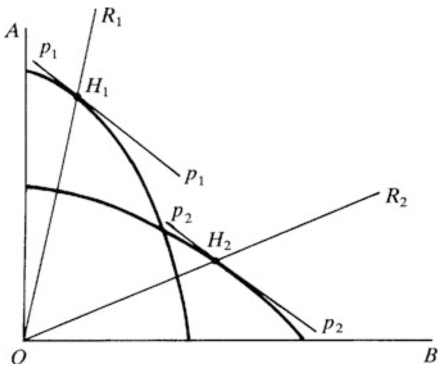


Fig. 4.6 Transformation curve and the Heckscher-Ohlin theorem

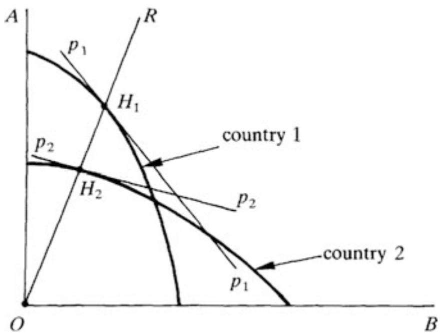


Fig. 4.7 Transformation curve and the Heckscher-Ohlin theorem: an alternative diagram

Let us consider a pre-trade (i.e. autarkic) situation and take a given commodity-price ratio which is identical in both countries ($p_1 p_1$ and $p_2 p_2$ are parallel, thus denoting the same price ratio p_B/p_A). Country 1 is at point H_1 on its own transformation curve and country 2 at point H_2. It can immediately be seen that, at the same relative price of goods, the ratio of the output of A to the output of B is greater in country 1 than in country 2 because the slope of OR_1 is greater than the slope of OR_2. This property holds for any common relative price of goods. An alternative way of looking at the same thing is based on Fig. 4.7. Let us consider, as before, a pre-trade situation and examine a given ratio of A–B, identical in both countries, for example, that represented by the slope of OR. Country 1 would then be at point H_1 on its transformation curve and country 2 at point H_2. The marginal rate of transformation is greater in country 1 than in country 2 (computed at H_1 and H_2 respectively). Commodity A would then be relatively cheaper in country 1 than in country 2, and vice versa for commodity B (we must bear in mind that in equilibrium the marginal rate of transformation coincides with the commodity-price ratio p_B/p_A). In other words, the *opportunity cost* of A in terms of B is lower in country 1, the capital-abundant country, has a production bias in favour of the capital-intensive commodity A, whilst the labour-abundant country 2 has a production bias in favour of the labour-intensive commodity B, in the sense that each country can expand its production of the commodity which is intensive in the country's abundant factor, at a lower opportunity cost than the other.

It is now easy to show that *each country exports the commodity which uses the country's more abundant factor more intensively.* This follow from the lemma and from the assumption that the structure of demand is identical in both countries (and independent of the level of income). In fact, with free trade and no transport costs, the commodity-price ratio (terms of trade) is the same in both countries. Now, according to the lemma, at the same relative price of goods country 1 (the capital-abundant country) will produce relatively more A (the capital-intensive commodity) and country 2 (the labour-abundant country) will produce relatively more B (the labour-intensive commodity): the ratio A/B is greater in country 1 than in country 2. But, given the assumption as to the structure of demand, at the same relative price of goods both countries wish to consume A and B in the same proportion: it follow that country 1 will export A (and import B, which will be exported by country 2) so that after trade the structure of the quantities of the goods available (the quantity available is given by domestic output plus imports or less exports) turns out be identical in both countries and equal to the structure of demand. This completes the proof.

As a spin-off *the terms of trade* will be determined, in much the same way as in Sect. 3.3, Fig. 3.6 (and will lie between the autarkic commodity-price ratios of the two countries)—we call it a "spin-off" because the main point of the Heckscher-Ohlin theory is to prove the basic proposition on the pattern of trade rather than to determine the terms of trade. This not surprising, because—as we have already noted in Sect. 3.7, and as is now obvious from the treatment in the present chapter—the Heckscher-Ohlin theory can be considered, from the purely analytical point of view, as a particular case of the neoclassical theory in which production functions and structures of demand are assumed to be internationally identical.

4.3 Factor Price Equalization

We propose now to show that if there is incomplete specialization the Heckscher-Ohlin model gives rise to *factor-price-equalization* (henceforth FPE); this result is usually stated as follows:

Theorem (FPE). *International trade in commodities and incomplete specialization, under the assumptions of the Heckscher-Ohlin model and notwithstanding the international immobility of factors, equalizes relative and absolute factor prices across countries.*

It should be stressed that the equalization concerns not only *relative* factor prices (p_L/p_K), but also *absolute* factor prices, that is, $p_{1L} = p_{2L}, \quad p_{1K} = p_{2K}$. To prove FPE we shall assume that international trade *does not bring about complete specialization,* so that each country continues to produce both goods; it is important to stress that this assumption, which is *additional* to those at the basis of the Heckscher-Ohlin theorem, is necessary to demonstrate the theorem under consideration.

Let us first recall from Sect. 4.1 that, thanks to the assumption of no factor intensity reversals, there is a one-to-one correspondence between the relative price of goods and the relative price of factors, which is the same in both countries. Secondly, with free trade, no transport costs, etc., the same good must have the same price in both countries (the *law of one price*), so the relative price of goods is the same in both countries. It follows that the *relative price of factors* is identical in both countries.

To arrive at *absolute factor price equalization* (which is what interests us) some more groundwork is necessary.

As a consequence of the identity between the relative price of factors and of the assumptions on technology, the optimum input combination in each sector is the same in both countries (but for a factor of scale): in other words, $(K/L)_{1A} = (K/L)_{2A}$ and $(K/L)_{1B} = (K/L)_{2B}$, as can also be read off Fig. 4.2a. With constant returns to scale, marginal productivities depend solely on the factor input ratio (see Sect. 19.1.3) and are independent of scale. It follows that the marginal productivities of the two factors in the two sectors are identical in both countries, namely

$$\begin{aligned} MPK_{1A} &= MPK_{2A}, \\ MPL_{1A} &= MPL_{2A}, \\ MPK_{1B} &= MPK_{2B}, \\ MPL_{1B} &= MPL_{2B}, \end{aligned} \tag{4.1}$$

where *MPK* and *MPL* denote the marginal productivities of capital and labour respectively, and the subscripts refer to the countries and commodities as usual.

The importance of the assumption of incomplete specialization should be noted here. In fact, if specialization were complete (for example, country 1 produces exclusively commodity A and country 2 commodity B), the quantities MPK_{1B} and MPL_{1B} could not be defined in practice (because commodity B is not produced in country 1), neither could be MPK_{2A} and MPL_{2A} (because commodity A is not produced in country 2); therefore Eq. (4.1) could not be written and the rest of the proof would fall.

Now, under perfect competition the equilibrium condition *value of the marginal product of a factor = price of the factor* must hold. In symbols (remember that p_A and p_B are internationally identical) we have, with reference, for example, to capital,

$$\begin{aligned} p_A MPK_{1A} &= p_{1K}, \\ p_A MPK_{2A} &= p_{2K}, \\ p_B MPK_{1B} &= p_{1K}, \\ p_B MPK_{2B} &= p_{2K}, \end{aligned} \tag{4.2}$$

from which—since the marginal productivities obey (4.1)—it follows that $p_{1K} = p_{2K}$. In a similar way it can be shown that $p_{1L} = p_{2L}$. This completes the proof of FPE.

Better to appreciate the importance of this theorem, it is sufficient to realize that it shows that free trade in commodities is a perfect substitute for perfect international mobility of factors.[2] Note that, if perfect international factor mobility existed as well, then perfect competition would necessarily lead to the full international equalization of factor prices. But in our models of the pure theory of international trade we have assumed an absolute *international immobility of factors* (see Sect. 1.1), so that it might seem that no reason exists for the equalization of their prices, which would not be equal except by sheer chance.

Contrary to this impression, the theorem under consideration shows that FPE, far from being an improbable event, is a necessary consequence of international trade in the assumed conditions. This came as a surprise to the very writers who first gave a rigorous proof of this theorem: see Samuelson (1948, p. 169).

This explains the great deal of attention paid by international trade theorists to this theorem, which can also be given a graphic treatment.

4.3.1 A Graphic Treatment

For this purpose, we bring together in one diagram (the Samuelson-Johnson diagram) the relationships between (K/L) and p_L/p_K. In the upper half of Fig. 4.8 we have reproduced Fig. 4.2a, in the lower half, Fig. 4.5a turned upside down. Given the international identity of production functions etc., Fig. 4.8 can refer to either country.

Let us denote by $\varrho_1 \equiv (K/L)_1, \varrho_2 \equiv (K/L)_2$ the relative factor endowment in the two countries, where $\varrho_1 > \varrho_2$ owing to the assumption that country 1 is capital abundant relative to country 2. The introduction of ϱ_1 and ϱ_2 makes it possible to determine the *admissible range of variation* of relative factor prices (p_L/p_K) in *each country separately considered*. If we consider, for example, country 1, given its relative factor endowment ϱ_1, the relative price of factors can vary between p_1' and p_1''. Note that at point p_1', country 1 would be completely specialized in the production of A. In fact, in general the overall capital/labour ratio is a weighted average of the capital/labour ratios in the two industries, that is (omitting the country subscript)

[2] It is also possible to show that the opposite is true as well, i.e. that perfect international mobility of factors is a perfect substitute for free international trade. In other words, in a hypothetical model in which commodities are immobile (no international trade), but factors are perfectly mobile between countries, the equalization of factor prices (caused by their perfect mobility) will bring about the equalization of commodity prices across countries notwithstanding their immobility. See Mundell (1957) and Sect. 6.8, p. 137. See also Svensson (1984) for an examination of whether goods trade and factor mobility are necessarily substitutes or may be complements in particular cases.

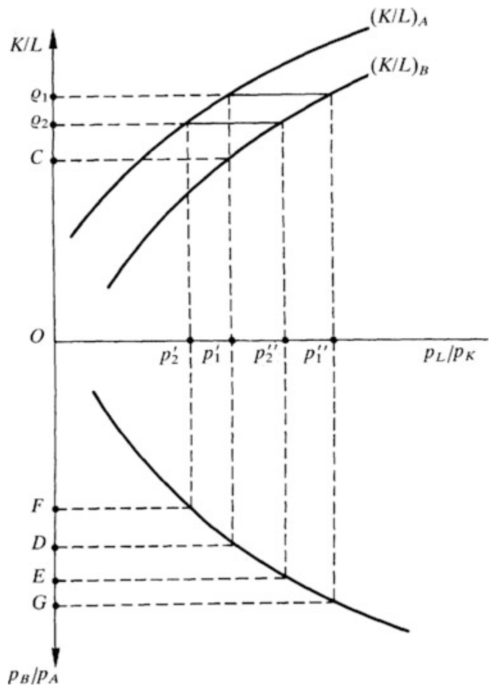


Fig. 4.8 The factor-price-equalization theorem

$$\frac{K}{L} = \frac{L_A}{L}\frac{K_A}{L_A} + \frac{L_B}{L}\frac{K_B}{L_B}, \tag{4.3}$$

where we have used the fact that $K_A + K_B = K$ by the assumption of full employment, which also ensures that the sum of the weights is one (because it also implies $L_A + L_B = L$). Now, if the relative price of factors is p_1', the capital/labour ratio in country 1's industry A is ϱ_1, whilst it would be $C < \varrho_1$ in industry B: but this is not possible, because (4.3) would not be satisfied (the sum of the weights is one); it is therefore necessary for the output of B to be zero in order for (4.3) to hold. It can be shown by similar reasoning that country 1 is completely specialized in B when the relative price of factors is p_1''.

A similar demonstration will show that country 2 is completely specialized in A when $p_L/p_K = p_2'$ and in B when $p_L/p_K = p_2''$; these values delimit the admissible range of variation of relative factor prices. It is now clear that only if the two ranges overlap and so admit of a common part (henceforth called "segment of equalization" for brevity) the equalization of relative factor prices (and so of absolute factor prices, if complete specialization does not occur) will be possible. This segment is $p_1' p_2''$ in our example; from the lower part of Fig. 4.8 it can be seen that the relative price of goods must fall in segment *DE*.

As can readily be seen from the diagram, the farther the relative factor endowments of the two countries are apart, the less probable is the presence of a segment of equalization. If ϱ_1 and ϱ_2 are so distant as to exclude the presence of such a segment, there will be complete specialization in at least one country and even the relative factor price equalization will be impossible. In general, various cases can be distinguished and classified as follows:

(a) A segment of equalization exists, and at the pre-trade equilibrium the relative prices of goods in the two countries are such that the corresponding relative prices of factors fall in this segment (in terms of Fig. 4.8, the relative prices of goods fall in *DE* in both countries before trade). In this case the equalization of the relative price of goods due to international trade brings about the equalization of the relative price of factors. To show this, we first observe that (terms of trade) that comes about as a consequence of international trade necessarily falls strictly between the two pre-trade equilibrium relative prices. In fact, if the terms of trade were equal to the pre-trade equilibrium relative prices of either country, this country would not obtain any benefit from trade and would not engage in international trade. If the terms of trade were lower than the smaller or higher than the greater pre-trade equilibrium price ratio, then one country would suffer a loss. As a matter of fact, we have shown during the analysis of the neoclassical theory (of which the Heckscher-Ohlin model can be considered as a particular case) that the terms of trade are always strictly included between the two autarkic equilibrium price ratios: see Figs. 3.6 and 3.7b.

    Given that the relative price of goods strictly falls between the two pre-trade equilibrium relative prices, the corresponding factor-price ratio must necessarily fall within the segment of equalization. Now, since specialization is not complete (the extreme points of the segment, which give rise to complete specialization, are in fact excluded), absolute factor price equalization will also occur.

(b) No segment of equalization exists. In this case complete specialization of at least one country is inevitable and even relative factor price equalization is excluded. In terms of Fig. 4.9, before trade, country 1's relative price of goods was in *DG* (for example at G') and country 2's was in FE (for example at F'), with the corresponding relative price of factors in $p_1' p_1''$ and $p_2' p_2''$ respectively. After the opening of trade, the (common) relative prices of goods will be included between G' and F': it may fall in $F'E$ or in ED, or in DG'.[3] If it falls in $F'E$, for example at point H, country 2 will produce both goods and the relative price of factors there will be p_{2H}. Country 1, on the contrary, will specialize completely in commodity A and the relative price of factors there will be p_1': it must, in fact, be stressed that, *when complete specialization obtains*, we can no longer use the one-to-one relation between relative factor prices and relative goods prices (which presupposes that both goods are produced domestically) and so—no matter what the terms of trade are—the relative price of factors will be that corresponding to the point of full specialization.

It can be checked by similar reasoning that, if the terms of trade fall in *ED*, country 1 will completely specialize in A and country 2 in B (the relative prices of factors will be p_1' and p_2'' respectively), whilst if they fall in DG', country 1 will produce both commodities and country 2 will completely specialize in B (the relative prices of factors will be: included between p_1' and $p_{1G'}$, and equal to p_2'', respectively).

[3] It cannot fall at F' or G' because, as stated repeatedly, the terms of trade cannot be equal to either pretrade autarkic equilibrium price ratio.

(c) A segment of equalization exists, but the pre-trade equilibrium relative prices of goods are *not* such as to make both countries' relative prices of factors fall within it: in terms of Fig. 4.8, $(p_B/p_A)_2$ is, for example, included in *FD*, whilst $(p_B/p_A)_1$ is, for example, included in EG. After trade, the terms of trade will be included between these price ratios as usual, but the outcome will be different, depending on where the terms of trade themselves happen to fall. If they fall in *DE* (excluding the extreme points D and E), both the relative and the absolute factor prices will be equalized as in case (a). But they may equally well fall in *FD* or in *EG*: in both instances the result will be the same as in case (b), that is one country will completely specialize (both cannot, however) and factor price equalization will be impossible.

Since in case (c)—differently from cases (a) and (b)—it is important to know the exact position of the terms of trade (an information that we can get only by exactly knowing the demand side), we must conclude that also in case (c) the result is, in general, ambiguous.

We conclude this section with three observations. Firstly, the essential role played by the assumption of absence of complete specialization in the factor price equalization theorem must be stressed again. Secondly, the presence or absence of a segment of equalization is related to the spread between the relative factor endowments of the two countries: as we have seen, the more distant ϱ_1 and ϱ_2 the more probable—*ceteris paribus*—the absence of such a segment and the complete specialization of at least one country. Thirdly, it is always possible, even in the absence of full factor price equalization, to state that international trade brings about a *tendency* to relative factor price equalization: it can in fact be readily seen from Fig. 4.9 that, *after* trade, the relative factor prices will, in any case, be closer than before trade.

4.3.2 *The Factor Price Equalization Set*

The FPE theorem has been proven in Sect. 4.3.1 under the assumption of incomplete specialization. In this section we address a related question, namely, under what conditions the free trade equilibrium is one of incomplete specialization. We have seen above that the degree of international specialization is positively related to differences in relative factors endowments. The conditions we search for will therefore concern such differences. To investigate this matter we follow the thought experiment proposed by Samuelson (1949, 1953) and known as the *integrated world equilibrium* approach.

Consider first a world economy constituted by only one country with endowments given by $\overline{K}$ and $\overline{L}$. The equilibrium for this economy, necessarily a closed economy, is identified by the equilibrium price of factors (p_K^*, p_L^*) and equilibrium value of output (A_W^*, B_W^*). We refer to this equilibrium as the integrated equilibrium. Imagine splitting this single-country world economy into a two-country world economy with free trade between them. The split is operated by arbitrarily allocating a portion of the world endowments to each country so as to exhaust world endowments. Clearly, there is an infinity of possible such allocations. We search for

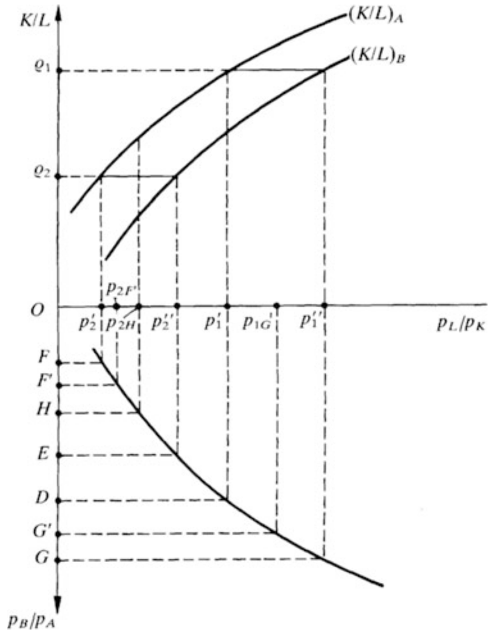


Fig. 4.9 A case of no factor price equalization

allocations such that the equilibrium factor prices of the two-country world economy are the same as those of the integrated equilibrium and such that the outputs of A and B in each country are positive and sum up to A_W^* and B_W^*, respectively. If such allocations exist they will be, by construction, allocations characterized by FPE and incomplete specialization. We therefore will have found the conditions on countries' factor endowments under which a free trade world economy implies incomplete specialization and FPE.

This thought experiment has a simple graphical representation[4] which we depict in Fig. 4.10.The base and height of the rectangle represent world endowments $\overline{L}$ and $\overline{K}$, respectively. The diagonal represents the vector of world endowments. Recall that in the integrated economy $\overline{L}$ and $\overline{K}$ are employed to produce the quantities A_W^* and B_W^*.

The next step is to find the sectorial employment vectors of the integrated equilibrium. The elements of such vectors are the employment of L and K in each industry, denoted L^A, $K^A$and L^B , K^B. Naturally, the sum of the sectorial employment vectors gives the endowment vectors, that is, $L^A + L^B = \overline{L}$ and $K^A + K^B = \overline{K}$. Recall that the slope of a vector is given by the ratio of its elements, thus, for instance, the slope of the sectorial employment vector measured on the L-axis

[4]This diagram is commonly attributed to Dixit and Norman (1980, pp. 109ff.), but earlier presentations can be found in Travis (1964, pp. 15ff.) and Lancaster (1957, pp. 31ff.).

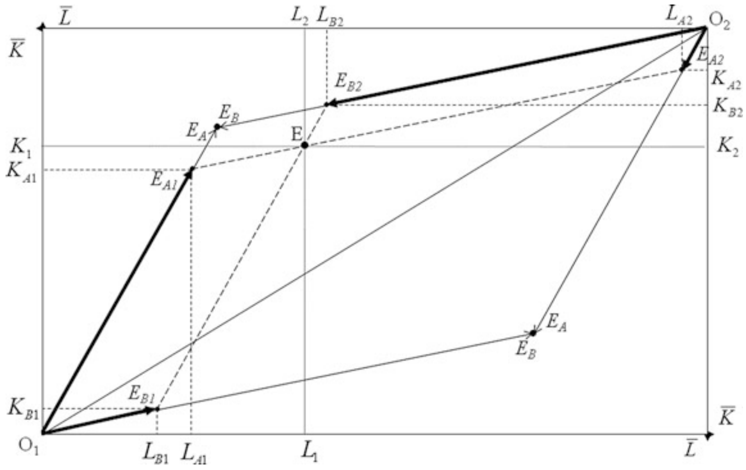


Fig. 4.10 The factor price equalization set

for industry A is given by K^A/L^A which is the capital intensity in industry A. The integrated equilibrium yields the equilibrium values of factor prices from which, as studied in Sect. 4.1.1, we can determine the factor intensity in each industry. Given the factor intensities we can draw the sectorial employment vectors by drawing two segments (from whichever corner) whose slopes correspond to the factor intensities and whose length is such that the sum of them give the vector of world endowments. In Fig. 4.10, the vectors O_1E_A and O_1E_B (or, equivalently, O_2E_A and O_2E_B) represent the sectorial employment vectors. Recalling the parallelogram rule of vectorial sum it is immediate that the sum $O_1E_A + O_1E_B$ gives the world endowment vector. The length of each sectorial employment vector is proportional to factor inputs and, therefore, reflects the quantity of output; the longer the sectorial employment vector the larger the sector output.

We now split the world into two countries and assume free trade between them. We refer to this equilibrium as to the free trade equilibrium. Any point in the rectangle represents a possible division of the world economy into two countries but not all the divisions will give rise to the free trade equilibrium we require. We require the free trade equilibrium to have the same factor prices and, therefore, the same factor intensities as the integrated equilibrium. We also require both countries to produce both goods.

To find any of the divisions satisfying the requirements we begin by arbitrarily assigning to each country a share of the integrated equilibrium outputs. This assures incomplete specialization. Let s_{A1}, s_{B1}, $s_{A2} = 1 - s_{A1}$, and $s_{B2} = 1 - s_{B2}$ be the arbitrarily chosen shares. The next step is to find the sectorial employment vectors for each country, denoted E_{A1} and E_{B1} and E_{A2} and E_{B2}. They must have the same slope as the sectorial employment vectors of the integrated equilibrium (because we require same factors price) and their length must be such that the corresponding factor inputs give outputs equal to $A_1^* = s_{A1} A_W^*$, $B_1^* = s_{B1} B_W^*$, $A_2^* = (1 - s_{A1})\, A_W^*$, $B_2^* = (1 - s_{B1})\, B_W^*$. Given these requirements, each country's employment vectors are a fraction of the integrated equilibrium employment vectors. Precisely, $E_{A1} = s_{A1} E_A$, $E_{B1} = s_{B1} E_A$, $E_{A2} = (1 - s_{A1})\, E_A$, $E_{B2} = (1 - s_{B1})\, E_B$. Let O_1 be the origin of measures for country 1 and O_2 that for country 2 in the two-country world economy. These vectors are shown in Fig. 4.10 as fractions of the world sectorial employment vectors. The sum of $O_1 E_{A1}$ and $O_1 E_{B1}$ gives the vector $O_1 E$ which represents total employment in country 1. Because of full employment, this vector necessarily represents the endowment vector for country 1 consistent with the arbitrarily chosen shares. Analogously, $O_2 E$ represents the resulting endowment vector for country 2. The vectors $O_1 E$ and $O_2 E$ have precisely the properties we have required; they represent an allocation of world endowments such that the resulting free trade equilibrium yields the same factor prices as the integrated equilibrium (i.e., FPE) and such that each country is incompletely specialized. It is easy to verify that the set of all such possible allocations is constructed as the sum of all possible fractions of the world employment vectors and, therefore, is represented graphically by the area demarcated by the parallelogram composed by the four vectors $O_1 E_A$, $O_2 E_B$, $O_2 E_A$, $O_1 E_B$. Any point inside the parallelogram represents a division of the integrated world economy such that factor prices equalize and countries are incompletely specialized. The borders of the parallelogram belong to the FPE set but imply that at least one country is completely specialized.

An alternative but equivalent way of constructing the FPE set is the following. Pick a point in the rectangle and consider the vector drawn from O_1 to the chosen point. If this vector can be decomposed into two vectors which are portions (including 0 and 1) of the integrated equilibrium employment vector then the chosen point belongs to the FPE set. If such decomposition is impossible then the chosen point does not belong to the FPE set. Naturally, drawing the vector from O_1 or O_2 is equivalent. Clearly, only points in the parallelogram allow to draw vectors that can be decomposed into portions of the integrated equilibrium employment vectors.

The conclusion of the analysis can be summarized in the following

Theorem (Factor Price Equalization Set). *The Factor-Price Equalization Set is the set of all weighted sums of the integrated-equilibrium employment vectors, where the weights take values between zero and one.*

4.4 The Factor Content of Trade and the Heckscher-Ohlin-Vanek Theorem

The Heckscher-Ohlin theorem states that each country exports the commodity which uses the country's more abundant factor more intensively. This theorem may be reformulated in terms of the *factor content of trade*. This reformulation, due to Vanek (1968), is instructive as it allows seeing the Heckscher-Ohlin theory in a different perspective and permits discussing some generalization in a simple way.

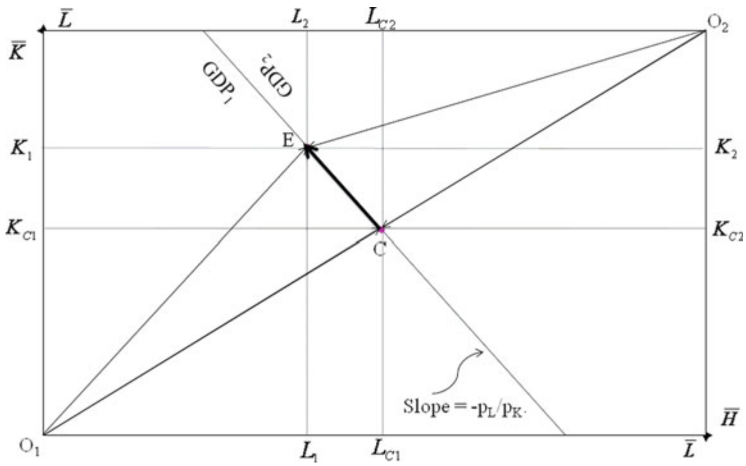


Fig. 4.11 The factor content of trade

The factor content of trade is defined as the quantity of factors used to produce the goods exported *minus* the quantity of factors used to produce the goods imported. We refer to these quantities as the factor services embodied in net trade. Thus, for instance, for a country who exports ten units of A and imports ten units of B, the factor content of trade is given by the factor services embodied in the ten units of A exported minus the factor services embodied in the ten units of B imported. Noting that net exports are given by production minus consumption, the factor content of trade may equivalently be defined as the vector representing the factor services embodied in the goods produced by the country minus the vector representing the factor services embodied in the production of the goods consumed by the country. The reformulation under examination predicts that in free and balanced trade the sum given by the capital services embodied in exports minus the capital services embodied in imports is positive for the capital abundant country and negative for the labour abundant country; signs reversed for L. More generally, we have the following:

Theorem (Heckscher-Ohlin-Vanek). *Each country is the net exporter of the services of its abundant factor and the net importer of the services of its scarce factor.*

Net exports are given by production minus consumption. Therefore, the factor content of trade is simply the vector representing the factor services embodied in the goods produced by the country minus the vector representing the factor services embodied in the production of the goods consumed by the country. The factor content of trade vector has a simple graphical representation.Figure 4.11 represents the free trade world economy discussed in Sect. 4.3.2, where O_1E and O_2E are the endowment vectors of country 1 and 2, respectively, and point E is assumed to be within the Factor Price Equalization set (not shown in the figure). Comparing the slopes O_1E and O_2E with respect to the $L-$axis shows that country 1 is relatively capital abundant. Since there is full employment, the endowment vectors O_1E and O_2E are also the total employment vectors. A line emanating from E whose slope is given by the relative price of L represents the GDP line (or budget constraint line) of each country since it is obtained by multiplying factor endowments by factor prices. The diagonal represents the vector of factor services embodied in the production the goods consumed by the world economy.

Since preferences are identical and homothetic and since trade is balanced, each country will consume a share of world production of goods equal to its share of world income. Therefore, the vector representing factor services embodied in the goods consumed by a country will necessarily lie on the diagonal and will necessarily be a fraction of it. Its length is given by the intersection of the GDP line with the diagonal. In Fig. 4.11, O_1C and O_2C represent such vectors for country 1 and 2, respectively. The vector CE is the factor content of trade vector for country 1. Its elements are $(L_1 - L_{C1}) < 0$ and $(K_1 - K_{C1}) > 0$. The first element is negative and the second is positive reflecting the fact that country 1 is relatively well endowed of K. The vector EC is the factor content of trade vector for country 2. In conclusion, we have found that, as stated in the Heckscher-Ohlin-Vanek theorem, each country is the net exporter of the services of its abundant factor and the net importer of the services of its scarce factor.

It is interesting to note that the H-O-V theorem does not require any information about the output of goods in each country or about the direction of trade. This will be important when discussing a generalization of the Heckscher-Ohlin theory in which there are more goods than factors. It will also be important when addressing the empirical verifications of the Heckscher-Ohlin theory.

4.5 Extensions and Qualifications

This section aims at analysing the consequences of dropping some of the basic assumptions examined in Sect. 4.1, in particular that concerning the structure of demand, that concerning the absence of factor-intensity reversals, and that concerning the presence of only two commodities and two factors. The assumption of internationally identical production functions cannot be dropped without altering the essence of the Heckscher-Ohlin theory. A list of studies on the extensions and qualifications of the Heckscher-Ohlin model includes Baldwin (2008); Bhagwati (1972); Bhagwati and Srinivasan (1983); Brecher and Choudhri (1982); Chacholiades (1978); Davis et al. (1997); Deardoff (1982); Dixit and Woodland (1982); Ethier (1982, 1984); Feenstra (2004); Hamilton and Svensson (1984); Harkness (1978, 1983); Harrod (1953); Helpman (1984a); Herberg et al. (1982); Horiba (1974); Johnson (1957); Leontief (1956); Maskus (1985); Neary (1984, 1985b); Pearce (1952); Samuelson (1967); Sarkar (1984); Takayama (1972).

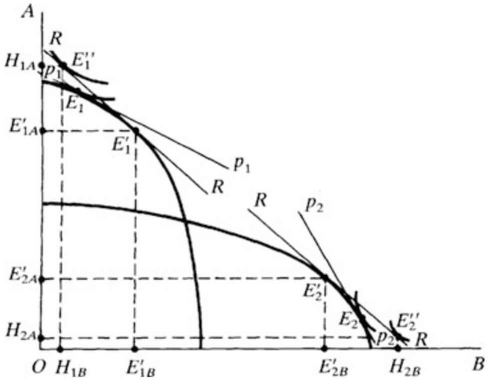


Fig. 4.12 Non-identical structures of demand

4.5.1 Non-identical Structures of Demand

If we drop the assumption of internationally identical structures of demand, the Heckscher-Ohlin proposition is no longer necessarily true. In fact, if a country has a strong preference for the commodity which uses the country's more abundant factor more intensively (remember that we are using the physical definition of factor abundance), it may happen that, when trade opens up, each country exports the other commodity, namely the one which is intensive in the country's less abundant factor. This is illustrated in Fig. 4.12, where the transformation curves and the social indifference curves of the two countries are brought together, in the same way as in Fig. 3.15.[5] The pre-trade equilibriumpoints are E_1 and E_2 for country 1 and country 2 respectively; the corresponding relative prices of goods are measured by the (absolute value of the) slope of p_1p_1 and p_2p_2. After trade begins, an intermediate price ratio (terms of trade) will obtain, for example, that measured by the slope of RR. The production point will be E_1' for country 1 which, however, given its strong preference for commodity A, will consume at E_1'', importing $E_{1A}'H_{1A}$ of A and exporting $E_{1B}'H_{1B}$ of B. Therefore, country 1 will import the commodity intensive in capital (the country's more abundant factor) and export the commodity intensive in labour (the country's less abundant factor). Similarly it can be seen that country 2 will produce at point E_2' and consume at point E_2'', importing $E_{2B}'H_{2B}$ (equal to $E_{1B}'H_{1B}$) of B and exporting $E_{2A}'H_{2A}$ (equal to $E_{1A}'H_{1A}$) of A: commodity B is intensive in labour (country 2's more abundant factor) and A in capital (country 2's less abundant factor). Thus the Heckscher-Ohlin proposition is contradicted.

[5]We refer the reader to that chapter for the problems related to the use of social indifference curves. With the occasion, we point out that Fig. 4.12 makes it possible to show the gains from trade in the same way as in Fig. 3.15. In the case of the Heckscher-Ohlin model with identical structures of demand, we can use the same diagram with the proviso that an *identical* family of social indifference curves (which, in addition, must be homothetic) must be used for both countries.

It should, however, be noted that this result *may*, and need not, occur: it is, in fact, possible—as the reader can ascertain graphically by experimenting with different families of social indifference curves—that the basic proposition remains valid even with different structures of demand, provided that, in each country, these are not too much biased towards the commodity which uses the country's more abundant factor more intensively. We can therefore conclude that the assumption of identical structures of demand is a sufficient, but not a necessary, condition for the validity of the Heckscher-Ohlin theorem.

It is important to stress that the possible invalidity of this theorem, because of different structures of demand, *does not invalidate* the factor price equalization theorem, which continues to hold within the limits clarified in the previous section. The latter theorem, in fact, does not depend on the assumption of identical demand structures, and as long as no factor-intensity reversal occurs and specialization is incomplete, the theorem under consideration remains valid.

However, the possible invalidity of the Heckscher-Ohlin theorem when demand structures are different, has led various authors to investigate the possibility of reformulating the theorem without that assumption. The answer is that it can be done, *provided that the Heckscher-Ohlin theorem is reformulated in terms of the price definition of factor abundance* (see Sect. 4.2). The reason is intuitive: in country 1, in the pre-trade autarkic equilibrium situation, the strong bias of tastes towards the capital-intensive commodity A implies that this factor, notwithstanding its relative abundance in physical terms, will be relatively scarce (less abundant) in economic terms, namely, will have a greater relative price than in country 2, where exactly the opposite situation obtains. Thus we shall have

$$p_{1K}/p_{1L} > p_{2K}/p_{2L}, \text{namely } p_{1L}/p_{1K} < p_{2L}/p_{2K}, \tag{4.4}$$

and so, in terms of the price definition of factor abundance, country 1 is *labour*-abundant relative to country 2. More rigorously, (4.4) can be arrived at by way of the one-to-one correspondence between relative factor prices and relative prices of goods. Figure 4.12 tells us that, in the pre-trade equilibrium situation, $(p_B/p_A)_1 < (p_B/p_A)_2$. Therefore—see Fig. 4.5a—we have $(p_L/p_K)_1 < (p_L/p_K)_2$, as was to be shown.

In conclusion, the Heckscher-Ohlin theorem is valid independently of the structure of demand (thus assumption 2 of Sect. 4.1 can be dropped), *if* the price definition of factor abundance is adopted. This is one of the motives which have induced some writers to prefer the price to the physical definition. It is interesting to point out that Ohlin himself used the price definition of abundance, though hinting at a physical definition: "... the real problem is to demonstrate what lies behind such inequality in prices, or, more precisely, to show in what way *differences in equipment* come to be expressed in differences in money *costs and prices*" (Ohlin, 1933, p. 13; p. 7 of the 1967 edition. Our italics).

However, arguments for the physical definition are not lacking. Relative factor abundance in physical terms is observable at any moment (provided of course that the factors can be measured unambiguously, but this is a general problem). On the contrary, relative factor abundance in price terms is not observable, as it is defined with reference to a hypothetical pre-trade autarkic equilibrium situation. Some authors (see, for example, Leamer, 1984, p. 2) even think that hypotheticals such as autarkic prices, that have no observable counterpart, are to be excluded from discussion.

4.5.2 *Factor-Intensity Reversals*

To investigate the consequences of the presence of factor-intensity reversals it is expedient to use the diagram which brings together the relationships between the capital/labour ratio and the factor-price ratio, and between the latter and the commodity-price ratio. We have reproduced Fig. 4.2b in the upper half of Figs. 4.13 and 4.5b, turned upside down, in the lower half.

Various cases must now be distinguished, according to the position of the relative factor endowments of the two countries. If these endowments are such that, in the interval between them, no factor-intensity reversal occurs, as is the case of ϱ_1 and ϱ_2, then the Heckscher Ohlin theorem remains valid, for any factor intensity reversal occurring *outside* the $\varrho_1 - \varrho_2$ interval is irrelevant: in the relevant stretch, commodity A is unambiguously capital-intensive relative to commodity B (in terms of Fig. 4.1b, only the part of the diagram to the left of the radiant of tangency must be considered). The factor price equalization theorem also remains valid (within the limits in which it is valid in general: existence of a segment of equalization, etc.).

If, on the contrary, relative factor endowments are separated by a point of factor-intensity reversal (as is the case of ϱ_1 and ϱ_2' in Fig. 4.13), then exportables have the same kind of factor intensity in both countries, so that the Heckscher Ohlin theorem is no longer valid or, to be precise, remains valid for one country only. Let us assume that the pre-trade equilibrium relative prices of commodities are $(p_B/p_A)_1$ and $(p_B/p_A)_2$; as we know, the terms of trade will fall at an intermediate point, for example R_s. Country 1 will export commodity A and country 2 commodity B[6]: now, as can be seen from the diagram, in country 1 the capital-intensive commodity is A and, in country 2, the capital-intensive commodity is B (owing to the factor-intensity reversal). Thus the Heckscher Ohlin theorem is valid for country 1, the capital-abundant country relative to country 2, but not for country 2, which is the relatively labour-abundant country. In this case also the factor price equalization theorem is invalid ,as no segment of equalization exists; besides, it can be seen that the relative price of factors moves in the same direction in both countries: from OD

[6]This cannot be directly seen from the diagram, but from an inspection of the transformation curves. More simply, as $(p_B/p_A)_1 > R_s$, country 1 will find it profitable, when trade begins, to give up A in exchange for B and similarly, as $(p_B/p_A)_2 < R_s$, country 2 will give up B in exchange for A.

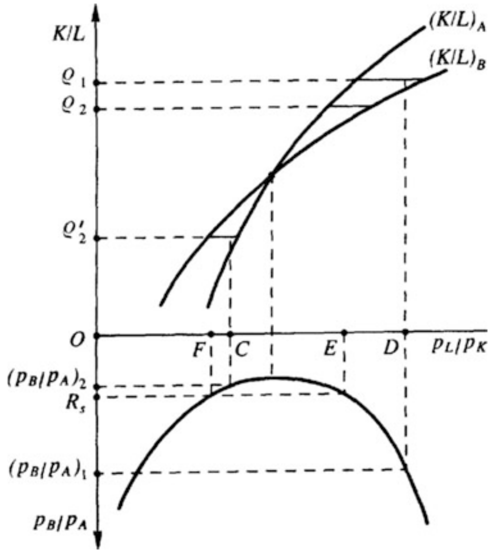


Fig. 4.13 Factor-intensity reversals, H-O and FPE

to OE in country 1 and from OC to OF in country 2. In Fig. 4.13 these movements bring the relative prices of factors nearer, because $DE > FC$, but in general, movements of this kind may equally well bring them farther apart. Therefore, as the relative price of factors moves in the same direction in both countries (either decreasing, as in Fig. 4.13, or increasing), it is no longer possible to state that, in general, there will be at least a tendency towards relative factor price equalization.

In Fig. 4.13 we have examined the case of a single reversal but, as we know, there may be two or more reversals. We give a list of results (which can be derived by graphic analysis):

(a) If there is an odd number of reversals occurring in the interval between the two countries' relative factor endowments, the same conclusions hold as shown above, with reference to Fig. 4.11;
(b) If there is an even number of reversals occurring in the interval between the two countries' relative factor endowments, then each commodity can be unambiguously classified as intensive in a given factor. However, the pattern of trade may not conform to the Heckscher-Ohlin theorem (for example, it may happen that the labour-abundant country exports the capital-intensive commodity). When this occurs, the relative prices of factors will move in opposite directions. On the contrary, when the pattern of trade conforms to the Heckscher-Ohlin theorem, the relative prices of factors will move towards each other, but will never coincide.

4.5.3 The Heckscher-Ohlin-Vanek Generalization

The model studied above is often referred to as the two-by-two version of the Heckscher-Ohlin theory since it counts only two goods and two factors. The two-by-two version is somewhat special since the dimensionality is low (two-by-two) and the number of goods equals the number of factors. In this section we investigate whether the results of the theory are robust to a generalization that allows for many goods and many factors.

Any such generalization gives only three possible dimensional structures: (a) more factors than goods, (b) equal number of goods and factors, (c) more goods than factors. The first dimensional structure is well illustrated by the *specific factor model* that for its importance deserves a separate discussion that we postpone to Sect. 6.2. The second and third dimensional structures may be treated together for our purposes. In what follows let N be the number of goods and M the number of factors and assume that $N \geqslant M \geqslant 2$ with $M > 2$ if $N = M$. This generalization is often called the Heckscher-Ohlin-Vanek generalization.

Consider first the effects of such generalization for the FPE theorem. The FPE set can be constructed using the same logic as in Sect. 4.3.2. Indeed, neither the requirements nor the logic of construction of the FPE depend on the number of goods and factors as long as $N \geqslant M$. Beginning by the integrated world equilibrium we note that it is unaffected by the existence of more goods than factors. The integrated equilibrium system of equations counts the same number of unknowns as there are equations regardless of the number of goods and factors. There will be N efficiency conditions (price = marginal cost), M equilibrium conditions in factor markets, and $N - 1$ equilibrium conditions in commodity markets. These equilibrium conditions determine M factor prices, N commodity outputs, and $N - 1$ commodity prices. The FPE is again given by the sums of all possible fractions of the sectorial employment vectors of the integrated equilibrium where the fractions are arbitrarily chosen shares of the integrated equilibrium outputs. It is useful to provide a graphical representation of the FPE set in this context where $N \geqslant M$. To this purpose consider the simple case where $N = 3$ and $M = 2$. Let A, B and D be goods (the letter C is reserved for consumption) and, as usual, let L and K be factors.The FPE is represented in Fig. 4.14 where the world sectorial employment vectors are O_1A, O_1B, and O_1D. The last two vectors are also represented by the vectors AB' and $B'O_2$. We recall that these sectorial employment vectors are drawn using the information on factor intensities and on total output of goods. Indeed, since output is proportional to inputs, the length of each sectorial employment vector represents (in the space of factors) the total industry output in the integrated equilibrium.

The next step is to choose arbitrarily a partition of the integrated equilibrium outputs keeping factors prices the same as in the integrated equilibrium. The vectors O_1A_1, O_1B_1, and O_1D_1 represent one such partition since they are fractions of the integrated equilibrium sectorial employment vectors. The last two vectors are also represented by the vectors A_1B_1' and $B_1'E$. The corresponding partition

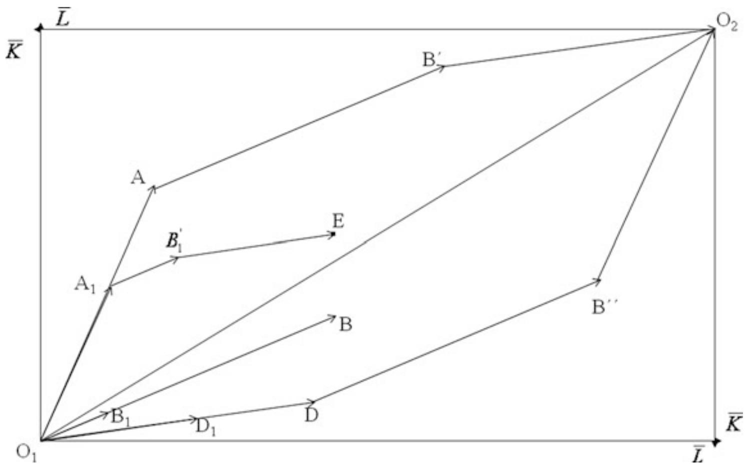


Fig. 4.14 The factor price equalization set: 3 goods and 2 factors

of world endowments is necessarily given by the vector obtained from the sum of O_1A_1, O_1B_1, and O_1D_1, represented by point E in Fig. 4.14. This partition of world endowments between countries satisfies FPE by construction. It is quite clear that the set of all FPE-compatible partitions is obtained from the sum of all possible arbitrarily chosen fractions of the vectors O_1A, O_1B, and O_1D. Graphically the FPE is represented by the area demarcated by the parallelogram $O_1AB'O_2B''DO_1$. In conclusion, the M-by-N generalization where $N \geqslant M$ has no effect on the validity of the FPE theorem.

Coming to the Heckscher-Ohlin theorem, the major nuisance resulting from having more goods than factors is that the model no longer determines the quantities of goods produced in each country. This is clear by noting that while the integrated equilibrium counts the same number of equations and unknowns (regardless of the number of goods and factors) the two-country free-trade equilibrium does not. The latter is composed by N efficiency conditions (price = marginal cost), $2M$ equilibrium conditions in factor markets (M conditions in each country), and $N-1$ equilibrium conditions in commodity markets. The endogenous variables are M factor prices, $2N$ commodity outputs (N in each country), and $N-1$ commodity prices. We therefore have $2M+2N-1$ equations and $M+2N+N-1$ unknowns. Since $N > M$ the free trade equilibrium counts more unknowns than equations, therefore it cannot determine the equilibrium values of all the endogenous variables. In particular, this means that when passing from the integrated equilibrium to the two-country free-trade equilibrium there is an infinity of production structures for the two countries that is consistent with the integrated equilibrium factor prices. This indeterminacy in production is particularly disturbing for the Heckscher-Ohlin theorem since it does not allow to relate output proportions to endowment proportions. Therefore, it is not possible to say which goods are exported by each country.

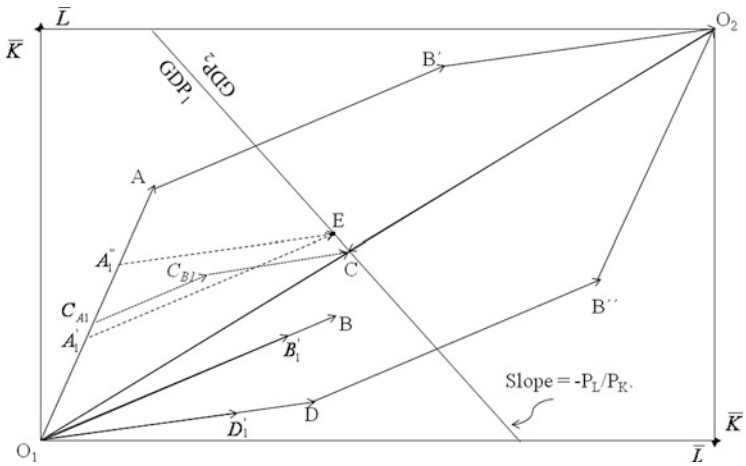


Fig. 4.15 The factor content of trade: 3 goods and 2 factors

However, the Heckscher-Ohlin-Vanek theorem (Sect. 4.4) remains intact.We show this in Fig. 4.15. Consider the endowment point E in the FPE set. This partition of world endowments is compatible with many production structures. In the figure we represents two extreme cases of production structures for country 1 (the choice of the country is irrelevant). The first case has output of A represented by O_1A_1', output of B represented by $O_1B_1' = A_1'E$, and output of D equal to 0. The second case has output of A represented by O_1A_1'', output of B equal to 0, and output of D equal to $O_1D_1' = A_1''E$. The sum of each of these two sets of vectors gives the endowment vector for country 1. As explained above, there is an infinity of equally possible alternatives (one of them is represented by the vectors O_1A_1, O_1B_1, and O_1D_1 in Fig. 4.14).

This example shows the indeterminacy of the production pattern when there are more goods than factors. It becomes impossible to associate relative outputs of goods with relative factor endowments. In addition, it is possible that the *K-abundant* country imports a *K-intensive* good. To see this recall from Sect. 4.4 that the consumption vector lies on the main diagonal and is identified by the GDP line; this is represented by point C in the figure. Then consider the consumption vectors for country 1 represented by O_1C_{A1}, $C_{A1}C_{B1}$, and $C_{B1}C$. Each sectorial consumption vector is in the same proportion to the corresponding integrated equilibrium sectorial employment vector as the vector of total consumption of the country is to the vector of total consumption of the world economy. Therefore, $O_1C_{A1}/O_1A = C_{A1}C_{B1}/O_1B = C_{B1}C/O_1D = O_1C/O_1O_2$. These proportions are required by the assumption of identical and homothetic preferences, which implies that each country consumes a share of output in world output (for each good) equal to its share of income in world income. It is clear that in the first case of production structure (represented by O_1A_1', $O_1B_1' = A_1'E$ and output of $D = 0$) country 1 imports good A in spite of the fact that A is *K-intensive* and country 1 is *K-abundant*. The Heckscher-Ohlin theorem stated in terms of production and export pattern of goods does not survive the $N \geqslant M$ generalization, but the Heckscher-Ohlin-Vanek theorem does. Indeed, the factor content of trade vector for country 1 is the vector CE, and for country 2 it is EC. It is clear that, in spite of the indeterminacy of the production pattern and in spite that the K-abundant country may end up importing the most L-intensive good, each country exports the services of its relatively abundant factor.

4.6 Empirical Studies

4.6.1 Leontief's Paradox

The empirical relevance of the Heckscher-Ohlin theorem has been the subject of very many studies, beginning with the pioneering one of Leontief (1953). By applying his *input–output* analysis[7] to the 1947 input–output table of the US economy, Leontief computed the total (direct and indirect) input requirements of capital and labour per unit of the composite commodity "US exports" and per unit of the composite commodity "US competitive import replacements"; in both cases the unit was one million dollars' worth of commodities at 1947 prices and composition. By "competitive import replacements" Leontief refers to "imports of commodities which can be and are, at least in part, actually produced by domestic industries", so that by replacing a unit of imports with a unit of domestic production, it is possible to find out "whether it is true that the United States exports commodities the domestic production of which absorbs relatively large amounts of capital and little labour and imports foreign goods and services which—if we had produced them at home—would employ a great quantity of indigenous labour but a small amount of domestic capital" (1953, p. 75). The principal findings of this analysis are summarized in Table 4.1, adapted from Leontief (1953):
As can be seen from the last column, it turned out that the United States exported labour-intensive commodities and imported capital-intensive ones. Now, since the United States was generally considered to be a capital abundant country relative to all its trading partners (remember that the data refer to 1947), Leontief's results were in sharp disagreement with the Heckscher-Ohlin theorem (according to which the US ought to have exported capital-intensive commodities), whence the "paradox", as it came to be known in the literature.

[7]For an explicit treatment of intermediate goods in the pure theory of international trade see below, Sect. 6.4.

Table 4.1 Domestic capital and labour requirements per million dollars of US exports and of competitive import replacements (of average 1947 composition)

	Capital (dollars, in 1947 prices)	Labour (man-years)	K/L
Exports	2, 550, 780	182.313	13, 991
Import replacements	3, 091, 339	170.004	18, 184

4.6.2 Explaining the Paradox

Leontief's analysis gave rise to wide debate, concerning both its statistical and theoretical aspects, and to a host of successive empirical studies, which still continue, with conflicting results. It would be impossible to survey this enormous amount of literature here, so we shall focus on some aspects only. Surveys of the initial debate aroused by Leontief's original analysis and of the empirical studies carried out up to the early 1960s are contained in Bhagwati (1964, pp. 21ff.) and Chipman (1966, pp. 44ff.). For subsequent surveys, see Stern (1975), Deardoff (1984), Kohler (1988), Leamer and Levinsohn (1995), Baldwin (2008). See also Leamer (1984), for an original treatment.

By simplifying to the utmost, it is possible to divide the attempts at explaining Leontief's paradox into two groups. The first includes all those works which maintain that serious mistakes or, at the very least, inaccuracies, were made in passing from the theoretical formulation to its empirical testing, so that the latter is vitiated and cannot be considered as a refutation of the Heckscher-Ohlin theorem. The second includes all those works which maintain that one or more of the basic assumptions are not fulfilled in reality, so that the theorem itself loses all validity: the empirical analysis must necessarily confirm this invalidity.

It is self-evident that, whilst the attempts that belong to the first group attempt to rescue the theorem, those belonging to the second are destructive of the theorem itself.

4.6.2.1 Mistakes in Calculations?

Considering the *first group*, we begin with the argument (set forth by Leontief himself, 1953, pp. 87ff.) according to which American labour was—at that time—more efficient than rest-of-the-world labour, so that, when the former was converted into equivalent units of the latter, the United States became a labour abundant country relative to the rest of the world. According to Leontief, it was plausible to assume a coefficient of conversion of three: "...in any combination with a given quantity of capital, one man-year of American labour is equivalent to, say, three man-years of foreign labour. Then, in comparing the relative amounts of capital and labour possessed by the United States and the rest of the world (...) the total number of American workers must be multiplied by three (...). Spread thrice as thinly as the unadjusted figures suggest, the American capital supply per 'equivalent worker' turns out to be comparatively smaller, rather than larger, than that of many other countries” (1953, pp. 87–88).

One must, of course, avoid the logical mistake of attributing the greater efficiency of American labour to the greater amount of capital per man employed in the United States, for by so doing one would commit a tautology; such greater efficiency is, in fact, attributed by Leontief to entrepreneurship, superior organization etc. in the United States relative to other countries. These elements, however, increase not only the productivity of labour but also that of capital, and so if these were to increase by the same proportion, the relative factor abundance would not change. Therefore, Leontief concludes (1953, p. 90), “. . . entrepreneurship, superior organization, and favourable environment must have increased—in comparison with other countries—the productivity of American labor much more than they have increased the efficiency of American capital”.

It should however be noted that subsequent studies did not confirm the coefficient of conversion of three that Leontief assumed. For example, Kreinin (1965) interviewed managers and engineers of about 2,000 US firms operating both at home and abroad, through questionnaires. These aimed at determining the amount of labour time required to produce one unit of the same output—with the same equipment and organization of labour—in plants in the United States and abroad. Most persons interviewed did in fact judge US labour more efficient than its foreign counterpart, but by 20 or 25 %; the resulting coefficient of conversion of 1.20 or 1.25 was far below the coefficient of 3 that, according to Leontief, would have made the USA a relatively labour abundant country.

Other researchers observed, in criticizing Leontief’s study, that it is wrong to consider two factors of production (physical capital and labour) only. For example, according to Diab (1956) and Vanek (1959), one must consider at least another factor, *natural resources*: for instance, the same equipment and the same workers with the same organization operating in the oil extractive industry will obtain better results in Venezuela or in the Arabian countries than in the United States, for the very simple reason that US oil-fields are less rich. Therefore if one neglects the natural resources factor, incorrect results will be obtained, whilst the paradox will disappear if this factor is taken into account. And in fact Vanek (1959), in addition to the data given by Leontief (Table 4.1 above), computed the input of (goods having a high content of) natural resources required to produce one unit of exports and one unit of import replacements: this input turned out to be $340,000 and $630,000 at 1947 prices, respectively. Therefore the United States imported goods intensive in natural resources (no matter whether this intensity was computed relative to capital or to labour), which was the relatively less abundant factor there, and exported goods intensive in capital and labour relative to natural resources (the first two factors being more abundant relative to the third). It followed that the Heckscher-Ohlin theorem, far from being refuted, was fully confirmed.

Other authors stress the importance of the *human capital* factor, which is that embodied in skilled workers, managers, engineers etc. as distinct from general or unskilled labour. Leaving aside the practical problems of the various methods of measuring human capital (capitalization of wage differentials; years of education;

professional qualifications; etc.), the consideration of this factor lends support to the hypothesis that US exports are intensive in human capital (a relatively abundant factor in that country) with respect to import replacements, in accordance with the Heckscher-Ohlin theorem: see, for example, Stern and Maskus (1981), who also cite similar results of previous studies; see also Lane (1985) and Charos and Simos (1988).

An important contribution is that of Casas and Choi (1984, 1985a), who were the first to point out that the Heckscher-Ohlin theorem—as all the theorems in the pure theory of international trade—implicitly presupposes a situation of balanced trade. Since in reality the trade balances are never in equilibrium, the paradoxical empirical results can be due to the non-verification of this essential condition. And in fact they maintain that the same data used by Leontief would have shown, under balance-of-trade equilibrium, that US exports were indeed more capital intensive than import replacements.

Finally it must be pointed out that, according to some writers (e.g. Clifton & Marxsen, 1984; Leamer, 1980; Williams, 1970), the test used by Leontief and subsequent writers is incorrect; by employing a revised test, they have shown that the pattern of US trade in 1947 was indeed in accordance with the Heckscher-Ohlin theorem (Williams, Leamer) which, in addition, turns out to be valid for many other countries (though not for all) in more recent times (Clifton and Marxsen). See also Leamer (1984) for an original study according to which "what emerges from the data analysis is a surprisingly good explanation of the main features of the trade data in terms of a relatively brief list of resource endowments" (p. 187). However, contrary to this result, Bowen, Leamer, and Sveinkauskas (1987), using the data on foreign trade of 27 countries in 1967, found that the Heckscher-Ohlin proposition was not confirmed.

4.6.2.2 Wrong Assumptions in the Model?

Let us now pass to the studies which belong to the *second group*, and begin with non-identical structures of demand. As we know (see Sect. 4.5.1) if the United States had tastes strongly biased in favour of the capital-intensive goods (the supposedly abundant factor), this might imply an import of these goods, whence the paradox. However, a study by Houthakker (1957) gives evidence for the contrary, namely for a similarity of the demand functions in different countries. Besides, it is a general phenomenon that, as per-capita income increases, society tends to spend more on labour-intensive goods such as services. It follows that, at the time considered by Leontief, the structure of US demand should have been biased in favour of labour-intensive goods relative to the rest of the world, that is, in exactly the opposite direction to that required for the paradox to occur.

Another important strand in the Leontief paradox problem is that consisting of those studies which aim to show that the phenomenon of factor-intensity reversals, far from being an exception, is the norm. The first systematic study in this sense is due to Minhas (1962) who, by using constant elasticity of substitution (CES) production functions, found that factor intensity reversals were quite frequent in reality. However, subsequent studies gave conflicting results (for example, Philpot, 1970, obtained results contrary to Minhas', whilst Yeung & Tsang, 1972, observed the presence of reversals), so that it is not possible to draw definite conclusions. It should however be noted that, as Fisher and Hillman (1984) have shown, the possible presence of factor intensity reversals at the level of single products or industries has no direct relevance for the aggregate ($2 \times 2 \times 2$) version of the Heckscher-Ohlin theorem.

In the traditional Heckscher-Ohlin theorem it is assumed that all countries produce (or can produce) the same goods. This is in disagreement with facts, as we shall see in Chap. 7; here we only wish to point out that, according to Brecher and Choudhri (1984), if one introduces new products in the Heckscher-Ohlin model, it is possible to give a satisfactory explanation of the Leontief paradox. We have so far examined some of the explanations of Leontief's paradox on the assumption that it exists. But this may not be so correct, since subsequent studies carried out with reference to both the United States and other countries have not systematically confirmed the presence of the paradox. As regards the United States, Stern and Maskus's (1981), already cited, confirmed the presence of Leontief's paradox by using the 1958 input-output table, whilst the paradox disappeared when the 1972 table was used. It should however be remembered that Stern and Maskus also take account of human capital (see above), so that their results are not directly comparable with Leontief's.

Wood (1994) argues that, contrary to the findings of most previous empirical tests, Heckscher-Ohlin theory provides an accurate explanation of the pattern of trade. The crucial point of his claim is that in testing this theory one should only consider internationally *immobile* factors, as this is the framework of the theory. Now, since capital is internationally mobile, all empirical tests that take capital into account and treat it as an immobile factor like land, do in fact mis-specify the theory. What Wood does is to examine the pattern of North-South trade in manufactures using a Heckscher-Ohlin model in which the factors of production are simply skilled and unskilled labour, which have a very low mobility between the North (the industrial countries) and the South (the developing countries). The empirical results are quite good, since he finds that the North (abundant in skilled labour) exports skill-intensive manufactures to the South (which is abundant in unskilled labour) in exchange for unskilled-labour-intensive manufactures. The importance of capital mobility in interpreting Leontief paradoxes is also stressed by Gaisford (1995). The role of capital mobility in the context of the Heckscher-Ohlin model is treated in Sect. 6.8.1.

The Heckscher-Ohlin theory assumes identical technology between countries. Yet, it is generally recognised that technology and factor supply differences can jointly determine comparative advantage. Harrigan (1997) proposes an empirical model aimed at jointly estimating the impact of different technologies and different factor endowments on international specialization and trade. He assumes Hicks-neutral technology differences across countries in addition to different factor endowments. The empirical estimation based on a data set of ten industrial countries

Table 4.2 Additional information on trade and endowments

Trade of goods and factor services	Values
Exports	\$16,678.4 million
Imports (competitive)	\$6,175.7 million
Net exports of capital services (K_T)	\$23,450 million
Net exports of labour services (L_T)	1.990 million man-years
Capital/labour intensity of trade (K_T/L_T)	\$11,783 per man-year

over 20 years for seven different manufacturing sectors show that technology differences are an important determinant of specialization, and that factor supplies alone cannot explain which industrial countries produce which goods.

As regards other countries, studies carried out in the years 1959–1962 by various authors (for a survey, see Bhagwati, 1964, pp. 24–25) with reference to Japan, India, East Germany, and Canada, in some cases confirmed Leontief's paradox and in others did not; similarly the article by Clifton and Marxsen (1984) already cited, shows that the pattern of trade in various countries (Australia, Ireland, Japan, Korea and New Zealand, besides the United States) conforms to the Heckscher-Ohlin theorem, whilst that of other countries (Israel, Kenya, and the United Kingdom) does not.

4.6.3 What Paradox? There Is No Paradox

In this section we reconsider the Leontief paradox in the light of the Heckscher-Ohlin-Vanek model. We follow the line of thought in Leamer (1980) since it is particularly instructive.

We have shown in Table 4.1 above the nature of the paradox: Leontief's (1953) study shows that the capital intensity of US exports is lower than the capital intensity of US imports. This result, at first sight, would imply that the US were a labour abundant country in 1947, and this is at odds with the sound opinion that the US were a capital abundant country. But this is not all. According to Leamer, Leontief reported additional findings as complementary information. These findings are summarized in Table 4.2, adapted from Leamer (1980, Table 2).

This table shows that the US had a trade balance surplus and was a net exporter of the services of both factors.

Leamer (1980, Table 3) supplements this information with results based on Travis (1964),[8] summarized here in Table 4.3. This table shows that the capital intensity in net exports was higher than in production and that the capital intensity in production was higher than in consumption.

[8]Net exports data are taken from Table 4.2. Production data are drawn from Travis (1964, Table 7 on p. 108). Consumption data are calculated using the identity Consumption = Production—Net Exports.

Table 4.3 Capital intensity of consumption, production, and trade

	Production	Net exports	Consumption
Capital ($ million)	328.519	23, 450	305, 069
Labour (million man-years)	47.273	1.99	45.28
Capital/labour ($ per man-year)	6, 949	11, 783	6, 737

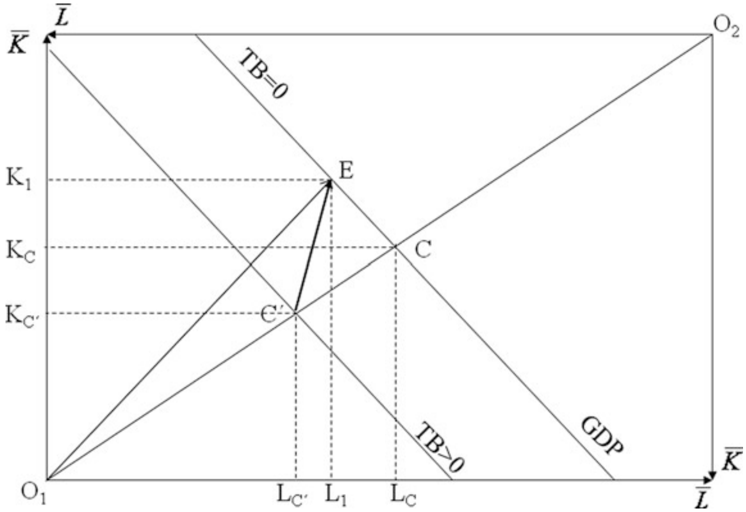


Fig. 4.16 The Leontief paradox reconsidered

Leamer made use of Tables 4.2 and 4.3. To understand Leamer's reasoning we have to bear in mind the Heckscher-Ohlin-Vanek model (see Sects. 4.4 and 4.5.3). Let us begin by identifying the factor content of trade vector in a situation of trade surplus.

A trade surplus takes place when production exceeds consumption. Thus, a country experiencing a trade surplus is consuming less than what it could, it is in fact saving part of its income. This implies that the vector of factor services embodied in consumption is smaller than the vector of factor services embodied in the maximum level of consumption that the country may achieve. In Fig. 4.16, the vector O_1C represents the latter and the vector O_1C' represents the former.

The two vectors have identical slope given the homotheticity of preferences. The line $TB = 0$ indicates the value of consumption corresponding to the situation of equilibrium of the trade balance. This line corresponds to the GDP line since all income is spent. The line $TB > 0$ indicates the value of consumption corresponding to the situation of trade balance surplus. With E being the endowment point, the vector $C'E$ is the factor content of trade vector for the country with the trade balance surplus (country 1).[9] Clearly, the country is a net exporter of the services of both factors. The slopes of the vectors $C'E$, O_1E, and O_1C represent, respectively, the capital intensity in net exports, in production, and in consumption. Remarkably, their ranking is precisely as found in the data reported in Table 4.3: namely, the slope of $C'E$, is larger than the slope of O_1E, which is larger than the slope of O_1C. It is surprising that for so long Leontief's findings did not stimulate investigation in the direction of comparing the factor content of consumption, trade, and production. The reason is that Leontief and many other scholars after him were not using the correct theoretical framework. They were thinking in terms of the two-by-two version of the Heckscher-Ohlin theory according to which we should find that the capital intensity in exports exceeds the capital intensity in imports for a capital abundant country. Yet, when there are more goods than factors the ordering of exports and imports by factor intensities is compatible with either ordering of relative factor abundance. In such a context, we have seen above (Fig. 4.14) a case where the capital abundant country imports a capital intensive good.

Therefore, the finding that the capital intensity in exports is smaller than in imports for a capital abundant country is not, per se, an invalidation of the Heckscher-Ohlin theory. We have seen above that a more robust prediction of the theory is formulated in terms of the Heckscher-Ohlin-Vanek theorem according to which, in balanced trade, each country exports the services of its abundant factor. Of course, if we observe the country being a net exporter of both factors it must be that it has a (sufficiently large) trade surplus. In this case we should observe the capital intensity in net trade to be larger than the capital intensity in consumption. Further, regardless of the trade balance, for a capital abundant country we should observe the capital intensity in production to exceed that in consumption, which is exactly the point made by Leamer. In conclusion, when using the correct theoretical framework, the paradox disappears.

4.6.4 Factor Content of Trade Studies

Seeing the Heckscher-Ohlin theory through the lenses of the factor content of trade has changed the way empirical research is conducted. We discuss in this section the logic of some of these verifications. The empirical verification consist in computing the factor content of trade from trade and technology data and comparing it with the factor content of trade resulting from the difference between factor content of production and the factor content of consumption. As discussed in Sect. 4.4, the two vectors should coincide; indeed, mathematically, from the former we obtain the latter.

It is convenient to discuss the matter by means of an example. Let the endowment of capital and labour in a country be, respectively, $K_1 = 10$ and $L_1 = 5$, and let

[9] Obviously, country 2 is running a trade balance deficit (its consumption exceeds production) and its factor content of trade vector is EC'.

the GDP of this country be, for instance, 8 % of the world GDP. Let $\bar{K} = 100$ and $\bar{L} = 80$ be world endowments. The country in question is therefore relatively well endowed with K. Let K_1^T and L_1^T denote the factor content of trade obtained from trade and technology data. K_1^T and L_1^T are obtained by multiplying the vector of net exports by the factor input per unit of output for each good. This should be equivalent to the factor content of trade obtained from endowments and consumption. Let $\hat{K}_1^T$ and $\hat{L}_1^T$ denote the latter. With reference to Fig. 4.11, we have $\hat{K}_1^T \equiv K_1 - K_{C1}$ and $\hat{L}_1^T \equiv L_1 - L_{C1}$. In our example they are $\hat{K}_1^T = 10 - 0.08 \times 100 = 2$, and $\hat{L}_1^T = 5 - 0.08 \times 80 = -1.4$. The two computations should give identical results, that is, we should find $\hat{K}_1^T = K_1^T$, and $\hat{L}_1^T = L_1^T$. This is in a nutshell the logic of the empirical studies based on the factor content of trade. The results of many such studies have shown that the two computations do not give identical results. In many cases even the signs do not match, that is K_1^T and $\hat{K}_1^T$ have opposite sign, likewise for L. Furthermore, K_1^T and L_1^T are often very small in absolute magnitude with respect to $\hat{K}_1^T$, and $\hat{L}_1^T$. This means that the observed volumes of trade are very small with respect to the volumes that we would expect to observe given factor endowment differences. This phenomenon is been dubbed by Trefler (1995) the "mystery of the missing trade". We discuss here three different ways to reconcile theory with data.

The first approach, suggested in Trefler (1993), consists in estimating the technological difference needed for the theory to fit the data perfectly and then verify the plausibility of these estimates against an alternative and independent indicator of technological differences. Thus, returning to our example, the first step is to multiply the endowment of each factor by parameters π_{Ki} and π_{Li} which reflect the productivity of that factor in country i. The resulting factor content of trade in our example becomes $\hat{K}_1^T = \pi_{K1} 10 - 0.08 * (\pi_{K1} K_1 + \pi_{K2} K_2)$, and $\hat{L}_1^T = \pi_{L1} 5 - 0.08 * (\pi_{L1} L_1 + \pi_{L2} L_2)$, and analogously for country 2. The equations $K_1^T = \hat{K}_1^T$, and $L_1^T = \hat{L}_1^T$ allow estimating the productivity parameters of each factor in country 1 relative to the same factor in country 2. Having done this estimation, the second step is to compare these estimates with alternative measures of productivity differences. Trefler does this by comparing the estimates of labour productivity differences with observed real wage differences between countries. One expects to see that higher wages correspond to higher estimates of productivity (in our example, if one finds $\pi_{L1} > \pi_{L2}$ then one should also observe the wage in country 1 to be higher than in country 2). Indeed the correlation between estimated productivity and wages found by Trefler is extremely high (he reports an estimated coefficient of 0.9). This result lends support to the technology-amended version of the Heckscher-Ohlin theory but is only indirect evidence. Evidence is indirect since the factor content of trade equation holds as an identity given the degrees of freedom generated by the insertion of the productivity parameters. A more direct evidence can be found in the second approach, proposed in Trefler (1995).

The second approach consists, broadly speaking, in restricting the technology differences to be uniform in the sense that all factors are assumed to be proportionally more productive in a country with respect to the other country.[10] Then the factor content of trade equation is no longer an identity. Using the second approach Trefler finds that nearly one half of the "missing trade" is explained by uniform productivity differences between countries. This result too, and even more directly than the first approach, gives support to a technology amended version of the H-O theory.

A third approach is suggested by Davis and Weinstein (2001). Their starting point is that in the presence of barriers to international trade there is no complete convergence of commodity prices; neither in absolute nor in relative terms. Therefore, as is clear from the study of Sect. 4.3.1, there is no complete equalization of factor prices either. In particular, in each country the relative price of the relatively scarce factor will be higher than what it would be if commodity prices had converged completely (see Fig. 4.8 in Sect. 4.3.1). As a consequence, the factor intensity for each industry will differ between countries: the K-abundant country will use more K-intensive techniques than the L-abundant country in all industries. This becomes important when computing K_1^T and L_1^T. In computing the factor content of exports and imports one should apply the techniques prevailing in each country. Davis and Weinstein find that the factor content of trade equation fits the data well when account is taken of the different techniques between countries. It is worth mentioning that this result does not require assuming different technologies between countries and, in this sense, it represents an even stronger evidence in support of the Heckscher-Ohlin-Vanek model.

4.6.5 Concluding Remarks

The first confrontation of the Heckscher-Ohlin theory with data has been rather traumatic since Leontief's results, at first sight, appeared as paradoxical. A number of explanations have been proposed for the paradox but the empirical performance of the theory remained far from satisfactory. Since the 1980s however the theory has fared pretty well. First, the correct interpretation of Leontief's result showed that there was no paradox after all. Second, factor content of trade studies have provided solid empirical support for the theory. The fact that the empirical performance of the Heckscher-Ohlin model improves when taking account of technology and demand differences between countries is particularly in line with what the neoclassical theory (see Sect. 1.2 and Chap. 3) had conjectured already one and a half century ago. Overall, this ancient theory proves to be very relevant to explain contemporary patterns of trade.

[10]In our example, for instance, if L is δ % more productive in country 1 than in country 2, so is K in exactly the same proportion δ. These are called Hicks neutral technology differences. For a more detailed definition see Sect. 13.5.1.

References

Baldwin, Robert E. (2008). *The development and testing of Heckscher-Ohlin trade models.*

Bhagwati, J. N. (1964). The pure theory of international trade: A survey.

Bhagwati, J. N. (1972). The Heckscher-Ohlin theorem in a multi-commodity case.

Bhagwati, J. N., & Srinivasan, T. N. (1983). *Lectures on international trade.*

Bowen, H. P., Learner, E. E., & Sveinkauskas, L. (1987). Multifactor tests of the factor abundance theory.

Brecher, R. A., & Choudhri, E. U. (1982). The factor content of international trade without factor-price equalization.

Brecher, R. A., & Choudhri, E. U. (1984). New products and the factor content of international trade.

Casas, F. R., & Choi, E. K. (1984). Trade imbalance and the Leontief paradox.

Chacholiades, M. (1978). *International trade theory and policy.*

Charos, E. N., & Simos, E. O. (1988). The effects of human capital and R&D upon international trade flows: Evidence from a multi-input, multi-output model for the United States.

Chipman, J. S. (1966). A survey of the theory of international trade: Part 3, the modern theory.

Clifton, D. S., Jr., & Marxsen, W. S. (1984). An empirical investigation of the Heckscher-Ohlin theorem.

Davis, D. R., & Weinstein, D. E. (2001). An account of global factor trade.

Davis, D. R., Weinstein, D. E., Bradford, S. C., & Shimpo, K. (1997). Using international and Japanese data to determine when the factor abundance theory of trade works.

Deardoff, A. V. (1982). The general validity of the Heckscher-Ohlin theorem.

Deardoff, A. V. (1984). Testing trade theories and predicting trade flows.

Diab, M. A. (1956). *The United States capital position in the structure of its foreign trade.*

Dixit, A. K., & Woodland, A. (1982). The relationship between factor endowments and commodity trade.

Ethier, W. J. (1982). The general role of factor intensity in the theorems of international trade.

Ethier, W. J. (1984). Higher dimensional trade theory.

Feenstra, R. C. (2004). *Advanced international trade: Theory and evidence* (chaps. 1–3).

Fisher, F. M., & Hillman, A. L. (1984). The commodity pattern of trade and the Heckscher-Ohlin theorem in the presence of aggregate and commodity-specific factor-intensity reversals.

Gaisford, J. D. (1995). International factor mobility, the factor content of trade and Leontief paradoxes.

Hamilton, C., & Svensson, L. E. O. (1984). Do countries' factor endowments correspond to the factor contents in their bilateral trade flows?

Harkness, J. P. (1978). Factor abundance and comparative advantage.

Harkness, J. P. (1983). The factor-proportions model with many nations, goods and factors: Theory and evidence.

Harrigan, J. (1997). Technology, factor supplies and international specialization: Estimating the neoclassical model.

Harrod, R. F. (1953). Factor-price relations under free trade.

Heckscher, E. F. (1919, 1949). The effect of foreign trade on the distribution of income. (English translation of the original 1919 article in 1919, Ekonomisk Tidskrift).

Helpman, E. (1984a). The factor content of international trade.

Herberg, H., Kemp, M. C., & Tawada, M. (1982). Further implications of variable returns to scale.

Horiba, Y. (1974). General equilibrium and the Heckscher-Ohlin theory of trade: The multi-country case.

Houthakker, H. S. (1957). An international comparison of household expenditure patterns, commemorating the centenary of Engel's law.

Johnson, H. G. (1957). Factor endowments, international trade and factor prices.

Kohler, W. K. (1988). Modeling Heckscher-Ohlin comparative advantage in regression equations.

Kreinin, M. E. (1965). Comparative labor effectiveness and the Leontief scarce factor paradox.

Lancaster, K. (1957). The Heckscher-Ohlin trade model: A geometric treatment.
Lane, J. (1985). An empirical estimate of the effects of labor-market distortion on the factor content of U.S. trade.
Leamer, E. E. (1980). The Leontief paradox, reconsidered.
Leamer, E. E. (1984). *Sources of comparative advantage: Theory and evidence*.
Leamer, E. E., & Levinsohn, J. (1995). International trade theory: The evidence.
Leontief, W. W. (1953). Domestic production and foreign trade: The American position re-examined.
Leontief, W. W. (1956). Factor proportions and the structure of American trade: Further theoretical and empirical analysis.
Lerner, A. P. (1933, 1952). Factor prices and international trade. (This article reproduces the text of a paper prepared for a seminar at the London School of Economics in 1933, as it was originally written).
Maskus, K. E. (1985). A test of the Heckscher-Ohlin-Vanek theorem: The Leontief common-place.
Minhas, B. S. (1962). The homohypallagic production function, factor intensity reversals, and the Heckscher-Ohlin theorem.
Mundell, R. A. (1957). International trade and factor mobility.
Neary, J. P. (1984). The Heckscher-Ohlin model as an aggregate.
Neary, J. P. (1985b). Two-by-two international trade theory with many goods and factors.
Ohlin, B. (1933). *Interregional and international trade*.
Pearce, I. F. (1952). The factor price equalization myth.
Philpot, G. (1970). Labor quality, returns to scale and the elasticity of factor substitution.
Samuelson, P. A. (1948). International trade and the equalisation of factor prices.
Samuelson, P. A. (1949). International factor price equalisation once again.
Samuelson, P. A. (1953). Prices of factors and goods in general equilibrium.
Samuelson, P. A. (1967). Summary on factor price equalization.
Sarkar, A. (1984). A dynamic assignment problem.
Stern, R. M. (1975). Testing trade theories.
Stern, R. M., & Maskus, K. E. (1981). Determinants of the structure of U.S. foreign trade 1958–1976.
Svensson, L. E. O. (1984). Factor trade and goods trade.
Takayama, A. (1972). *International trade: An approach to the theory* (Pts. I and IV).
Travis, W. P. (1964). *The theory of trade and protection.*
Trefler, D. (1993). International factor price differences: Leontief was right!
Trefler, D. (1995). The case of the missing trade and other mysteries.
Vanek, J. (1959). The natural resource content of foreign trade, 1870–1955, and the relative abundance of natural resources in the United States.
Vanek, J. (1968). The factor proportions theory: The n-factor case.
Williams, J. R. (1970). The resource content in international trade.
Wood, A. (1994). Give Heckscher and Ohlin a chance!
Yeung, P., & Tsang, H. (1972). Generalized production function and the factor-intensity crossovers: An empirical analysis.

Chapter 5
The Four Core Theorems

5.1 Introductory Remarks

As clarified in Sect. 1.2, Ricardian comparative-cost theory, neoclassical theory, and Heckscher-Ohlin theory together form the body of the traditional theory of international trade. However, the factor-proportion theory is often identified with "the" traditional theory, and the Heckscher-Ohlin theorem, together with the factor-price-equalization (FPE) theorem and two additional theorems (the Stolper- Samuelson theorem and the Rybczynski theorem), are said to constitute the four core theorems of the traditional theory of international trade.

Be it as it may, the purpose of the present chapter is to complete the treatment of the previous chapter by examining the Stolper-Samuelson and Rybczynski theorems. Both of them are comparative statics theorems, as they examine the effects of a change in some data on the general equilibrium of the economy. It is important to note that they are general theorems, in the sense that they also hold for a closed economy; but we shall be concerned with their ultimate impact on open economies.

5.2 The Heckscher-Ohlin and Factor-Price-Equalization Theorems

These have been extensively examined in the previous chapter. We only add that the factor-price-equalization theorem, though usually presented as a corollary of the Heckscher-Ohlin theorem (and we have followed standard practice in the previous chapter), is valid independently of the latter. In fact, what is really essential for FPE is the absence of complete specialization, given internationally identical technology of the constant-returns-to-scale type. When this is true, it does not matter whether international trade is due to different relative factor endowments and/or to different demand conditions (as could be the case under the neoclassical theory). See Samuelson (1948, 1949, 1967).

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_5,

5.3 The Stolper-Samuelson Theorem

The *Stolper-Samuelson theorem* (Stolper and Samuelson, 1941) states that the increase in the relative price of a commodity favours (in the sense that it raises the unit real reward of) the factor used intensively in the production of the commodity. This can be simply shown by using the Heckscher-Ohlin theory treated in Chap. 4.

Without loss of generality we can assume that commodity A is capital intensive while the labour-intensive commodity is B. Suppose now that the domestic relative price p_B/p_A increases: given the one-to-one relation between the relative price of goods and the relative factor price, it follows that p_L/p_K increases (this is due to the fact that in our case this correspondence is monotonically increasing: see Fig. 4.5a). This shows that the relative price of labour increases, but the theorem asserts something more, i.e. that the "real price" of labour (p_L/p_A, if we use commodity A as numéraire) increases, and to prove this more passages are required.

The increase in p_L/p_K causes the capital/labour ratio to increase in both sectors (see Fig. 4.4). Since the production functions are homogeneous of the first degree, the marginal productivities are functions solely of the factor *ratio* (see Sect. 19.1.3) and, more precisely, *MPK* is a decreasing function, and *MPL* an increasing function, of K/L. Now, as we have just shown, K/L has increased in both sectors; it follows that the marginal productivity of labour (and so its unit real reward, which in perfect competition coincides with *MPL*) increases. This completes the proof of the theorem.

In this proof of the Stolper-Samuelson theorem we have used the Heckscher-Ohlin theorem and, in particular, we have implicitly assumed no factor-intensity reversals (this is required for the one-to-one correspondence between relative price of goods and the relative price of factors), but it is important to note that the former theorem *does not depend* on the latter in any essential way. It is in fact possible to prove the Stolper-Samuelson theorem in its general formulation independently of the Heckscher-Ohlin theorem, which dispenses us to examine what happens when there is factor-intensity reversal.

Let us then assume that the domestic relative price of commodity B increases. We also assume that, in the interval under consideration, commodity B is unambiguously labour-intensive (which does not exclude the presence of factor-intensity reversals elsewhere). The increase in p_B/p_A causes a shift on the transformation curve towards a point where more B and less A is produced (see, for example, Fig. 3.5), so that resources will have to be reallocated from the latter to the former industry. But, since B is more labour intensive than A, it follows that—*at given relative factor prices*—the proportion in which capital and labour become available as a result of the decrease in the production of A does *not* coincide with the proportion in which the expanding sector B is prepared to absorb them.

In fact, at the given factor price ratio, labour and capital are made available by sector A in a *lower* proportion than that required by sector B. There follows, at the global level, an excess demand for labour and/or an excess supply of capital, with the consequence that p_L/p_K increases. As this ratio increases, cost-minimizing firms will substitute capital for labour in both sectors, that is, they will choose techniques with a higher K/L ratio. Since the marginal productivity of labour is an increasing function of this ratio, the theorem is proved.

An implication of the Stolper-Samuelson theorem is the so-called *magnification effect* (Jones, 1965). This effect states that the increase in the *nominal* price of the benefited factor is proportionally *greater* than the increase in the commodity price. In fact, since under perfect competition we have $p_B MPL_B = p_L$ or $MPL_B = p_L/p_B$, it is obvious that the increase in MPL_B following an increase in p_B must be accompanied by an increase in p_L proportionally greater than the increase in p_B.

The relevance of the Stolper-Samuelson theorem for international economics lies in its use for the examination of the redistributive effects of tariffs. A tariff, in fact, normally causes an increase (with respect to the international price ratio) in the domestic relative price of the good on which the tariff is levied, and hence income redistribution effects due to the change in real factor rewards. This will be dealt with in depth in Sect. 10.5.1.1.

5.4 The Rybczynski Theorem

The point of departure for examining the effects of an increase in factor endowments is Rybczynski's theorem (Rybczynski, 1955) according to which the increase in the quantity of a factor (given the other) will cause an increase in the output of the commodity which is intensive in that factor and a decrease in the output of the other commodity, at unchanged commodity and factor prices.

The proof of this theorem can be given by using the box diagram (see Sect. 3.1). In Fig. 5.1 the initial box is O_AHO_BG and, given the commodity prices, we shall take E as the initial equilibrium point. The labour-intensive commodity is A, for the labour/capital ratio in its production, O_AD/ED, is higher than the labour/capital ratio in the production of B, O_BF/EF; given the form of the locus of efficient points O_AO_B, this property holds at all points of this locus. Let us now assume that the quantity of labour increases from O_AG to O_AG'. The new equilibrium point will be E', as this is the only point lying along the ray O_AE such that the straight-line segment drawn from this point to the new origin O_B' is parallel to O_BE. That the new equilibrium point must be characterized by this property can be shown as follows.

Since commodity and factor prices are, by assumption, unchanged, the marginal rate of technical substitution (equal, in equilibrium, to the factor-price ratio) must also be unchanged, that is, the common slope of the A- and B- isoquants at the new equilibrium point must be equal to that at the previous equilibrium point. Now, given the property of radiality of homogeneous production functions (see Sect. 19.1), the A isoquant through E' has the same slope as the A isoquant through E; similarly, as $O_B'E'$ is parallel to O_BE, the isoquants of B have the same slope along ray $O_B'E'$ as they had along ray O_BE, and, therefore, the B isoquant through E' has the same slope as the B isoquant through E. It follows that the isoquants through E' have

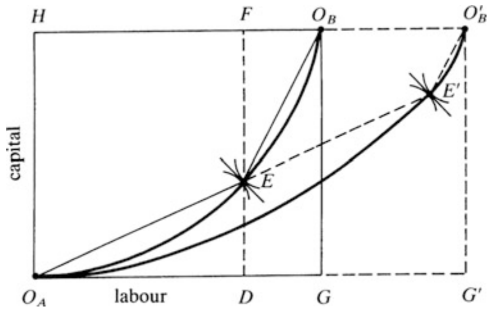


Fig. 5.1 Rybczynski's theorem

the same slope as the isoquants through E, and this shows both that E' lies on the new efficiency curve (as it fulfils the conditions of efficiency) and that E' is the new equilibrium point.

Now, since the more distant an isoquant is from its origin the greater is the production level it represents, and since $O_AE' > O_AE$, $O'_BE' < O_BE$, it follows that the output of A (the labour-intensive commodity) has increased as a consequence of the increase in the quantity of labour, whilst the output of the capital-intensive commodity B has decreased. This completes the demonstration.

An implication of the Rybczynski theorem is the so-called *magnification effect* (Jones, 1965), according to which the output of the expanding sector increases more than proportionally to the increase in the factor. This can easily be checked in Fig. 5.1. With constant returns to scale, the isoquant index is proportional to the distance from the origin, hence we can measure the increase in the output of A by EE'/O_AE, which is clearly greater than the proportional increase in labour (given by GG'/O_AG).

5.4.1 An Alternative Diagram

An alternative representation of this theorem can be given by a diagram which uses the transformation curve. At the same time we shall also explain an important corollary of Rybczynski's analysis, namely that the increase in the quantity of a factor (at unchanged quantity of the other factor) will cause a decrease in the relative price of the commodity that is intensive in that factor.In Fig. 5.2, TT is the initial transformation curve which shifts to $T'T'$ as a consequence of the increase in the quantity of labour, Q and Q' are the two equilibrium points (production points in the case of an open economy) at the same commodity price ratio ($R'R'$ and RR are parallel). Since A is the labour-intensive commodity, its output will increase and the output of B will decrease, that is, Q' must be situated to the left of Q'' (which is the point at which the output of B is the same as that at Q). However, point Q' is only hypothetical. Since the $R'R'$ line is higher than the RR line, and since each of these can be interpreted as an isoincome line, $R'R'$ represents a higher national

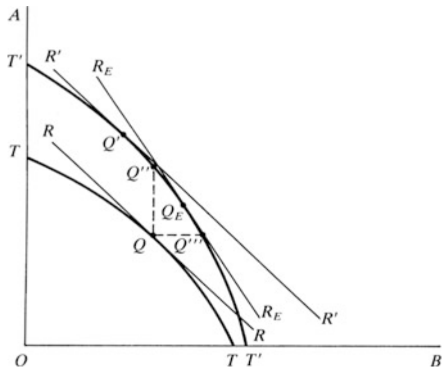


Fig. 5.2 Rybczynski's theorem and relative price of goods

income at constant prices (that is, at the same prices existing at the initial equilibrium point Q) than that represented by *RR*. Now—if we exclude inferior goods—this increase in income will cause an increase in the demand for *both* commodities; since, as we have seen, the output of B is lower, there will be an excess demand for this commodity which will cause an increase in its relative price (p_B/p_A) and, consequently, in its output. Therefore the new equilibrium point will be found in the stretch $Q''Q'''$ of the curve $T'T'$: only there, in fact, is the output of both A and B higher than at E. It can also be seen from the figure that at any point included in this stretch, for example Q_E, the relative price of A is lower, as this price is measured by the (absolute value of the) slope of the $R_E R_E$ line with respect to the A axis, which is smaller than the analogous slope of the *RR* line.

This holds in a closed economy. But what about an open economy? To answer this question we must distinguish between a small and a large economy, and take the structure of trade into consideration. In all cases we keep the assumption that no good is inferior.

1. Suppose that A is the import good and B the export good. The domestic demand increases for both commodities. Since at Q' the output of B is lower while its domestic demand is higher, the domestic excess supply (i.e., the supply of exports) decreases. Thus in the international market there will be a decrease in the supply of B. We now must distinguish whether the country under consideration is small or large. In the former case the decrease of the supply of B will have a negligible effect on the international market, so that the international price ratio (terms of trade) p_B/p_A will not change, and consequently the domestic price ratio will not change (in the model, it is equal to the terms of trade). Thus the production point remains at Q'.

   On the contrary, if the country is large, the decrease in its supply of exports will cause an excess demand for B in the international market, hence an increase in p_B/p_A, and the production point will move somewhere between Q'' and Q'''.
2. Consider now the case in which the import good is B. Since at Q' the output of B is lower while its domestic demand is higher, the domestic excess demand increases. Thus in the international market there will be an increase in the

demand for B. In the case of a small country, this increase will have a negligible effect, so that the international price ratio (terms of trade) p_B/p_A will not change, and the production point remains at Q'.

On the contrary, if the country is large, the increase in its demand for B will cause an excess demand for B in the international market, hence an increase in p_B/p_A, and the production point will move somewhere between Q'' and Q'''.

In conclusion, if we exclude the case of a small country, the result in an open economy is the same as in a closed economy.

The relevance of Rybczynski's theorem in international trade theory lies in its use to examine the effects of international factor mobility (see Sect. 6.8.1), and to examine the effects of growth, when the cause of growth is an increase in factor endowments (see Sect. 13.4).

References

Jones, R. W. (1965). The structure of simple general equilibrium models.
Rybczynski, T. M. (1955). Factor endowments and relative commodity prices.
Samuelson, P. A. (1948). International trade and the equalization of factor prices.
Samuelson, P. A. (1949). International factor price equalization once again.
Samuelson, P. A. (1967). Summary on factor price equalization.
Stolper, W. F., & Samuelson, P. A. (1941). Protection and real wages.

Chapter 6
Some Refinements

6.1 Introduction

Explicitly or implicitly, all theoretical models so far examined have a common set of assumptions: perfect competition, rigid supply of ubiquitous and internationally immobile productive factors, absence of intermediate goods, absence of transport costs, certainty, absence of illegal trade (such as smuggling), and so on.

These are undoubtedly assumptions which do not correspond to reality, so that it is legitimate to ask what happens when they are relaxed. In this chapter we shall be concerned with the introduction of those elements which can be dealt with from inside the traditional theory, of which they are in fact a refinement (the examination of the case of non-constant returns to scale is also a refinement, for which see Sect. 3.5). In Part III we shall examine the consequences of introducing non-competitive elements and other alternative explanations of international trade, which can be fitted only partially (if at all) into the framework of the traditional theory.

Although the various topics treated in this chapter may seem unrelated to one another, there is a common thread running through them, which is to show how far one can go while remaining in the context of the traditional account of trade in a competitive setting with constant returns to scale. This adaptability may be one of the reasons why the traditional theory is still alive and well after the advent of the new explanations of international trade (see Part III).

6.2 The Specific Factors Model

Factors of production have been so far assumed to be ubiquitous in all sectors. It is however possible that, alongside with these all-purpose factors, other factors exist which are *specific* to each sector. This means that they can only be used in the sector of pertinence and not elsewhere. For example, the (physical) capital

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_6,

required to produce computer microprocessors is quite different from that used to produce textiles, and they are not interchangeable in the short run. Long-run interchangeability is of course possible, as the (Marshallian) long run is, in fact, defined as a period of time sufficient to allow all factors to be in free intersectoral mobility. In the long run, capital can move from the textile to the microprocessor sector via depreciation without replacement in the former and new investment in the latter.

Thus the models so far examined can be considered as long-run models, while the specific factors model is more appropriate for the short run.

Although it maintains the basic two-sector setting, the specific factors model[1] is actually *a three-factor* model. In fact, besides the ubiquitous homogeneous factor (say, labour), two additional and different factors are needed to represent specificity. These may be, for example, capital and land, if we wish to consider manufacturing and agriculture as our two sectors. We remain in the traditional framework and assume that the specific factors are two different capital goods (say, K^A and K^B). Thus commodity A is produced using labour and K^A, while commodity B is produced using labour and K^B.

Apart from this, the model's setting is identical with the traditional one: perfect competition, production functions homogeneous of the first degree, etc.

As we have already seen in previous chapters, perfect competition implies the equilibrium condition *value of the marginal product of a factor = price of the factor*. Labour mobility implies that the wage rate is equalized between sectors. Hence we can write

$$\begin{aligned} p_A MPL_A &= p_L, \\ p_B MPL_B &= p_L, \end{aligned} \tag{6.1}$$

where MPL_A, MPL_B are the (physical) marginal products of labour in the two sectors, and p_L is the nominal wage rate. Letting $w = p_L/p_A$ denote the real wage rate in terms of commodity A, and $p = p_B/p_A$ the commodity price ratio, we have

$$\begin{aligned} MPL_A &= w, \\ pMPL_B &= w, \end{aligned} \tag{6.2}$$

hence

$$MPL_A = pMPL_B, \tag{6.3}$$

[1]This model was widely used prior to the predominance of Heckscher-Ohlin theory (see, for example, Haberler, 1936), which pushed it into the background: see Bhagwati et al (1998, Chap. 7). It was simultaneously and independently revived by Samuelson (1971), who called it the Ricardo-Viner model, and Jones (1971)

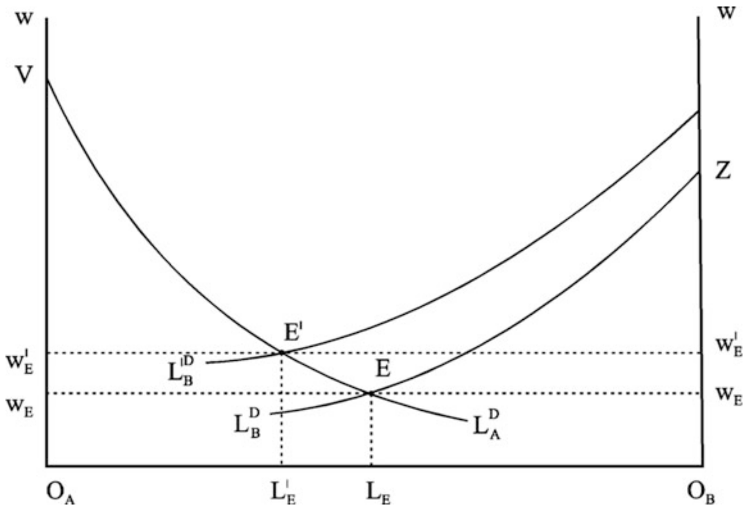


Fig. 6.1 The specific factors model

which determines the optimal allocation of labour between the two sectors and hence—since the two stocks of specific capital are also fully employed—the outputs of the two commodities for any given p.

Equation (6.3) can be given a simple graphic representation. In Fig. 6.1, the total amount of labour is measured by the segment $O_A O_B$. The quantity of labour used in sector A is measured from the origin O_A, while that used in sector B is measured from O_B. In the ordinate we show the real wage rate w. Curves L_A^D, L_B^D represent the demand-for-labour schedules in the two sectors, derived from Eqs. (6.2) for a given p. The equilibrium condition (6.3) obtains at point E. This determines the equilibrium real wage rate w_E and the optimal allocation of labour, which consists of $O_A L_E$ employed in sector A and $O_B L_E$ employed in sector B.

Since the area below a marginal product curve is total product, in sector A total labour income is the area $O_A w_E E L_E$, while total income of the specific capital K_A is the residual area $w_E V E$. Similarly in sector B labour receives $O_B w_E E L_E$ and the residual $w_E Z E$ goes to K^B.

Let us now determine the general equilibrium situation of the economy, which can be done through the general-equilibrium supply and demand curves (see Sect. 3.2). Let us note that the transformation curve cannot be derived from the two-dimensional box diagram as shown there, because the presence of three factors would require a three-dimensional diagram. It is however easy to show that the general-equilibrium supply curve is an increasing function of the appropriate relative price. Consider, for example, an increase in p. Given the second equation in (6.2), the L_B^D curve shifts upwards to position $L_B'^D$. This means that the amount of labour employed in sector B increases (from $O_B L_E$ to $O_B L'_E$), and hence the output of commodity B increases while that of commodity A decreases.

Let us assume that the general-equilibrium commodity price ratio is that corresponding to curve L_B^D, say p_1, and consider the introduction of international trade in a two-country framework. The condition for international trade to take place is that p_2, the closed-economy commodity price ratio in country 2, is different from p_1. Without loss of generality we can assume that $p_2 > p_1$, hence the post-trade price ratio p^* will be somewhere in between. Thus we can take E' as the post-trade equilibrium in country 1. In country 2 there will be a downward shift of the demand for labour in sector B, since $p^* < p_2$. This shows that there will be an increase in sector B's output in country 1 and in sector A's output in country 2.

What about the influence of trade on factor prices? A central result in traditional trade theory is factor price equalization (FPE, see Sect. 4.3). This is no longer true in the present context. Due to specific factors, marginal productivities are no longer equalized across countries. It remains true that, with constant returns to scale, marginal productivities only depend on the factor input ratio, but this ratio need no longer be equal across countries even with internationally identical production functions.

Take, for example, MPL. In the traditional $2 \times 2 \times 2$ model, MPL_{1A} depends on L_{1A}/K_{1A}, while MPL_{2A} depends on L_{2A}/K_{2A}. Since L_{1A}/K_{1A} and L_{2A}/K_{2A} turn out to be equal for the reasons explained in Chap. 4, it follows that $MPL_{1A} = MPL_{2A}$, etc.

In the present model, MPL_{1A} depends on $L_{1A}/\overline{K}_1^A$, while MPL_{2A} depends on $L_{2A}/\overline{K}_2^A$, where $\overline{K}_1^A, \overline{K}_2^A$ are the total amounts of the specific factor K^A existing in the two countries. There is no reason why these two ratios should be equalized.

That FPE does not hold should come as no surprise if we recall that even in the context of the traditional theory a model with more factors than commodities does not yield FPE (see Sect. 20.4).

The other basic results of the traditional $2 \times 2 \times 2$ model are the Rybczynski and Stolper-Samuelson theorems.

Let us begin with the Rybczynski theorem. We first note that it makes little sense to talk of factor intensity in the presence of specific and hence not comparable capital stocks. It is however possible to reformulate the theorem in the sense that an increase in a specific factor causes an increase in the output of the commodity in which it is employed and a decrease in the output of the other commodity. This can easily be shown in terms of Fig. 6.1. Take for example an increase in K^B. With constant returns to scale and decreasing marginal productivities, an increase in a factor must have a positive effect on the marginal productivity of the other factor (see Sect. 19.1.3). This means that for a given p the L_B^D curve shifts upwards, for example to position $L_B'^D$.

The new equilibrium point is E', where less labour is allocated to sector A (hence a lower output of A) and more to sector B (whose output increases both because more labour is employed there and because of the increase in its specific capital).

The outcome is however different when the ubiquitous factor is considered. An increase in labour (see Fig. 6.2) shifts the origin O_B to O_B'. Thedemand-for-labour

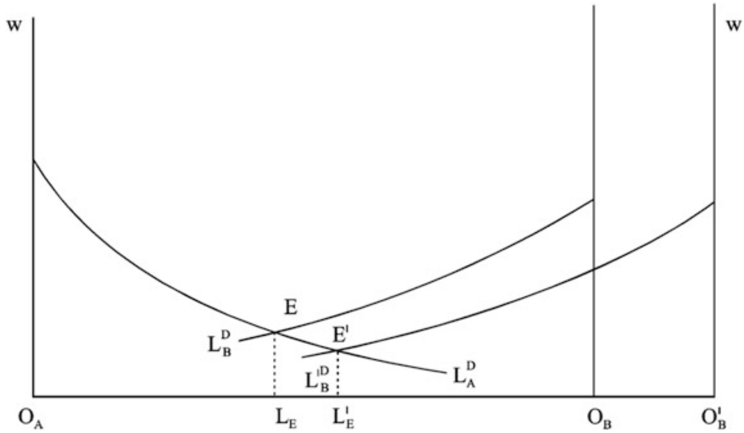


Fig. 6.2 The specific factors model and Rybczynski's theorem

schedule in sector B is now $L_B^{\prime D}$, which is the same as the curve L_B^D but referred to the new origin. The equilibrium point shifts from E to E', where more labour is employed in *both* sectors ($O'_B L'_E > O_B L_E$, and $O_A L'_E > O_A L_E$). Hence an increase in the ubiquitous factor brings about an increase in the output of *both* commodities.

Let us finally consider the Stolper-Samuelson theorem in its general formulation (see Sect. 5.3) according to which the increase in the relative price of a commodity raises the unit real reward of the factor used intensively in the production of that commodity. Again noting that it makes little sense to talk of factor intensity in the presence of specific and hence not comparable capital stocks, the Stolper-Samuelson theorem can also be reformulated in terms of specific factors. Let us then consider the reward of the specific factor used in the sector producing the commodity whose relative price increases.

For this purpose we can use Fig. 6.1, where we see that an increase in p (the relative price of commodity B) causes more labour to be used in sector B and less in sector A. The (specific) capital to labour ratio decreases in sector B and increases in sector A. Since the marginal productivity of capital (which is the real unit reward of capital) depends negatively on the capital to labour ratio, it follows that the marginal productivity of capital increases in sector B and decreases in sector A.

The effect on the ubiquitous factor is however ambiguous. The wage rate does, in fact, increase in terms of commodity A (from w_E to w'_E), but declines in terms of commodity B (since the marginal productivity of capital is higher there, the marginal productivity of labour is lower). Whether wage earners are better or worse off depends on the composition of their expenditure, a result that has been dubbed the *neoclassical ambiguity* in trade theory.

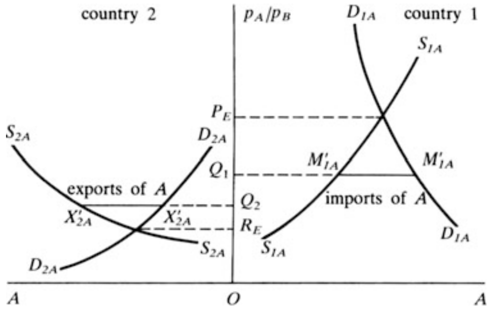


Fig. 6.3 The cost of transport: diagram 1

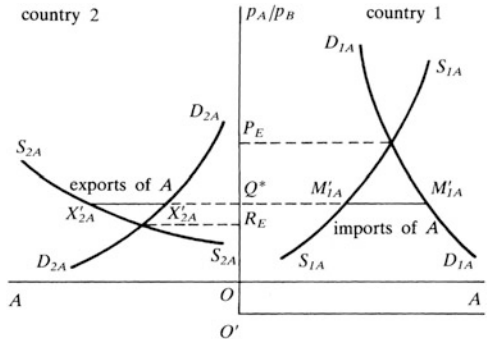


Fig. 6.4 The cost of transport: diagram 2

6.3 Transport Costs and International Trade

If we assume that the total cost of transport increases in proportion to the quantity of goods transported, i.e., that the cost of transport per unit of the commodity transported is constant, we can deal with the problem simply by taking up Fig. 3.6 again. The presence of constant unit cost of transport means that the price of a good in the importing country will be higher than the price of the same good in the exporting country by an amount equal to the given unit cost of transporting the commodity.

In Fig. 6.3 we have traced the same curves already analysed in Fig. 3.6. Note however that on the vertical axis we now measure p_A/p_B instead of p_B/p_A, so that the form and position of the curves has changed. Equilibrium is established when the relative price of A is OQ_2 in country 2 (the exporting country) and OQ_1 in country 1 (the importing country). The difference between OQ_2 and OQ_1, equal to segment Q_1Q_2, represents the given unit cost of transport, and segment $X'_{2A}X'_{2A}$ has the same length as segment $M'_{1A}M'_{1A}$.

An alternative way of showing the same phenomenon is described in Fig. 6.4 (the Cunynghame-Barone diagram: see Cunynghame, 1904, and Barone, 1908). This is derived from Fig. 6.3 simply by lowering the axis where the quantity of the importing market is measured (or, what amounts to the same thing, by raising the axis of the quantity of the exporting market) by an amount OO', corresponding to the given unit transport cost ($OO' = Q_1Q_2$ in Fig. 6.3). The equilibrium prices in country 1 and country 2 can be read in this diagram as $O'Q^*$ and OQ^* respectively, because at these prices the excess demand for A in country 1 ($M'_{1A}M'_{1A}$) is exactly equal to the excess supply of A in country 2 ($X'_{2A}X'_{2A}$). It goes without saying that $O'Q^* = OQ_1$, and $OQ^* = OQ_2$, where OQ_1 and OQ_2 are the values referred to in Fig. 6.3.

In the treatment so far, we have limited ourselves to considering the cost of transport relative to good A. As one may well assume that there are also transport costs for good B, the diagrams used above are no longer valid, because in calculating the relative price p_A/p_B (or p_B/p_A according to the case) it is necessary to take into account the unit cost of transport both for A and B. The analysis of the general case can be more easily conducted in terms of offer curves if we simplify by assuming that the cost of transport is expressed in terms of the good transported, of which this cost constitutes a given proportion (the assumption made above, of transport costs proportional to the quantity of good transported is also maintained), let it be c_A for good A and c_B for good B. This means that only a proportion of the good exported is received as an import by the importing country, the difference being in fact consumed by transport. This method of calculating transport costs was introduced by von Thünen (1826, chap. 4) and Samuelson (1954). Von Thünen assumed that the cost of transporting grain largely consists of the grain consumed during the transportation by the horses pulling the carriage. Samuelson assumed that only a fraction of exports reaches the country of destination as imports, just as only a fraction of ice exported reaches its destination as unmelted ice. The Samuelson ice similitude was subsequently called in the literature the *iceberg* assumption.

Now, if we use k_A to indicate the proportion of A received by the importing country, then obviously the relationship $c_A + k_A = 1$ must be valid; similarly, $c_B + k_B = 1$. We now see how the offer curves are modified as a consequence of introducing transport costs in the manner described above. We must remember (Sect. 3.4) that OG_1 is the offer curve of country 1, which imports good A and exports good B, while OG_2 is the offer curve of country 2, which imports B and exports A. In order to examine the shifts in these curves, we must first establish whether we want to work with c.i.f. or f.o.b. curves. If we consider the cost of transport relative to good B, we can modify the offer curve of country 1 to indicate that this country offers a smaller amount of good B considered as c.i.f. (cost, insurance, and freight, that is, delivered at destination in country 2) in correspondence to any given amount of A it demands, because part of the original quantity of B is consumed by transport. Or else we can modify the offer curve of country 2 to denote that it demands a greater amount of good B considered as f.o.b. (free on board that is, excluding the cost of transport) in correspondence to any given amount of A supplied, because a part of B is consumed by transport. The same can be said for the cost of transport relative to good A (in the c.i.f. case, country 2's offer curve shifts, while in the case of f.o.b., it is the offer curve of country 1 which shifts).

In Fig. 6.5 we have considered the c.i.f. curves. Thus, in consequence of the transport costs of good B, OG_1 shifts to OG'_1: if we consider for example the

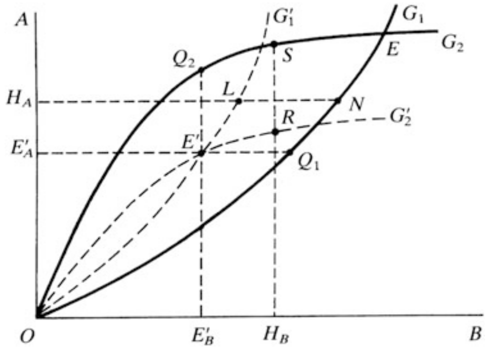


Fig. 6.5 The cost of transport and offer curves

given demand for imports OH_A, country 1 will be prepared to offer $H_A L$ (rather than $H_A N$) of B c.i.f., LN representing the cost of transport: from what we said at the beginning, it will be $LN = c_B \cdot H_A N$ and $H_A L = k_B \cdot H_A N$. Given that c_B is assumed to be constant (and therefore also k_B), the shift of OG_1 towards the A axis will be equiproportional. Similarly, OG_2 will shift equiproportionally to OG_2' as a result of the cost of transport for good A (we have $RS = c_A \cdot H_B S$ and $H_B R = k_A \cdot H_B S$).

The new equilibrium is established at E', and the terms of trade with prices calculated c.i.f. are given by the slope of OE'. Country 1 exports $E_A' Q_1$ of good B, receiving in exchange $E_B' E' = OE_A'$ of good A (so that the domestic price ratio is given by the slope of OQ_1); segment $E' Q_1$ represents the cost of transport of good B, so that the quantity of that good effectively received by country 2 is $E_A' E' = OE_B'$.

Country 2 exports $E_B' Q_2$ of good A and receives in exchange $E_A' E' = OE_B'$ of good B (so that its domestic price ratio is given by the slope of OQ_2); segment $E' Q_2$ represents the cost of transport of good A, so that the quantity of that good effectively received by country 1 is $E_B' E' = OE_A'$.

The difference between the exports of one country and the imports of the other is made up of the quantity of the good exchanged that is consumed as transport cost.

This type of analysis could be extended to an examination of other problems (f.o.b. terms of trade, transport services supplied by only one of the two countries for both goods, and so on: see Mundell, 1957a), but we do not propose to pursue the matter further, partly because a more general analysis should eliminate the simple assumption that transport costs are translated in terms of consumption of the good transported. In effect for that type of analysis the problem would need reformulating in terms of at least four variables (that is, the two transport services in addition to the two goods), so that it would no longer be possible to make use of diagrams, but a complex mathematical treatment would be required (see Sect. 22.2).

We can however offer some general considerations as to the effects of the presence of transport costs. Apart from the obvious fact that (still assuming perfect competition) the price of any good traded will be higher in the importing country than in the country of origin, two other effects deserve mention.

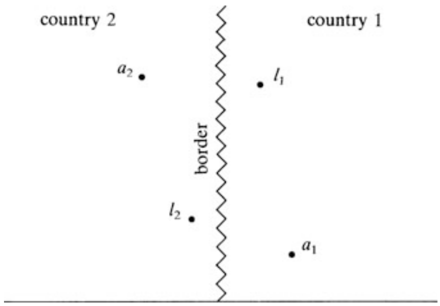


Fig. 6.6 The cost of transport as a determinant of international trade

The first is that *the presence of transport costs can impede the trading in goods which, in the absence of those costs, would be traded internationally*. If there is a gap between the prices of a certain good (expressed in a common unit of measurement) that ensure equilibrium between domestic demand and supply in each of the potential trading countries and if this gap is less than, or equal to the unit transport cost, there will be no international trade in that good. This can be easily be seen in terms of Figs. 6.3 or 6.4 (but the result is also valid in cases other than those shown in these figures): if the unit transport costs are equal to, or greater than, the gap between OR_E and OP_E, good A will not be traded. Transport costs, in other words, can prevent trading in a good that, in theory, should be tradable, just as a decrease in these costs can make a good tradable which had not been previously.

The second result is that *some international trade can be directly caused by the cost of transport* (without this having anything to do with technology, tastes, or factor endowments). Transport costs, then, become a *determinant* of international trade and can explain the apparently paradoxical fact that *a country is sometimes both importer and exporter of the same good*. Let us suppose that the two countries have a long common border and that both produce steel (in mills situated respectively at a_1 and a_2) which they subsequently transform into steel plate (in the mills situated at l_1 and l_2).

Technology, tastes and factor endowments are absolutely identical in the two countries. However, if we assume that, other things being equal, the cost of transport increases with distance, country 1 may find it cheaper to get its supply of steel from a_2, rather than a_1, because a_2 is nearer to l_1 (country 1 thus imports steel from country 2) and, in the same way, country 2 might find it cheaper to import steel from country 1 because a_1 is nearer to l_2 than is a_2 (Fig. 6.6).

Phenomena of this kind can be put in a general framework in the theory of location. Although location theory is beyond the scope of the present work, its relations with international trade and transport costs deserve a treatment, which we postpone to Chap. 16.

6.4 Intermediate Goods

As we have seen more than once, the traditional theory of international trade is based on a model in which two final goods (A and B) are produced employing two primary factors of production (K and L). In reality, production requires not only primary factors, but also intermediate goods. We have already come across intermediate goods in the empirical tests of the Heckscher-Ohlin theory (Sect. 4.6), and we shall meet them again in the theory of tariffs (Sect. 10.7) and in Sect. 6.4.1. One of the refinements of the traditional theory has been explicitly to consider these goods. Actually, the change in the price of a traded intermediate good influences relative factor and commodity priced (Djajić, 1983).

A first way of introducing intermediate goods (Vanek, 1963, Hamilton and Svensson, 1983) is to suppose that each existing product in the economy can be utilized as both an intermediate and a final good. Thus, in our simple model with two goods and two primary factors, the situation will be that good A is produced by using both K and L and certain quantities of itself and/or of good B, in the form of intermediate goods. The same can be said of good B. By subtracting the quantity of it used overall as an intermediate good in the economic system from the amount of good A produced, we have the net production of that good as a final good available to satisfy consumer demand.

Another way to tackle the problem (Batra & Casas, 1973) is to introduce pure intermediate goods, that is goods which are utilized exclusively as intermediate goods and are, therefore, physically different from final goods. Pure intermediate goods may or may not be traded internationally, but it is obviously of more interest when they are.

To deal with the case of intermediate goods which correspond physically to final goods, Samuelson (1965) suggested the expedient of considering the productive system as a "black box" with an input of primary factors of production and an output of the *net* quantity of final goods. The problem then is to define a net production function for each good, that is, a production function which has as its only inputs the total amount of primary factors, and as output the net quantity of each final good. By "total" amount of primary factors we mean the amount *directly* and *indirectly* necessary to produce a given net quantity of the final good. The indirect requirements of primary factors of production refer to the quantity of these primary factors required to produce the intermediate goods which enter into the production of the final good.

It is clear that if the expedient were feasible, one could argue in terms of net production functions; so that—if these have the same properties as traditional production functions, where intermediate goods are assumed absent—the theory of international trade given in the previous chapters would not require any modification.

It has in effect been demonstrated by Samuelson and others (see Sect. 22.3.1) that this is true (provided there are no joint products), so that the four core theorems (see Chap. 5) are still valid even in the presence of intermediate goods.

Those who support the second approach, however, object that in this way we lose sight of the fact that a large slice of international trade concerns those goods (semi-finished products, raw materials, etc.) which are used exclusively as inputs in the production of others goods and are thus *pure* intermediate goods. Traditional theory, further refined by the introduction of net production functions, cannot explain this phenomenon, and this represents a major weakness.

In order to examine the consequences of the second approach, it is necessary at the very least to introduce a third good, the pure intermediate one which is produced (by means of primary factors) exclusively to be used in the production of two final goods. In this case it is also possible to define derived production functions, which connect the production of final goods exclusively with the quantity of primary factors (directly or indirectly) required. So, the traditional theory, reformulated in terms of these new production functions, remains valid.

It is however clear that this method of solving the problem, if formally correct, is something of a piece of wizardry which leaves the initial problem unsolved, that is *how to explain international trade in intermediate goods*. Trade in intermediate goods cannot in fact be explained by reducing the model to a scheme of final goods/primary factors, from which intermediate goods have actually been eliminated! It is therefore necessary to work within the initial scheme with three goods. As the primary productive factors are always the two traditional ones (K and L), we must ask ourselves whether it is possible to classify the goods in order of factor intensity (measured as usual by the capital/labour ratio) and apply the traditional theory in its extended form to more than two goods. The answer is no, unless further qualifications and conditions are introduced and it is easy to understand why.

The traditional theory with two primary productive factors and three final goods is not applicable because the third good is not a final good but an intermediate one and, besides, in the definition of factor intensity, it is necessary to distinguish between *apparent* (or *net*) factor intensity and *total* (or *gross*) factor intensity. Apparent factor intensity is that obtained by considering the quantity of capital and labour directly required in the production of a given good. Total intensity is obtained, on the other hand, by considering the quantity of capital and labour *directly* and *indirectly* required in the production of that given good. The quantities of K and L indirectly required are those which enter into the production of the intermediate good. Total factor intensity, therefore, is obtained from the derived production function defined above. As regards the intermediate good, total and apparent factor intensities coincide, because the indirect requirements of K and L are zero, thanks to the simplifying assumption that the intermediate good itself is produced by means of primary factors only.

It is obvious that the classification of goods can be different according to whether apparent or total intensity is used, so that when the two classifications do not coincide problems arise which prevent the application of the traditional theory (see Sect. 22.3.2).

However, even when there is no discrepancy between the two classifications the structure of trade (that is which of the three goods are exported and which imported) is generally indeterminate, unless further restrictions are introduced.

Let us suppose, for example, following Batra and Casas (1973), that initially international trade in intermediate goods is forbidden. If we assume that there is no discrepancy between the two classifications, we can apply the Heckscher-Ohlin theorem and, having also assumed absence of complete specialization, the factor-price-equalization theorem will be valid (Sect. 4.3). Thus, the intermediate good (given the international identity of the production functions) will have the same price in the two countries. Consequently, once international equilibrium has been established, even if the prohibition of international trade in intermediate goods is eliminated, there will be no incentive for this trade.

However, we cannot exclude the possibility that this trade will take place in some direction,[2] without production and world demand for final goods being (initially) altered. But, as a result of the trade in the intermediate good, the transformation curves of the two countries shift—that of the country which is a net importer of the intermediate good outwards and that of the country which is a net exporter of this good inwards. Let us suppose that country 1 has a relatively plentiful supply of capital and that it is possible univocally to classify good A as the good with relatively high capital intensity. Let us also assume that country 1 exports the intermediate good: then, at the given prices, the shift of the transformation curves means that production of both A and B will decrease in country 1 and increase in country 2.

Consequently (remember that tastes, etc., are internationally identical), it is possible that in the end country 1 will import both good A and good B in exchange for the intermediate good, so that the Heckscher-Ohlin theory (according to which country 1 that has a relative abundance of capital ought to export good A) does not apply. It has been demonstrated by Batra and Casas (1973) that the condition for this theory to apply is that one of the three goods (whether a final or intermediate one) is a non-traded good and, in addition, that the apparent capital intensity of this good lies between the apparent intensities of the two traded goods.

The treatment of intermediate goods carried out in this section has important empirical implications. We have in fact seen in Sect. 4.6, that the studies of Leontief (and his followers) on the Heckscher-Ohlin theorem make use of total (direct and indirect) capital and labour requirements, that is, they take into account what we referred to above as total (or gross) factor intensity. When the intermediate good is not exclusively produced domestically but is (completely or in part) imported, then, to define the total factor intensity of final goods, it is necessary to take account not only of the requirements of capital and labour in producing intermediate goods of domestic origin, but also of the capital and labour requirements in producing goods for export, thanks to which the imported intermediate goods are obtained, by way of international trade (Riedel, 1976; see also Hazari, Sgro, & Suh, 1981, Pt. 2).

[2]Given the assumptions (internationally identical production functions, absence of transport costs, etc.), if the intermediate good has the same price in both countries, then, as we said, there will be no incentive to trade in it, in the sense that it will make no difference to producers of final goods in any country to use the domestically produced or the foreign intermediate good. But precisely because there is no difference, the possibility cannot be excluded that someone might use the nationally produced intermediate good and someone else the foreign produced one.

6.4.1 Intermediate and Capital Goods in the Neoclassical Theory

We know that the traditional theory of international trade in its basic version considers economic systems in which internationally immobile *primary* factors produce, without other inputs, final consumption goods, which are internationally mobile and traded. There is no room, in this version, for produced means of production (fixed and circulating capital). The stock of capital K, which appears in the version under examination, serves only to give it a (illusory) sense of realism: actually, many treatments eliminate the problem by avoiding all consideration of capital and introducing land (clearly a primary factor) as the other factor of production besides labour.

This version of the theory can be all too easily criticized, but it would *not* be correct to conclude from these deserved criticisms, without further analysis, that the whole neoclassical theory is invalid. We must at this point distinguish the problem of intermediate goods (circulating capital) from that of fixed capital goods.

As regards intermediate goods, these can be rigorously introduced into the traditional theory, as we have shown above. This part of the criticism then collapses.[3]

Much more difficult is the problem of fixed capital (henceforth, for brevity, we shall omit the adjective "fixed"), with regard to which two aspects must be distinguished: that of capital as produced means of production and that of capital as a collection of physically heterogeneous goods. If we assume that capital is a single physically homogeneous good (the terminology to indicate it is varied: meccano sets, treacle, jelly, etc.) which is used in conjunction with labour to produce both itself and consumption goods, no particular difficulty arises, and this aspect can be dealt with in the context of the traditional theory, as shown in Sects. 14.1 and 28.1.

The really serious difficulties arise when one must account for the fact that in reality no single physically homogeneous capital exists, but a collection of physically heterogeneous capital goods with varying proportions among themselves (if these proportions were constant, one could easily define a basket of capital goods in the fixed proportions, and consider it as a single homogeneous good).

This aspect will be examined in the next section; it is as well to inform the reader here that it also concerns the new theories of international trade (see Part IV), insofar as they also have to deal with heterogeneous fixed capital.

It is also important to point out that we have briefly dealt with this methodological debate in this chapter because neoricardian theories can be classified as "orthodox" in the sense that they also accept the basic assumptions of the traditional theory of international trade (as contrasted with the "new" theories), namely perfect

[3]Some problems might arise in time phased economies, i.e. in economies where production takes time. In this case a difference in the periods of production could give some trouble; see, however, Ethier (1979); see also Chacholiades (1985). A similar observation holds for the case of a homogeneous fixed capital good.

competition, product homogeneity, constant returns to scale (in the particularly simple form of a set of fixed technical coefficients).

6.4.1.1 The Methodological Debate Between Neoclassical and Neoricardian Theories

The problem mentioned at the end of the previous section is nothing but a reflection, on international trade theory, of the debate which has been going on for many decades regarding the theory of value and distribution. It is outside the scope of the present work to enter into this methodological debate, for which we refer the reader to the sources quoted in the References at the end of the chapter. Our task is briefly to examine the repercussions of this debate on the theory of international trade, hence our treatment will be no more than a very brief guide to the literature.

According to one line of thought (Parrinello, 1970; Steedman, 1979; Steedman Ed., 1979) the impossibility, in the presence of heterogeneous capital goods, of defining a measure of aggregate capital independently of distribution, mines the foundations of the neoclassical theory of international trade and in particular of the Heckscher-Ohlin theorem (it would become logically impossible, in fact, to determine factor intensities and factor endowments) and of the related theorems (factor-price equalization, etc.).

This line of thought therefore attempted to extend to international trade the analytical apparatus used to criticize the traditional (neoclassical) theory of capital and distribution in a closed economy. This apparatus, though set up in relation to the debate mentioned above, is related to the vision of the classical economists, in particular of David Ricardo, and this explains the adjective neoricardian in the title of this section and of the chapter. The main contributions in this direction are undoubtedly interesting, but in this line of thought it is not yet possible to find a complete model which can be considered as *the* neoricardian theory of international trade generally accepted by neoricardians (for a critical evaluation of Steedman (1979) and Steedman Ed. (1979), see Dixit (1981)).

According to a completely opposite line of thought (Ethier, 1979) it is perfectly possible to account for heterogeneous capital goods in the context of the traditional theory of international trade and reformulate its propositions in such a way that they remain valid. As we have seen in Chap. 5, the main propositions of the traditional theory are contained in four basic theorems: the Heckscher-Ohlin theorem, the factor-price equalization theorem, the Stolper-Samuelson theorem, and the Rybczynski theorem. Now, according to Ethier, the presence of heterogeneous capital goods does not vitiate the essence of these theorems, duly reformulated to account for such a presence. The numerous counterarguments of the neoricardian literature implicitly contain violations of the basic assumptions of the traditional model (such as, for example, factor-intensity reversals), so that their results can be fully dealt with in the context of the neoclassical theory: “The four basic theorems of the modern theory of international trade, formulated in a timeless context, are insensitive to the nature of capital and remain fully valid in a time-phased world with a positive interest rate. The numerous counterarguments of recent years are simply old friends in disguise: phenomena that can be (and for the most part have been) fully analysed in timeless models" (Ethier, 1979, p. 236). Nothing new under the sun, then? The neoricardians, of course, do not agree, and criticize Ethier (see Metcalfe & Steedman, 1981), who, however, maintains his position (Ethier, 1981). For a general survey of the controversy between the neoricardian and the neoclassical theory of international trade see Smith (1984). See also Robinson (1954), Sraffa (1960), Samuelson (1962), Various Authors (1966), Spaventa (1968), Garegnani (1970), Harcourt (1972), Hahn (1982), Schefold (1985), Pasinetti (1977, 1981), Mainwaring (1984, 1988, 1991), Chacholiades (1985), and Parrinello (1988).

6.5 Elastic Factor Supply

In traditional theory the supply of factors is assumed completely rigid: in other words, all of the quantity of capital and labour existing in the economy is supplied, whatever the rewards might be. It is a convenient assumption introduced for the sake of simplicity; in effect, if it is removed, the analysis is much more complex. Let us assume that labour supply is elastic with respect to the real wage rate, while retaining the assumption of a rigid supply of capital. We know from micro-economic theory that the labour supply curve is not necessarily upward sloping through its entire range with respect to the real wage rate: even in normal cases it can at a certain point bend back (that is, with further increases in the real wage rate, the supply of labour decreases, for example, because workers opt for more leisure: this point is thoroughly dealt with in Laffer and Miles (1982), chap. 8). This is all that is necessary to create the problems mentioned above, which can be summed up as follows:

1. The supply (production) of goods is no longer necessarily an increasing function of the appropriate relative price. We have seen in Sect. 3.2.1 that the supply of good B increases with the increase in the relative price p_B/p_A, while the supply of good A decreases (an increase in p_B/p_A is equivalent to a decrease in p_A/p_B). In the case of variable labour supply, the supply of goods may have an abnormal behaviour, that is, be a decreasing function of the appropriate relative price.

   An intuitive explanation of this phenomenon follows. Let us consider the productive side of the neoclassical model, which must be modified to take account of the fact that the quantity of labour is determined endogenously, not exogenously, and let us see what the effects of an increase in p_B/p_A are. Let us assume that good B is relatively more labour-intensive: consequently, on the basis of the Stolper-Samuelson theorem (Sect. 5.3), the increase in p_B/p_A causes the real wage rate to increase. Now, if we find ourselves in the backward bending branch of the labour supply function, the increase in real wages will cause a *decrease* in the supply itself. The decrease in labour supply determines, on the basis of Rybczynski's theorem (Sect. 5.4),[4] a decrease in the output of the labour-intensive good (in this case, good B) and an increase in the quantity produced of the other good (A). Note then that, with an increase in p_B/p_A, the supply of B decreases and the supply of A increases.

2. The offer curve can be anomalous, in the sense that there is a greater demand for imports when their price increases and vice versa. This is a possible consequence of the phenomenon described in the previous point (1). Remember (Sect. 3.4.1) that the offer curve is constructed starting from domestic excess supply and demand, so that the demand for imports coincides with the domestic excess demand for the importable good. Let us assume that A is the importable good: normally, the demand for A increases with the decrease in p_A/p_B (that is with the increase in p_B/p_A) and the supply of A decreases with the decrease in p_A/p_B, so that the excess demand for A (the demand for imports) increases with the decrease in its relative price. Let us assume that, for the reasons seen in point (1), the supply of A increases with the decrease in p_A/p_B. If this increase is greater than that of demand, the excess demand for A decreases with the decrease in its relative price and, conversely, it increases with the increase in p_A/p_B. This reasoning ignores possible effects of labour-supply variability on demand. These effects are due to the fact that this variability can produce anomalous effects on income and therefore on demand (for example, an increase in real wage rate that causes a reduction in labour supply can determine a reduction rather than an increase in workers' income). See Sect. 22.4.

   When the offer curves are anomalous, all the results of the pure theory of international trade based on the assumption that these curves are normal must be revised, whence the complications mentioned at the beginning (for example, equilibrium may be unstable).

6.6 Non-traded Goods

In the real world, each country produces goods that are not the object of international trade, that is, goods neither for export nor import. There are plenty of reasons why certain goods are not traded: prohibitive import duties (Sect. 10.3), embargoes (Sect. 10.6.4), prohibitive costs of transport (Sect. 6.3), etc.: all of which may justify the existence of non-traded goods.[5]

[4]Rybczynski's theorem has been discussed with reference to an increase in the quantity of factors, but it has symmetrical validity in the case of a decrease, as can easily be established from the box diagram if a shortening rather than a lengthening of the side representing the quantity of labour is considered.

[5]See, however, Padoan (1977) for an interesting criticism of the concept of non-tradable goods itself.

Alongside these cases, in which barriers to trade are due to obstacles which, if removed or reduced, might result in the goods concerned being traded, there are goods which in any case would not be traded, on account of differences of tastes or for reasons inherent in the nature of the goods (many services, for example, are intrinsically nontradable). According to some economists (for example, Kemp, 1969b, p. 134), in most industrialized nations the amount of non-traded goods represents more than half of the national product.

There thus seems to be a very real need to enrich and extend traditional analysis so as to include non-tradable goods. This means that it is necessary to introduce a third good into the standard two-good model, that is, in fact, the non-tradable good, which is produced by means of the same primary factors (K and L) used in the production of tradable goods.

It is often stated that, while prices of traded goods are determined on the international market (and so, in the case of a small country, are exogenously given), the prices of non-traded goods are determined exclusively by the conditions of domestic supply and demand. This is inexact for the simple reason that—assuming the right conditions occur for absence of factor-intensity reversals—the one-to-one correspondence between relative prices of goods and relative prices of factors (Sect. 4.1.1), together with the assumption of perfect competition and free internal mobility of factors, means that the relative price of the non traded good can be determined precisely, starting from the given terms of trade.

Let A, B, and N be three goods, of which the third is not traded, and let us consider the relative prices of goods B and N with respect to A. Given the terms of trade $p_B/p_A = p^*$, the relative price of the factors (p_L/p_K) used in sectors A and B is determined. This relative price, given the assumption of perfect competition and free domestic mobility of factors, is valid also for the N sector. Consequently, assuming that in the sector of the non-traded good the relation between the relative price of factors and the relative price of the good is also one-to-one, the relative price p_N/p_A is determined.

It is possible to give a simple diagram of this chain determination. Let us assume, for example, that good A has a capital intensity greater than both B and N, so that both the relation between p_B/p_A and p_L/p_K and that between p_N/p_A and p_L/p_K are increasing. The relation between p_B/p_A and p_L/p_K is taken from Fig. 4.5a; in the same way we can obtain the relation between p_N/p_A and p_L/p_K.

In the upper half of Fig. 6.7 we have shown the relation between p_B/p_A and p_L/p_K, while in the lower half we have given that between p_N/p_A and p_L/p_K turned upside down.

Given p^* (terms of trade), the relative price of the factors is determined at OF and so (lower half of Fig. 6.7) the relative price p_N/p_A is determined at OG. It can be seen from the diagram that at every value of p^* there corresponds one and only one value of p_N/p_A. This single-value correspondence will occur even if the relations between the relative price of goods and the relative price of factors are decreasing (either or both), provided that these relations are monotonic (absence of factor-intensity reversals).

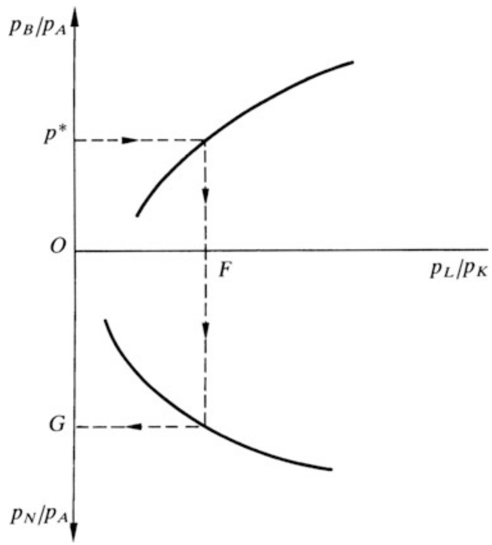


Fig. 6.7 Relative prices of traded and non-traded goods

But there is more to it than that: not only the relative price, but also the *absolute price* of the non-traded good is determined by the international market for traded goods if the factor-price-equalization theorem (see Sect. 4.3) holds. First note that, given p_L/p_K, the optimum factor combination for the production of good N is determined and (given the assumption of first-degree homogeneous production functions) independent of the scale; thus the technical coefficients K_N/S_N and L_N/S_N, where S_N is the quantity of good N produced, once given p_L/p_K, are constant. Now, as in perfectly competitive equilibrium the value of the product is equal to the sum of factor rewards, we have

$$p_N S_N = p_K K_N + p_L L_N, \tag{6.4}$$

from which, by dividing both sides by S_N, we get

$$p_N = p_K \frac{K_N}{S_N} + p_L \frac{L_N}{S_N}. \tag{6.5}$$

The technical coefficients are given, as shown above and, if the factor-price-equalization theorem is valid, p_L and p_K are also given at the level of the corresponding prices of factors in the rest of the world. It then follows from (6.5) that p_N is completely determined.

The statement that the price of non-traded goods are determined exclusively by domestic supply and demand conditions is therefore wrong if approached from the view-point of traditional theory enriched by the introduction of a third sector, which produces a non-traded good. One way to validate this statement—apart from the cases of factor-intensity reversals, etc.—is to drop the assumption of perfect competition, and so admit that factors can have *different* (relative and absolute) prices in the various sectors, and/or that the price of the non-traded good should be fixed without respecting condition (6.5).[6] Another possibility is that there exist *specific* productive factors (see Sect. 6.2) in each sector.

At this point we must ask what is the relevance for international trade theory of the introduction of the non-traded goods sector, seeing that, on the basis of the argument so far, this sector is influenced by, but seems not to influence, the foreign sector? In effect, this impression is false, because the presence of sector N has a considerable influence on the offer curve (relative to goods A and B) of the country considered and thus also on the determination of the terms of trade (once the assumption of the small country is abandoned).

In fact, the presence of sector N can give the offer curve an anomalous behaviour, for example because the demand for imports increases (instead of decreasing) when the terms of trade worsen and decreases (instead of increasing) when the terms of trade improve.

Let us assume that A is the imported good, so that the demand for imports is given by the domestic excess demand for that good. If p_B/p_A increases (this represents an improvement, as p_A/p_B decreases) the excess demand for A in the two-good model increases for two reasons. On the one hand, with normal functions, the increase in p_B/p_A causes an increase in the demand for A. On the other, it causes an increase in the production of B and therefore a decrease in the production of A, which gives up resources to sector B. We shall now see what may happen in the three-good model.

As we have seen above, to every given p_B/p_A there corresponds a given p_N/p_A; let us now assume that when p_B/p_A increases p_N/p_A decreases.[7] The decrease in p_N/p_A, in a context of general equilibrium, also has effects on the demand for A, but to avoid further complications we shall assume that the effect of p_B/p_A prevails anyway, so that the demand for A increases when p_B/p_A increases. We now come to the production side: in a context of general equilibrium the supply of each good is also a function of all the relative prices, but, for simplicity's sake, we shall assume that following the decrease in p_N/p_A the supply of N decreases in any case. This makes resources available which flow into the other sectors, i.e., not only into sector B, but also into sector A (provided the decrease in p_A/p_B is less than the decrease in p_N/p_A, so that the production of A is more profitable than that of N). Thus an increase in the production of A is possible and, if this increase is greater than the increase in demand, the excess demand for this good (that is, the demand for imports) decreases.

In the same way, we can establish the possibility of an increase in the demand for imports when p_A/p_B increases.

---

[6]It is clear that by doing this we move outside the context of the traditional theory: the problems that derive from abandoning the assumption of perfect competition will be dealt with in Part III.

[7]In terms of Fig. 6.7, this means for example that the relationship between p_N/p_A and p_L/p_K is monotonically decreasing rather than monotonically increasing.

The possibility of an abnormal behaviour of the offer curve opens up a whole series of problems which have been dealt with in earlier chapters: for example, international equilibrium can be unstable (Sect. 3.4.2), the Metzler and Lerner cases in the theory of tariffs can occur (Sect. 10.5.2), etc.

It is interesting to note in conclusion that the presence of a non-traded good has an influence on the offer curve in a way similar to what we saw in the case of variable supply of factors examined in Sect. 6.5. This will come as no surprise if we observe (Kemp, 1969b, p. 134) that the non-traded goods sector serves as a sort of reservoir which can release factors to the international sector, or absorb factors from it, in response to variations in prices. Finally, the validity of the four core theorems (see Chap. 5) in the presence of non-traded goods is examined in depth by Ethier (1972).

6.7 Natural Resources, "Dutch Disease", and De-industrialization

The phenomenon of the contraction of the traditional manufacturing sector, due to the rapid expansion of the extractive sector, was observed in various countries and was labelled de-industrialization or "Dutch disease". "Dutch" because it occurred in Holland among other countries, due to the rapid development of the natural gas extractive industry. The same phenomenon was observed in Australia (extraction of minerals), and Britain and Norway (following the extraction of oil from the North Sea).

To analyse this phenomenon on a proper theoretical basis, we must use a model with at least three goods (one exported, one imported and one non-traded good) and certain specific factors in the production of each good besides the traditional unspecific or general factors, which move freely from sector to sector. We can see at once that it is an extremely complex model, not to be dealt with by using traditional diagrams. Still, it is possible to make it less complex and more tractable, by the fairly simple use of an expedient introduced for other purposes by Salter (1959) and subsequently adapted by various economists (Snape, 1977; Corden & Neary, 1982; etc.), for the examination of the problem in hand.

This expedient consists in first assuming that we have to deal with a small country for which, therefore, the terms of trade are given. The relative price of the exported and imported goods is therefore exogenously given for the country in question, so that we can apply Hicks' theorem (1939, 1946) by which, if the relative prices of a given group of goods remain constant as the quantity of the goods themselves varies, the different goods in the group can be treated as a single whole, that is, as if they were a single good.

Thanks to this expedient,[8] we can get a two-sector model: the sectors of traded and non-traded goods. Thus, starting from the three goods A, B, and N[9] (see Sect. 6.6), we can argue in terms of two goods, say, C (all traded goods) and N (the non-traded one).

We now come to the productive factors. Following the specific factors model (see above, Sect. 6.2) we assume that each sector utilizes a specific factor (for example, a particular kind of capital) besides labour, which is the only general factor and moves freely from sector to sector. The price of N is therefore determined by domestic supply and demand, as the presence of specific factors prevents the application of the argument developed in Sect. 6.6.

Let us assume that there is a rapid expansion in the traded goods sector, for example, following a boom due to technical progress in the extraction of natural resources. We must distinguish two effects of the boom (Corden & Neary, 1982). The first is the *resource movement effect*: the boom in the extractive sector causes the marginal productivity of the general factor to grow and attracts it away from the other sectors (the basic model is always that of full employment of factors), with a series of adjustments in the rest of the economy. If the extractive sector uses relative little of the general factor, these adjustments will not be very appreciable, and the second effect will have the greater impact (as happened in Britain: see Corden & Neary, 1982).

The second is the *spending effect*: greater real income from the boom induces a greater expenditure on the various goods (none is assumed to be an inferior good). This in turn causes an increase in the price of N (without influencing the prices of A and B, as these are given by the international market) and a further chain of effects.

To analyse these effects we use the familiar diagram of the transformation curve; given our assumptions, we can argue in terms of goods C and N. In the initial situation, given the conditions of internal supply and demand, a certain price of N is determined with respect to C, for example that given by the slope of $P_h P_h$ in Fig. 6.8, and therefore equilibrium is found at point Q.[10] The boom in the extractive sector causes the transformation curve to shift to $T'T'$: note that, as nothing has happened in the N sector, the intercept with the N axis does not change in the new curve.[11]

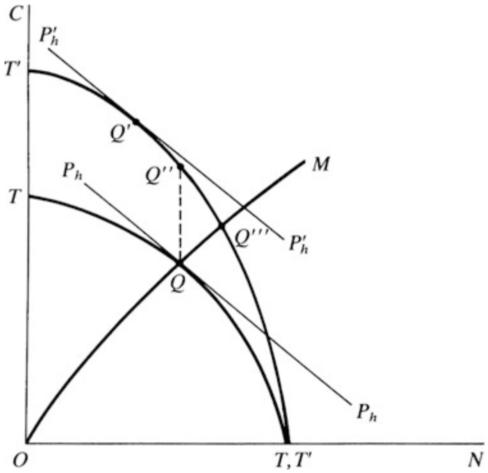
[8]It is self-evident that this expedient cannot be used for a country which enjoys monopolistic power, for example by way of a cartel. On cartels see Sect. 10.6.3; on the role of natural resources in trade models in general see Kemp and Long (1984).

[9]Each of the three goods A, B, and N can in turn be considered as a group inside which relative prices are constant. This explains why we can talk of "good N" and "non-traded goods" without making distinctions.

[10]Note that as N is a non-traded good, in equilibrium the production point and the consumption point coincide. In fact, point Q can also be determined by the tangency between the transformation curve and a social indifference curve, from which the relative price is determined, as is the common slope of the two curves at the point of tangency.

[11]The reader will note the analogy between Figs. 6.8 and 13.11 in Sect. 13.5.2. In effect, the extractive boom can be assimilated to the case when technical progress occurs in sector C.

Fig. 6.8 Effects of an extractive boom on the commodity market



Assuming for the time being that the price of N is unchanged, the new point of equilibrium will be at Q', where $P_h' P_h'$, parallel to $P_h P_h$ is tangent to the new transformation curve. The initial effect of the *movement of resources* is represented by the shift of the production point from Q to Q', with a reduction in the production of non-traded commodities. If we wish to examine the further repercussions by abstracting from the spending effect, we assume that the income elasticity of the demand for N will be zero, so that the income-consumption curve is a vertical line which passes through Q and Q'', to denote the invariability of the demand for good N. By comparing Q'' with Q' it can be seen that there is excess demand for N which brings about an increase in the relative price of that commodity. In the graph, the slope of $P_h' P_h'$ with respect to the N axis increases, so that point Q' moves towards Q''; but without reaching it: with the increase of p_N / p_C, in fact, the demand for N decreases so that equilibrium will be found at an intermediate point between Q' and Q''.

The effect of the resource movement is therefore to reduce the production of good N, though to a lesser degree than the initial reduction.

Let us now consider the *spending effect* and, so as to abstract from the resource movement effect, let us assume that the transformation curve shifts in such a way that, at the given initial relative price, the tangency between $T'T'$ and $P_h' P_h'$ occurs exactly at Q''. Assuming that N is not an inferior good, the demand for it at the given initial relative price increases as a consequence of the increase in income, moving along an income-consumption curve such as OM, which intersects $T'T'$ at Q'''. If we compare Q''' with Q'', we note that there is excess demand for N, which will lead to an increase in the relative price of that commodity, so that point Q'' moves towards Q''', without however actually reaching it, because the increase in p_N / p_C causes the demand to decrease. The point of equilibrium will be between Q'' and Q'''.

The spending effect acts therefore to increase the output of N. The total effect will be given by the sum of the resource movement effect and the spending effect;

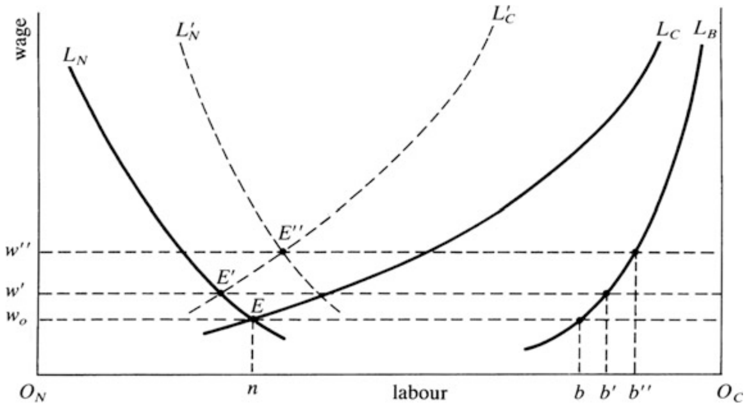


Fig. 6.9 Effects of an extractive boom on the labour market

because one is negative and the other positive the balance can in general have any sign and so the production of N can either increase or decrease. In any case, the production of C increases but for our purposes it is necessary to determine the variations in the outputs of the two traded goods, namely that of natural resources and that of manufactured goods; without any loss of generality, we can say that they are, respectively, commodities A and B.

For this purpose, it is sufficient to look at the *labour market*. In fact (see Sect. 6.2), since we have assumed that labour is the only mobile factor, while the others are specific factors, fully employed in each sector, to find out whether the production in one sector increases or decreases it is sufficient to find out whether employment increases or decreases in that sector. Let us therefore consider Fig. 6.9, where the segment $O_N O_C$ represents the total quantity of labour in existence. The quantity of labour used in sector N is measured from the origin O_N, while that used in sector C is measured from O_C. In the ordinate we show the wage rate (expressed in terms of good B). Curves L_N, L_B, L_C, represent the various functions of demand for labour in the initial situation. The demand for labour is a decreasing function of the wage rate on the basis of the well-known relationship $p_i MPL_i = p_L$ where MPL_i indicates the marginal productivity of labour in sector $i = A, B, N$; p_L is the wage rate and p_i is the price of commodity i. As we have expressed wages in terms of B, in order to draw the curves in Fig. 6.9 it is necessary also to know p_A and p_N. Now, p_A, like p_B, is given by the international market, while p_N is determined by Fig. 6.8. We have therefore all the elements necessary to construct Fig. 6.9. Note that the labour-demand curve shifts upwards both when the marginal productivity of labour increases (due to technical progress) and when the price of the commodity increases.

Let L_B be labour demand in the manufacturing sector; if we add the demand for labour in sector A (not shown in the diagram, so as to simplify) to L_B, we obtain the total demand for labour in the sector of traded goods, L_C. The L_N curve represents instead the demand for labour in the sector of nontraded goods.

Given the assumptions of full employment and mobility of labour (which imply an equal wage rate in all sectors), the wage rate will be w_0 and employment will be $O_N n$ in the sector of non-traded goods and $O_C n$ in the sector of traded goods, of which $O_C b$ in the production of manufactures and bn in the extractive sector.

The boom in sector A is the equivalent of an increase in the productivity of labour in that sector so that, at unchanged prices of the goods (which is the equivalent of considering the movement from Q to Q' in Fig. 6.8), the demand curve L_A shifts (at each given wage there is a greater demand for labour) and the total demand curve in the sector of traded goods shifts from L_C to L'_C. The new equilibrium point in the labour market is E', to which a wage rate w' corresponds; it can also be seen that employment has decreased both in sector N and in sector B, while it has obviously increased in sector A. However, point E' is only a temporary equilibrium point for, as we have seen above, in the final equilibrium situation the price of non-traded goods increases relative to those of traded goods and thus the labour demand curve in sector N shifts towards the right, for example to L'_N, and the wage rate further increases to w''. Employment in sector B decreases further (point b''). In the diagram we have assumed that E'' is to the right of E, so that employment (hence production) in sector N increases, but point E'' could also be to the left of E, so that employment (and thus production) in sector N might also decrease, as we already knew. The important result that we obtain is that in *any case* employment (and so output) in sector B decreases (*de-industrialization*): in fact, as point E'' will in any case be on L'_C to the right of E', point b'' will always be to the right of b'.

It goes without saying that, as the output of C has increased, the output of A—given that the output of B has decreased—must have increased.

We shall now see what happens to factor rewards. The wage rate expressed in terms of manufactured goods increases, but it is uncertain what happens to the real wage rate, if by "real" wage rate we mean workers' purchasing power, that is the nominal wage rate divided by a general price index. As the price of traded goods is a given constant, while the relative price of non-traded goods has increased, the purchasing power of wages in terms of non-traded goods might also have decreased. In fact, if we indicate the nominal wage rate by p_L and since $p_L/p_N = (p_L/p_B)(p_B/p_N)$, the increase in p_L/p_B can be more than compensated for by the decrease in p_B/p_N (if p_N/p_B increases it is obvious that p_B/p_N decreases); it follows from this that p_L/p_N can decrease. We thus have

(i) p_L/p_B increases;
(ii) p_L/p_A increases (as p_B/p_A is given by the terms of trade, if p_L/p_B increases p_L/p_A also increases);
(iii) p_L/p_N can either increase or decrease.[12]

[12]Given that $MPL_N = p_L/p_N$ and that MPL_N is a decreasing function of employment in sector N, it follows that p_L/p_N increases (decreases) if employment and therefore production of sector N decreases (increases).

When p_L/p_N increases, the real wage is bound to increase, but if p_L/p_N decreases, the real wage rate can either decrease or increase, according to the greater or smaller share of non-traded goods in workers' consumption.

As far as the rewards for other factors—the specific factors—are concerned, the only certainty is that the reward (in terms of good B) of the specific capital of sector B decreases. In fact, as employment in this sector decreases, the marginal productivity of specific capital of the sector itself decreases.[13]

In sector N, on the other hand, we do not know whether employment increases or decreases, so that the marginal productivity of specific capital in this sector can either increase or decrease. Also in sector A the reward of the specific factor can move in either direction, insofar as it is necessary to consider, besides the effects of employment, also the effects of technical progress on the marginal productivities of the factors. It is therefore possible (even if this involves not very plausible values of the parameters) for the benefits of the extractive boom to spread to other factors, to the point where there is a decrease in the reward of the specific factor used in the extractive sector.

We can then conclude that a boom in the extractive sector will have the following effects:

1. Production and employment in the extractive sector increase while production and employment in the traditional manufacturing sector decrease (*de-industrialization*); production in the non-traded goods sector, on the other hand, may either increase or decrease;
2. The price of non-traded goods increases. As the price of traded goods is given by the international market, the general price level in the country concerned increases[14];
3. The direction in which the real rewards of the various factors (labour and specific factors) move is usually indeterminate a priori.

It is important to stress the fact that these results have been obtained assuming a *single* general factor that is mobile between sectors, while the others are immobile specific factors. This assumption can be relaxed, for example, by introducing the mobility of capital between the two sectors of traded goods (while the sector of non-traded goods continues to use a specific factor in addition to labour) or even that capital and labour are common factors to all sectors and are freely mobile between these. By modifying the assumptions the results change, and it is no longer certain whether de-industrialization will come about: for a detailed examination of the various possible cases, see Corden and Neary (1982); see also

[13]We recall from the properties of first-degree homogeneous functions—see Sect. 19.1.3—that the marginal productivity of a factor is an increasing function of the quantity of the *other* factor. Thus the marginal productivity of capital decreases (increases) if the quantity of labour employed decreases (increases).

[14]This is inflation of the type contemplated by the so-called Scandinavian model of inflation. See, for example, Lindbeck (1979).

Bruno and Sachs (1982), Long (1983), van Wijnbergen (1984), Corden (1984a), and Findlay (1995, pp. 172–73).

6.8 International Factor Mobility and Trade in Factors

The international immobility of productive factors is, as we know, one of the concepts around which the traditional theory of international trade revolves. In effect, it would be possible to argue that, in a situation of free and perfect international mobility, of both goods and factors, the need for a theory of international trade disappears, as the whole world would become a single integrated system.

In reality there is never perfect international mobility either of goods or factors, but the assumption of absolute immobility of factors is undoubtedly inexact, so that it is important to analyse the consequences of introducing international mobility of factors into traditional theory.

Before going on, however, a few terminological caveats are in order.

Firstly, although ‘international factor mobility’ and ‘trade in factors of production’ are often used synonymously, we prefer to keep them distinct for the following reasons.

International factor mobility remains rooted in the traditional model, in the sense that we are always in the context in which final goods are produced by means of *primary* factors. The only difference from the traditional model is that the assumption of international factor immobility is dropped: factors can freely move at both the national and international level. If, say, capital moves from country 1 to country 2, and labour from country 2 to country 1, we may say for short that country 1 has ‘exported’ capital and ‘imported’ labour, but we must keep in mind that these primary factors are not ‘traded’ in the sense in which commodities are traded.

In fact, as we know from previous chapters, commodity trade depends on the conditions of demand and supply, where supply implies *production* in an essential way. The $2 \times 2 \times 2$ simple general equilibrium model that forms the basis of the traditional theory of international trade is not a pure exchange model, but a model with production and exchange. Primary factors of production, by definition, are not produced. This is why we prefer not to speak of factor trade when we are in the presence of the mere international mobility of primary factors. Both capital and labour can be considered under this heading, land being immobile by its very nature.

Trade in factors, on the other hand, implies that we are dealing with factors which are themselves *produced* means of production and, in addition to being internationally mobile, can be traded as any other good. This practically restricts the picture to (physical) capital in its various forms, both fixed and intermediate.

Our distinction is neither semantic nor whimsical, as it has important consequences. Suffice it to point out that, in the case of mere factor mobility, when

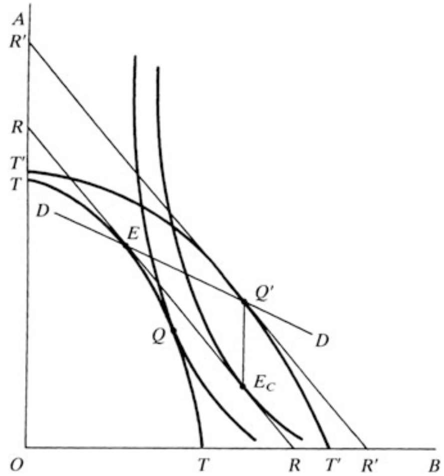


Fig. 6.10 International factor mobility

factor prices are equalized through factor movements, factors *stop moving*. On the contrary, in the case of trade in factors, when the prices of traded factors are equalized through free trade, these factors (in their quality of traded goods) *continue to move* as any other traded commodity.

This is why these apparently equivalent topics are treated in separate sections.

6.8.1 Factor Mobility, and the Theorems of International Trade

Let us assume that productive factors can shift from one country to another as a result of income differentials: each factor thus will tend to move to that country where the reward is highest.[15]

The point of departure in our analysis is therefore the existence of different rewards for the same factor in different countries and that, note, is the same as saying that *the conditions of the factor-price-equalization theorem* (Sect. 4.3) *are not satisfied*. As we know, there are various reasons why factor price equalization may not occur: one of these is the presence of obstacles to international trade which prevent the equality of prices of goods in the various countries. We shall therefore assume that a duty is levied and, to simplify the analysis, that the country which levies the duty is a small one, so that the terms of trade on the international market

[15]Naturally, this does not mean that factors may not shift for other reasons (unemployment in the country of origin; the political or social situation in the country of origin; the possibility of more fully realizing one's own potential, rewards apart—the so-called brain drain—etc.), but only that these reasons cannot be considered by the traditional theory, based on the assumption of full employment, etc.

are unchanged. Proceeding step by step, we begin—see Fig. 6.10—from an initial situation of international free trade and immobility of factors.

The situation in Fig. 6.10 gives rise to the terms of trade represented by the slope of *RR*,[16] to the production point E and the consumption point E_C, so that the country imports commodity B and exports commodity A. Working within the Heckscher-Ohlin model, we assume that this result is due to the fact that the country concerned is relatively abundant in labour and that A is the labour-intensive good. Given the absence of complete specialization and taking the other conditions to be fulfilled, the factor price equalization theorem is valid (Sect. 4.3), so that the real reward (marginal productivity) of each factor is equal at home and abroad.

If at this point all obstacles to international mobility of factors are removed, the factors will not shift, because there are no income differentials. But by introducing a duty the situation changes. Following Mundell (1957b and 1968, Chap. 6) we assume that the duty is prohibitive (we shall see later that the conclusions do not change even when the duty is not prohibitive) and that free international mobility applies only to capital.

The introduction of a prohibitive tariff on B shifts the production point, to coincide with the consumption point, at Q. As the domestic relative price of B has increased, it follows from the Stolper-Samuelson theorem (Sect. 5.3) that the real reward (marginal productivity) of the factor used relatively intensively in the production of B, i.e., capital, increases. Given free international mobility of capital, this will flow from the rest of the world towards the country concerned and will continue to do so until the income differential has disappeared: as the prices of goods and factors remain unchanged in the rest of the world, this means that, in the country concerned, the reward of capital (and, therefore, labour, given the assumption of first-degree homogeneity and international identity of the production functions) must return to its original pre-tariff level.

In the final equilibrium situation, therefore, the income of national factors must be the same as the initial national income and, furthermore, the domestic relative price of goods must be the same as the initial pre-tariff one: in fact, given the assumption of absence of factor intensity reversals, there exists a one-to-one relationship between the relative price of goods and the relative price of factors (Sect. 4.1.1). This means that in the final equilibrium situation the income of national factors coincides with the original isoincome line, *RR*.

Let us now point out the following important implication of free international factor mobility. When the marginal productivities of factors and therefore also their incomes are equalized, both the relative and the absolute prices of commodities must be equalized, given the assumption of international identity of the production functions. This confirms what was said in Sect. 4.3, note 2, that free international mobility of factors constitutes a perfect substitute for free international mobility of

[16] We must also remember that the position of *RR* represents the level of national income, measured by the intersection with the vertical axis (in terms of A) or by the intersection with the horizontal axis (in terms of B).

commodities and leads to the equalization of the prices of the commodities, despite the fact that these are immobile (see, however, Svensson, 1984 and Markusen & Svensson, 1985, for an examination of whether goods trade and factor mobility are necessarily substitutes or may be complements in particular cases). It is, as it were, a "*commodity price equalization theorem*", dual to the factor price equalization theorem.

It is important at this stage to note what happens to income earned by the foreign owners of the capital which flowed in from abroad: for simplicity's sake, we shall assume that it is entirely repatriated to the country of origin, so that income spent in the country we are concerned with always coincides with income received by national factors. This income is clearly less than the value of the product, because a part of the latter is handed over to the foreign capitalists.

The increased production is naturally made possible by the use of a greater quantity of capital which has flowed in from abroad, therefore the transformation curve shifts upwards and to the right (see $T'T'$). When there is an increase in capital, Rybczynski's theorem can be applied (see Sect. 5.4) on the basis of which—with the same factor prices—there is an increase in the domestic production of the commodity which is relatively intensive in its use of the increasing factor (that is commodity B) and a decrease in the production of the other (that is to say, commodity A). This means that Rybczynski's line (see Sect. 14.2) has a negative slope, that is point Q' at which $R'R'$, parallel to RR, is tangent to $T'T'$, must be further down to the right with respect to point E.

We must now demonstrate that the situation represented by Q' (as a production point) and E_C (consumption point) is indeed that of final equilibrium. That the final consumption point is E_C derives from the fact already discussed, that the income of national factors coincides with the initial isoincome line RR and from the assumption that all the income accruing to foreign capital is repatriated, so that the income spent at home must be that accruing to national factors. Consequently, the final consumption point must be identical with the initial one. That the final production point is Q' derives from the fact already discussed that the difference between the value of the product and the income paid to national factors constitutes the reward of foreign capitalists. It is therefore necessary for the country to produce at a point (which must lie along Rybczynski's line) such that, when the foreign capitalists' reward has been deducted, it is able to consume at E_C without trade (given the existence of the prohibitive tariff). Since the difference between the value of the product and the income paid to the national factors can be measured by the vertical distance between $R'R'$ and RR, it becomes clear that by producing at Q', which lies vertically above E_C, and by paying $Q'E_C$ to the foreign capitalists, the country can consume at E_C. At any other point along the DD line, to the right or to the left of Q', the structure of production would not be such as to permit the country to consume at E_C without trade (after the foreign capitalists have been rewarded).

At this point the tariff becomes irrelevant! When the prices of factors and commodities have been equalized between the country in question and the rest of the world, and when the production-consumption situation, given by Q' and E_C, has been stabilized, even if the tariff is eliminated, there is not the slightest incentive to move commodities, so that there will be no international trade, nor any incentive to cause an outflow of the foreign capital.

As we said above, these results do not change even if the initial tariff is not prohibitive: however small the tariff may be, it always leads to the disappearance of trade. Going back to the initial situation, we assume that the tariff introduced is not prohibitive, so that trade goes on. The increased domestic reward of capital causes more to flow in from abroad. Since we have assumed that the country considered, let's call it country 1, is relatively labour abundant, in the initial situation we find $(L/K)_1 > (L/K)_2$. The inflow of K from the rest of the world (country 2) to country 1 leads to a continuous decrease in $(L/K)_1$ and increase in $(L/K)_2$ to the point where the two ratios become equal: once the difference between the relative factor endowments has been eliminated international trade will cease. Another way of getting the same result is to observe that, with the inflow of capital into country 1, the production of the importable good B (which is relatively capital intensive) will grow and the production of the exportable commodity A will be reduced to the point where the structure of production will coincide with the structure of demand (cessation of international trade). One consequence of the outflow of capital is that in country 2 the output of the exportable commodity B (which is relatively capital intensive) is reduced[17] and the production of the importable commodity A increases. Thus in country 2 (the large country compared to the small country 1), the price ratio p_B/p_A increases, once the trade flows have ceased (but not the outflows of capital, because the difference in reward persists) and therefore the marginal productivity of capital (Stolper-Samuelson theorem) increases in country 2 and decreases in country 1, until they are equalized. At this point capital movements also cease.

Among the other causes of international factor movements due to different rewards, we must list complete specialization, factor intensity reversals, etc. The principal conclusions of the analysis are as follows (Kemp, 1964, chap. 9; for further analysis see Sect. 22.7):

1. The removal of impediments to international factor movements gives rise to an improvement in the world productive efficiency;
2. The terms of trade can move in any direction or else remain the same;
3. If at least one of the trading countries levies a duty, then the final equilibrium will be characterized by the absence of trade;
4. If, on the other hand, there is free trade, the final equilibrium will be characterized by an increase in specialization in the various countries compared to the initial situation and at least one country will be entirely specialized.

---

[17]It should be remembered that Rybczynski's theorem is valid for both increases and decreases of a factor: the production of a commodity with a relatively intense use of a factor varies in the same direction as the quantity of this factor.

A related issue is whether the four core theorems of the traditional theory (the Heckscher-Ohlin, factor-price-equalization, Rybczynski, and Stolper-Samuelson theorems) remain valid under the assumption of factor mobility. The answer is yes, provided that the number of goods and mobile factors is at least as large as the total number of factors (Ethier & Svensson, 1986; Wong, 1995, chap. 4; see also the Appendix to the present section).

On international factor movements in general see Jones (1967), Hill and Méndez (1983), Various Authors (1983), Jones and Dei (1983), Ruffin (1984), Norman and Venables (1995), and Wong (1995).

Further light on the question can be thrown by using the specific factors model treated in Sect. 6.2 (see Neary, 1995; Wong, 1995, chap. 4, sect. 4.10). For clarity of exposition we shall separately treat the movements of capital and the movements of labour.

6.8.2 International Movements of Labour (Migration)

To examine the effects of an inflow of labour in the specific factors model, it is expedient to use Fig. 6.2.

As we have seen in Sect. 6.2, an increase in labour, which is the ubiquitous factor, shifts the origin O_B to O'_B. The demand-for-labour schedule in sector B is now L'^D_B, which is the same as the curve L^D_B but referred to the new origin. The equilibrium point shifts from E to E', where the wage is lower. We also note that more labour is employed in *both* sectors ($O'_B L'_E > O_B L_E$, and $O_A L'_E > O_A L_E$), hence an increase in the ubiquitous factor brings about an increase in the output of *both* commodities. Since both industries have more workers but fixed amounts of the respective specific factor, the wage in both industries declines because of the diminishing marginal productivity of labour.

Thus the specific factors model predicts that an inflow of labour will lower the wage in the country where the workers are migrating to. It also predicts that the output of both industries will increase. What about the returns ("rentals") of the specific factors? We begin by observing that the (specific) capital to labour ratio *decreases* in both industries because more labour is employed in each of them. Since the production functions have been assumed to be homogeneous of the first degree, it follows that the marginal productivity of capital is a decreasing function of the capital/labour ratio. Hence a *decrease* in this ratio will cause an *increase* in the marginal productivity of capital.

In conclusion, the owners of (specific) capital will benefit from the reduction in wages due to immigration. Thus we should not be surprised that owners of capital normally support more open borders, that provide them with foreign workers with a consequent reduction in wages.

It should be noted that the above results are valid in the short run (actually, the specific factors model is a short-run model). The long run effects can be analysed in the context of Rybczynski's theorem (see Sect. 5.4). Further analysis of migration is contained in Hazari and Sgro (2001).

As regards the actual migration flows all over the world, see International Organization for Migration (IOM).

6.8.3 International Movements of Capital

The theory of international capital movements focuses on the movement and renting of physical capital, and can be treated much in the same way as we have done for international labour movements. Take for example an inflow of capital specific to sector B (K^B increases), and consider Fig. 6.1.

With constant returns to scale and decreasing marginal productivities, an increase in a factor must have a positive effect on the marginal productivity of the *other* factor (see Sect. 19.1.3), ceteris paribus. This means that for a given p the L_B^D curve shifts upwards, for example to position $L_B'^D$.

As we have seen in Sect. 6.2, the new equilibrium point is E', where less labour is allocated to sector A (hence a lower output of A) and more to sector B (whose output increases both because more labour is employed there and because of the increase in its specific capital).

What about factor rewards? As is obvious from the diagram, the wage rate increases from w_E to w_E'. Consider now sector A. As we have just shown, the amount of labour employed in that sector decreases, which implies a decrease in the marginal productivity of the other factor, namely in the rental of the specific capital K^A.

As regards sector B, we are in the presence of two opposite effects. On the one hand, the marginal productivity of K^B increases because of the increase in the amount of labour employed in sector B. On the other, the marginal productivity of K^B decreases because of the increase in the amount of K^B due to the capital inflow. However, since product prices are assumed fixed, the increase in the wage must be offset by a *decrease* in the rental on capital in both industries,[18] hence the rental of the specific capital K^B also falls.

[18]This follows from the fact that the proportional change in the price of each good is a weighted average of the proportional changes in factor prices in each sector, the weights being the share of each factor in the value of output of that sector. With fixed prices of goods, if a factor price increases, then the price of the other factor must decrease. More formally, consider Eq. 22.6 derived in Appendix 22.1, that we reproduce here for the reader's convenience:

$$\begin{aligned} \theta_{K^A A} p_{K^A}^* + \theta_{LA} p_L^* &= p_A^*, \\ \theta_{K^B B} p_{K^B}^* + \theta_{LB} p_L^* &= p_B^*, \end{aligned}$$

where the $\theta' s$ denote the factor shares in each sector, and the asterisks denote relative changes. With fixed prices of goods, $p_A^* = p_B^* = 0$, so that, given $p_L^* > 0$, both $p_{K^A}^*$ and $p_{K^B}^*$ must be negative.

6.8.4 Foreign Direct Investment and Multinational Corporations

We must now point out that this theory does not cover the phenomenon of multinational corporations (MNC) that carry out foreign direct investment (FDI). Nowadays FDI is absolutely predominant, so that a new theory is called for.

The firm that carries out foreign direct investment is usually a big corporation that operates in a market with high product differentiation. For such a firm, foreign direct investment is often an alternative to exporting its product(s), because the ownership of plants abroad facilitates the penetration in foreign markets. Multinational corporations, also called multinational enterprises (MNE) are firms that undertake foreign direct investment, namely investment by which the firm (called the parent company) acquires a substantial participation in the equity of a foreign firm, or sets up a foreign subsidiary (the controlled foreign firm and the subsidiary are both called affiliates of the parent company).

Direct investment is defined *horizontal* when the foreign affiliate produces goods and/or services similar to those that the parent company produces for its domestic market. It is defined *vertical* when it refers to a geographic *fragmentation* of the productive process in stages. This term identifies the segmentation of a previously integrated productive process in two or more distinct stages, called fragments (or segments) of the productive process, localized in plants situated in different countries. Vertical MNC produce intermediate goods in a country and export them in another country where they are used to produce final goods. In such a case, since the intermediate goods remain within the same firm but cross the border, there is *intrafirm* international trade. According to UNCTAD, a significant percentage[19] of international trade is intrafirm, and the greater part of FDI is horizontal. It is also possible that mixed horizontal-vertical FDI takes place.

There are several reasons for the proliferation of MNC First, the progress in production techniques has made it possible to fragment the production process in distinct segments that can be located in different places. Second, the progress in transport technologies has made less and less expensive the transfers of goods (both intermediate and final) between distant locations. Third, the progress in the service links has facilitated the coordination among the various stages of the productive process. The service links are activities like transport, insurance, telecommunications, quality control, coordination management, that make possible the interaction among the foreign affiliates, and between the foreign affiliates and the parent company. Finally, the improvement in the knowledge of the culture and of the legal and institutional system of other countries has made it easier to set up economic activities (in particular production activities) beyond the national boundaries.

[19] Percentages are subject to change over time. Updated values can be found in UNCTAD's World Investment Report.

6.8.4.1 Types and Determinants of FDI

The starting point of the theory of MNC is the observation that firms which operate in a foreign country bear higher costs than the domestic firms of the foreign country. Therefore, for a firm to become multinational, there must be benefits that offset such higher costs. These benefits are summarized in the classification OLI (acronym of *O*wnership *L*ocation *I*nternalization) due to Dunning (1977, in Ohlin et al.; see also Markusen, 2002), still useful to understand the incentives for a firm to internationalize.

(a) *Ownership* advantages. These advantages are specific to a given firm and consist of the competitive advantage that the firm has over its competitors regardless of its location. Multinational corporations usually own a particular type of capital called *knowledge capital*. It consists of human capital (managers, engineers, financial experts, etc.), patents, know-how, reputation, trademarks, etc. The main characteristics of knowledge capital are:

  1. It can easily be transferred to foreign affiliates at a low cost. For example, managers, engineers and other skilled workers can visit the foreign affiliates or communicate with them from the parent company through fax, phone, e-mail, teleconferencing, etc.
  2. It can be used repeatedly and in different places without depreciating: chemical formulae, blueprints, reputation etc. are very costly to produce but, once created, they can serve the foreign affiliates without losing value or productivity. This means that knowledge capital possesses some of the characteristics of public goods (essentially the non-rivalry in consumption), so that it can be considered as a public input for the firms that owns it.

(b) *Location* advantages. These advantages are specific to a given country or region, and are due to competitiveness in factor prices or to proximity to markets. With production facilities localized near final consumers, multinational enterprises cut transport costs. Furthermore, MNE can decide to localize stages of the production process which are relatively intensive in a certain factor, in a country where this factor is cheaper than in the parent company's country. This advantage is related with the principle of comparative advantage due to different relative factor endowments (see Chap. 4). For example, unskilled labour is normally cheaper in developing countries than in industrialized countries. Therefore, a MNE whose parent company is located in an industrialized country will find it profitable to move the production stages intensive in unskilled labour in a developing country, while keeping the production stages intensive in skilled labour in the parent company's country.

    Finally, location advantages may derive from the possibility of avoiding trade barriers, such as import duties levied by the foreign country. Vertical multinationals may find it optimal to export intermediate inputs and knowledge capital to a foreign affiliate for the assembly, and from there to export the final product to the parent company's country.

(c) *Internalization* advantages. Ownership and location advantages could in principle be also reaped through agreements (such as licences) with foreign firms. However, the same characteristic of knowledge capital that makes it easily transferable also makes it easily dissipated. For example, licencees may absorb the knowledge capital and then defect and set up a business on their own, or they can ruin the trademark's reputation in order to satisfy their greed for gain. Therefore multinational enterprises prefer to transfer know-how etc. internally, to maintain the value of knowledge capital and prevent its dissipation.

6.8.4.2 Effects of FDI

In Sect. 6.8.4 we have examined the effects of a movement of (physical) capital in the context of the specific factors model. However, we have warned that this view does not cover the phenomenon of MNC that carry out FDI. Better to understand this statement, it is enough to consider the fact that a direct investment does *not* necessarily mean an increase in the physical capital stock of the host country. If, for example, the multinational corporation x of country 1 buys the majority of the equities of corporation y in country 2 (previously owned by country 2's residents) the only thing that has happened is an inflow of *financial capital* (the payment for the equities) into country 2, whose stock of physical capital is *exactly the same* as before. It goes without saying that insofar as the multinational x subsequently transfers entrepreneurship, known-how, etc., to y, there will be "real" effects on country 2, but this is a different story. It has indeed been observed that direct investment is strongly industry-specific: in other words, it is not so much a flow of capital from country 1 to country 2 but rather a flow of capital from industry α of country 1 to industry α of country 2.

Here we give a brief treatment of the effects of FDI on the home country and on the host country.

Effects on the home country Exports of the home country may either increase or decrease. They will decrease to the extent that the domestic firm which becomes multinational shifts abroad the production of a commodity that it produced domestically for export. But exports may also increase if the internationalization of the domestic firm is a success and so enables this firm to sell abroad more of the goods whose production has been kept at home.

The effects on domestic employment may also act in two opposite directions. In general, as treated under point b (location advantages) above, a MNE whose parent company is located in an industrialized country will tend to shift the production stages intensive in unskilled labour toward developing countries. Hence there will be a rearrangement of the labour force in the home country against unskilled labour and in favour of skilled labour (employees in the administrative, financial, marketing, R&D sectors, etc.).

Besides, the fact that some enterprises become multinational will have effects on the enterprises that remain domestic. These effects may be both positive and negative. They will be positive to the extent that the internationalization of some enterprises generates externalities on the productivity and the competitiveness of the whole economic system. The MNE, having access to technologies present in the host countries, may "import" them in the domestic country and spread them throughout the domestic productive system. In addition, domestic enterprises may benefit from the situation if they are domestic suppliers of the MNE. In fact, these suppliers see an increase in their business with positive effects on domestic employment. However, it is very likely that the MNE will replace some domestic suppliers with suppliers located in the countries where the MNE has delocalized some stages of the productive process. In this case the domestic enterprises will have to reduce their business with negative effects on domestic employment.

A further effect on the home country is that on the sectorial composition of its productive system. As we have already said, in the case of vertical FDI the various segments of the productive process are shifted to foreign countries where the factor of which these segments make intensive use is cheaper, there will be a sectorial recomposition of the domestic productive system according to the logic of comparative advantage.

Finally, the tax revenue of the domestic country might be negatively affected by the internationalization of domestic firms. In fact, the MNE will shift some of its productive activities to countries where taxation is lower: thus the foreign affiliates' profits will be taxed a first time in the host country and a second time in the domestic country of the parent company but only insofar as they are repatriated and only if the tax rate in the domestic country is higher than that in the host country. In this last case the tax rate will be an average of the rates of the two countries. Hence the tax revenue in the domestic country falls.

Effects on the host country The effects of FDI on the host country are, in the first place, those on its entrepreneurial system and on employment. Among the positive effects we must recall that MNE transfer into the host country technology and managerial skill often not available locally. However, this transfer sometimes does not occur because of the presence of a dual market, one in which MNE operate and the other in which local enterprises operate. The former is characterized by the access to advanced technologies, know-how, contractual power, network of international relations, etc. None of this is available to the latter.

A second category of effects are those called pro-competitive. In general, MNE are considered more efficient than the local enterprises of the host country. Therefore, the operation of MNE in the local market of the host country may stimulate the competitiveness of the local entrepreneurial system. But this positive effect cannot be taken for granted. In fact, it may happen that the entry of much more efficient firms in a preexisting market causes difficulties to the local firms, which are unable to cope with the higher competition and have to leave the market (a *crowding-out* effect).

Then there are the effects on the host country's employment, which are the other side of the coin of the effects seen above on the home country. If the country of destination of FDI is, as it often happens, a developing country, there will be an increase in the demand for unskilled labour. It should however be noted that workers which are unskilled from the point of view of the MNE might be considered skilled from the point of view of the developing country, in the sense that the MNE might in any case request a process of training to perform tasks for which the local workers are not prepared. The effects on employment are ambiguous. Usually there will be a decrease in the employment in the local enterprises (due to the crowding out effect mentioned above) and an increase in the employment in the plants of the MNE.

In addition to the effects on the level of employment, there may be effects on its volatility. MNE are generally considered as footloose enterprises, in the sense that, when the international situation makes it profitable, they can leave the host country since they have no long-run interests there. This implies that the employment generated by MNE in the host country may change in relation to the changes in the international economic situation. The effects on the level of wages are also ambiguous.

Finally, other effects on the host country are:

(i) The exploitation of the local economy, for example when the outflow of repatriated profits is greater than the inflow of FDI;
(ii) The possible decrease in its sovereignty, when the affiliate follows the directives of the parent company rather than those of the local government;
(iii) The possible checkmating of its economic policies (for example a restrictive monetary policy can be nullified by the subsidiary which has recourse to the financial market of the country of residence of the parent company).

6.8.5 Offshoring

The term offshoring refers to the decision by a firm to realize one or more stages of the production process abroad. Such stages may involve physical production of goods (typically intermediate inputs) or instead concern only immaterial services which can conveniently been carried out at distance (such as call centers, accounting services, etc.). A firm may relinquish the ownership of offshored activities (*foreign outsourcing*) or retain ownership (in this latter case we are in the presence of FDI by a multinational enterprise).

We shall return to offshoring in Chap. 17 where we shall study the effect of offshoring on wage inequality. Here we study a simplified version of the model proposed in Grossman and Rossi-Hansberg (2008) which extends the Heckscher-Ohlin set up by including the possibility of offshoring. The model highlights a fundamental trade off: offshoring is attractive for firms because it allows hiring some factors more cheaply abroad than at home but carries higher supervision and coordination costs since the different stages of the production process take place far from each other.

In the present context K and L denote, respectively, skilled labour and unskilled labour instead of capital and labour. Each factor of production performs one and only one type of task. Tasks performed by skilled labour are denoted K-tasks while tasks performed by unskilled labour are denoted L-tasks. Production of each good requires performing each of the L-tasks and each of the K-tasks once. Let N_L and N_K be the number of L-tasks and K-tasks, and let t index tasks and assume $N_L = N_K \equiv N$. Let a_{fi} denote the input of factor f needed to perform a typical f-task in industry i of country 1; where $f = K, L$ and $i = A, B$. Goods have identical technology in terms of task inputs because they all require performing each task once. Nevertheless, goods differ in factor intensity because the parameters a_{fi} differ between goods. This assumption parallels that of different factor intensities between goods typical of the Heckscher-Ohlin model. Factor markets are assumed to be perfectly competitive and factors may freely move between industries though they are immobile between countries. In the present context, offshoring is assumed to be possible for L-tasks only. Let w_1 and w_2 denote the price of unskilled labour in countries 1 and 2, respectively. To make things simple, assume that there is a technology disadvantage of country 2 relative to country 1 represented by the parameter $\chi > 1$. Any task, when performed by firms of country 2 requires a factor input which is χ times the factor input used by firms of country 1. Therefore, any equilibrium of incomplete specialization will be such that $w_1 > w_2$ and such that offshoring, if at all, takes place from country 1 to country 2. Goods markets are perfectly competitive, trade in goods is free; thus goods prices are identical between countries. Factors price equalization does not take place given the technological difference represented by χ.

When a firm offshores a task it uses the technology available to it in its own country. Nevertheless, performing a task abroad comes at an additional costs. Such cost has a generic component which applies to all tasks and a component which differ across tasks. Specifically, when an L-task is performed in country 2 by a firm of country 1 it requires $a_{Li}\beta\delta_t$ units of L. It is assumed that $\beta\delta_t > 1$ for all t so that performing a task abroad requires larger labour input than performing it at home. The parameter β is a shifter that applies to all tasks. It could represent, for instance, the additional cost of communication when passing from face-to-face to remote communication. In this interpretation, a decline in β would represent an improvement in remote communication technology. The parameter δ_t is a parameter specific to task t. It may represent the cost of remote communication related to each specific task over and above the cost β. Tasks are ordered in such a way that $\delta_{t''} > \delta_{t'}$ for any $t'' > t'$. The fact that $\beta\delta_t > 1$ runs against offshoring but lower wages of unskilled labour in country 2 run in favour of it. In equilibrium the following no-arbitrage condition must hold:

$$\underbrace{w_1}_{\text{Unit labour cost in 1}} = \overbrace{\beta\delta_t \underbrace{w_2}_{\text{Unit labour cost in 2}}}^{\text{Overall unit cost of labour in country 2 for firms of country 1}} \tag{6.6}$$

Given wages, Eq. (6.6) determines the task t^* such that the cost of performing the task at home is the same as performing it abroad; that is, it determines the number of offshored tasks.[20] Note that if parameters δ_t were the same for all tasks than either all tasks or none would be offshored. Instead, the fact that parameters δ_t increase with t gives rise to the possibility that some but not all tasks are offshored. Equation (6.6) also shows two quite intuitive relationships. First, the larger the wage difference between countries the larger the number of offshored tasks, ceteris paribus. Second, ceteris paribus, the number of offshored tasks increases as the shift parameter declines.

Consider an initial equilibrium in which both countries produce both goods and in which there is some offshoring. Then consider an improvement in remote communication technology, represented in the model by a fall in β. The first consequence of such fall in β is that the number of offshored tasks increases as we have seen by inspection of Eq. (6.6). This, in turn reduces the marginal cost in both industries. It is interesting to note that the decline in marginal cost due to offshoring is equivalent to an increase in productivity of unskilled labour in country 1. To see this it is convenient to spell out the marginal cost of production in country 1, denoted mc_i^1, which is

$$mc_i^1 = \underbrace{w_1 a_{Li}\left(N - t^*\right)}_{\text{Cost of home}L\text{-tasks}} + \underbrace{w_2 a_{Li}\beta \sum_{t=1}^{t^*}\delta_t}_{\text{Cost of offshored}L\text{-tasks}} + \underbrace{r_1 a_{Ki} N}_{\text{Cost of }K\text{-tasks}} \tag{6.7}$$

The first addendum on the right hand side of (6.7) is the contribution of L-tasks performed at home to the marginal cost; there are $(N - t^*)$ L-tasks performed at home, each of them requires a_{Li} units of L whose unit price is w_1. The second addendum is the contribution to marginal cost of L-tasks performed abroad each of which costs $w_2 a_{Li}\beta\delta_t$. The third addendum is the contribution to marginal cost of K-tasks, where r_1 is the price of skilled labor in 1. From Eq. (6.6) we obtain $w_2 = w_1/\beta\delta_t$ which substituted into Eq. (6.7) yields

$$mc_i^1 = w_1 a_{Li}\left(N - t^* + \frac{1}{\delta_{t^*}}\sum_{t=1}^{t^*}\delta_t\right) + r_1 a_{Ki} \tag{6.8}$$

The term in parenthesis declines as t increases, which makes the marginal cost fall as t increases. Inspection of Eq. (6.8) reveals the channel through which offshoring affects the economy. Since the term in parenthesis multiplies the input coefficients a_{Li}, we can interpret offshoring as a gain in productivity of unskilled labour, it is as if unskilled labour had become more productive in country 1. It is

[20] As an example take $w_1 = 1.8$, $w_2 = 1$, $\beta = 1.5$, $t = 1 \ldots 10$, and $\delta_t = \{1.1,\ 1.2,\ 1.3,\ \ldots,\ 2\}$. Then Eq. (6.6) gives the equilibrium value for δ_t equal to 1.2, to which it corresponds $t^* = 2$. This means that two L-tasks are performed abroad and the remaining eight L-tasks are performed at home. Or, 20 % of L-tasks are offshored.

intuitive then that as a result of a decline in β the world supply of the labour intensive good will increase and that its relative price will decline. In this model, offshoring is equivalent to an increase in productivity of the home factor concerned by the offshoring activity. As per the effects on factors prices they will move according to the usual Stolper-Samuelson mechanism; the fall in the relative price of the L-intensive good runs against the relative price of unskilled labour. The real wage of skilled labour unambiguously increases and the real wage of unskilled labour may increase or decrease depending on whether the "productivity" effects dominates the adverse effect of changes in goods prices. Thus, it is possible that all factors gain from offshoring. Note that this is different from the result of liberalization of trade in goods where the relatively abundant factor gains and the relatively scarce factor loses. The reason is that the scarce factor experience the equivalent of an increase in productivity.

6.8.6 Factor Trade

Capital as a produced and traded means of production has been considered in Sect. 6.4 (as intermediate capital, together with standard immobile primary fixed capital), and will be considered in Sect. 14.1 (as fixed capital). We refer the reader to these sections, where international trade in such capital goods is analysed.

6.9 International Trade under Uncertainty

An implicit assumption in the models of international trade so far examined is that each economic agent should have precise knowledge of all the relevant data as well as the outcome of every action initiated by him. If we look at the neoclassical model treated in Chap. 3, for example, this amounts to the assumption that once the equilibrium price has been determined, production and trade occur immediately and simultaneously or, alternatively, that they take place in the future with certain outcome. In reality all economic activity is permeated by uncertainty and this is particularly true in international trade, where agents often have to make decisions without knowing the precise value of specific and crucial variables, as, for example, the terms of trade. In this regard, one only needs to remember the instability of international prices of raw materials and the consequent problems that it may create for the producing countries, which are often underdeveloped and base their development policy on forecasts as to the income from the export of these raw materials.

It must also be remembered that in the real world many production processes take a certain amount of time, in the course of which stochastic factors beyond the control of economic agents may intervene, in such a way as to alter the expected results radically. The classic example comes from agriculture, where once a certain quantity of inputs have been used, the quantity of produce obtained depends on the weather conditions during the period of production. But problems of uncertainty may exist even on the side of consumption and on that of factor endowments. As far as factor endowments are concerned, adventitious and uncontrollable events may alter them (for example, a flood can put land out of use) and, in the same way, in the field of consumption, demand should be seen as probabilistic (in the above example, a consequence of the flood will be that landowners' income will decrease and so will their consumption of commodities, etc.).

Uncertainty can thus fall indiscriminately on any of the three basic determinants of international trade: technology, factor endowments, and demand. One might well ask whether the results of the international trade models examined in previous chapters hold true even when uncertainty of one sort or another is introduced or—if they are no longer true—whether it is possible to replace them by different, but determinate results. At the present state of the art, there is no satisfactory answer to the question except by making extremely restrictive assumptions. For example, let us consider the Ricardian model treated in Chap. 2. As we know, one of the findings of that model, once the necessary and sufficient conditions for international trade have been met, is that it is to the advantage of each of the two countries to specialize in one of the two goods, and precisely in the one in which the country has a relatively greater advantage (or a relatively lesser disadvantage).

We now introduce uncertainty, but only insofar as it affects production. This means that—using the same symbols as in Sect. 2.2—the quantity of commodity x produced with the employment of a given amount of labour is uncertain and the same applies for the quantity of commodity y. We assume that this state of affairs can be represented formally by introducing a stochastic variable ε (with mean one) in multiplicative form: in other words, as far as x is concerned, we shall have

$$x = \left(\frac{1}{a_1} L_1\right) \varepsilon. \tag{6.9}$$

We now introduce a further simplifying assumption, namely that uncertainty in the production of y can be represented by means of the same stochastic variable, so that

$$y = \left(\frac{1}{b_1} L_1\right) \varepsilon. \tag{6.10}$$

In this extremely simplified case it is obvious that it will be worthwhile for country 1 to specialize in the production of the commodity in which it has a comparative advantage (in our case commodity x). In fact, independently of the value assumed by the stochastic variable, the *ratio* between the quantities depends exclusively on the comparative cost (which is certain), as can be seen from the fact that, by calculating the ratio y/x, the variable ε (which appears in multiplicative form both at the numerator and the denominator), will cancel itself out, so that we shall again have Eq. 2.1 in Sect. 2.2. In other words, the stochastic variable has an influence only on the absolute level of the quantities produced and leaves their ratio unchanged.

The type of uncertainty mentioned is defined in the literature as *scalar uncertainty* and can be applied to any theory of international trade without altering the results. As Dumas (1980) observes, the only difference between a traditional production function

$$Y = F(K, L), \tag{6.11}$$

and a production function affected by scalar uncertainty lies in the introduction of a multiplicative stochastic variable, which causes the quantity of output also to be stochastic:

$$Y_s = \varepsilon_s F(K, L). \tag{6.12}$$

In the last formula the subscript s refers to "states of nature" (supposedly of finite number, say S) to which the various values of the stochastic variable ε correspond.

When technology is affected only by scalar uncertainty, the ratios between the quantities of a given commodity produced in different states of nature will be independent of the input combination, as can readily be seen from the fact that

$$Y_i/Y_j = \varepsilon_i F(K, L)/\varepsilon_j F(K, L) = \varepsilon_i/\varepsilon_j, \tag{6.13}$$

where i and j indicate any two states of nature.

In the case of scalar uncertainty it can be shown (see Sect. 22.8) that all the theorems of the traditional theory remain valid.

Unfortunately, as soon as the assumption of scalar uncertainty is dropped to move to more general cases (so-called *generalized uncertainty*), the situation becomes very complicated and it is not easy to demonstrate the truth of the traditional theorems (see Sect. 22.8). For a general analysis of the traditional trade model under uncertainty see Hoff (1994), see also Casprini (1979), Kemp (1976), Helpman and Razin (1978), Pomery (1979).

6.10 Illegal International Trade and the Economic Theory of Smuggling

The presence of smuggling implies a situation in which there are restrictions to trade (tariffs, quotas, etc.). It is in fact obvious that where there is free trade for all commodities there will be no scope for smuggling.

The traditional opinion was that smuggling, apart from any ethical judgement, improves economic welfare because it constitutes a (total or partial) avoidance of tariffs (or quantitative restrictions, etc.) and amounts to the (total or partial) removal of these obstacles to free trade. This action, like any other removal of restrictions to trade, increases welfare.

This opinion is, however, mistaken for two reasons. First of all, because one must consider that the thesis, according to which the removal of an obstacle to free trade definitely improves social welfare, implicitly assumes, as it does, that this removal in itself is free of costs: this is not the case with smuggling, which obviously involves costs additional to legal trade. In the second place, the basic thesis shows itself to be invalid in the light of the theory of second best (see Sect. 11.6), because in a real situation in which several violations of Pareto-optimum conditions are present, the elimination of any one of these violations may have any effect (positive or negative) on welfare. It is thus necessary to go beyond a generic statement of the above kind and construct appropriate models in which smuggling activity is explicitly incorporated in the traditional theory together with the activities of legal trade. These models have given results which for the time being are not clearly defined. This comes from the fact that the different ways in which smuggling is formally introduced will produce different results.

Like any kind of economic productive activity in the broad sense, smuggling requires the use of resources which involve costs for anyone who undertakes it. The root problem therefore is how to formalize this activity.

A first possible way was introduced by Bhagwati and Hansen (1973; but, as the authors recognize, the basic idea was already contained in an article by Cesare Beccaria in 1764, which was the first attempt to analytically examine smuggling. See also Bhagwati Ed., 1974). They assume that smuggling is an activity which "uses" one (or both) commodities—we are in fact in the context of the standard two-commodity model—and does not utilize productive factors, which means that the real costs of smuggling consist exclusively in the loss of part of the smuggled goods (through confiscation, etc.). Note incidentally that this assumption is similar to the one adopted in the traditional treatment of transport costs (Sect. 6.3). In this way it is possible to remain within the bounds of the two-commodity and two-factor model.

To analyse the effects of smuggling within this framework, it is expedient to take Fig. 10.3 from Sect. 10.5.1 (see that section for the diagram's construction details) and introduce a representation of smuggling into it (Fig. 6.11).

The price charged by smugglers will be intermediate between the international price and the domestic price inclusive of duty. It will be higher than the international price on account of the real costs of smuggling,[21] but it will be lower than the legal domestic price (the international price plus duty) because otherwise consumers would not buy smuggled commodities. We also assume, for simplicity, that the price charged by smugglers is independent of the level of smuggling, so that the illegal domestic price is constant.

It is therefore possible to represent the illegal domestic relative price as the slope, say, of the line $P_S P_S$, which, as we have said, is intermediate between that of *RR*

[21] One should, *de rigueur*, add the smuggler's profits to the real costs of smuggling (these profits disappear if one assumes that there is a situation of perfect competition between the smugglers themselves), in which case, to avoid problems of the assessment of the welfare associated with those profits, one may assume that the smugglers are non-residents.

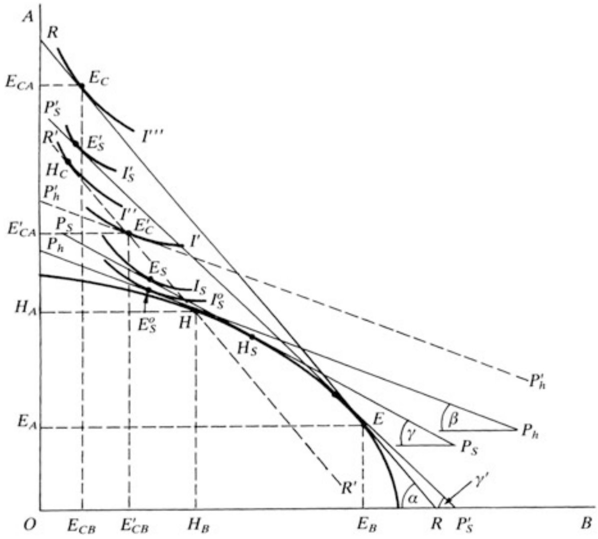


Fig. 6.11 Smuggling and social welfare

and P_hP_h as $\tan\beta < \tan\gamma < \tan\alpha$.[22] As the public can now trade in the two commodities at the relative price $\tan\gamma$, the production point will be H_S and the consumption point E_S, to which there corresponds an indifference curve I_S lower than I'. Smuggling has thus caused a reduction in welfare. But it is quite possible–still on the condition that the illegal domestic relative price is intermediate between the legal one and the international relative price–that the said illegal domestic relative price is $\tan\gamma'$, in which case the consumption point will be E'_S, with welfare, represented by I'_S, *greater* than I'. It is clear that the nearer the illegal domestic relative price is to the legal one (i.e., the greater are the costs of smuggling)[23] the more probable it becomes that smuggling will bring about a reduction in welfare. In effect it is possible to state that in the purely hypothetical case in which the relative price of smuggling is equal to the legal domestic relative price, there would surely

[22]We must remember that A is the imported commodity, so that the tariff and the smuggling, which imply a greater domestic price for A than its international price, cause the relative domestic price p_B/p_A to be lower than the international price. Note that, while in the case of a tariff international trade takes place at the given international terms of trade and consumers react to signals received from the domestic relative price (see Sect. 10.5.1), in the case of smuggling, international trade also gives rise to the same domestic relative price, given the assumption that part of the commodities, after being traded on the international market, is lost through smuggling.

[23]And the greater are the smugglers' profits, see footnote 21.

be a reduction in welfare, as one can see from the fact that when $P_S P_S$ coincides with $P_h P_h$ the consumption point is E_S^0 which is on I_S^0 lower than I'.[24]

In the model we have given, smuggling and legal trade cannot coexist. In fact, when the illegal domestic relative price is more favourable to consumers than the legal price, everyone will turn to the smugglers and, on the assumption that these will not modify their price, legal trade will disappear. In reality this does not happen, and the co-existence of legal and illegal trade can be introduced into the model under examination in various ways, for example, by assuming that the price charged by the smugglers is increasing with the increase in the amount of smuggling on account of increasing costs. In the case of co-existence, it has been shown (Bhagwati & Hansen, 1973) that smuggling necessarily causes a reduction in welfare.

The analysis has been carried out so far without any account taken of the purposes for which the tariff was introduced (that is, by limiting the argument to a discussion of the de facto situation, in which the tariff is present as a historical accident). But it may also be assumed that the tariff was introduced for very precise ends, for example to protect a national industry from outside competition (see Sect. 11.2) and to obtain a given level of domestic production of the commodity in question. It can be seen then that a tariff in the absence of smuggling—while still suboptimal—is better than a tariff in the presence of smuggling (Bhagwati & Hansen, 1973). In that case, smuggling causes a reduction in welfare, as can be seen intuitively from the fact that its presence prevents (totally or in part) the achievement of the objective of production.

A second way in which smuggling can be analysed (Sheikh, 1974) is to assume that smuggling—in addition to the costs due to the risk of confiscation, etc., of the commodities smuggled—also implies the use of the same primary factors of production (capital and labour) employed in legal activities. This use is in any case indirect, in the sense that there is a third commodity produced with these factors, which is then utilized exclusively to make the smuggling possible (one can imagine for example a specific activity of transport used for smuggling: then, besides commodities A and B, we shall have commodity C).

With this way of introducing smuggling, the results obtained by Bhagwati and Hansen are no longer valid. In particular, it is no longer true that there are some cases in which smuggling necessarily reduces welfare (the case of the co-existence of legal trade and smuggling and that of a tariff introduced for a production objective), because it can be seen that also in these cases smuggling can both worsen and improve welfare. The difference in results is due to the fact that, as we are now dealing with a two-factor and three-commodity model (see above), the activity of smuggling modifies the form of the transformation curve, so that the quantities

[24] This is on the assumption that all trade is carried out by way of smuggling. If, on the other hand, legal trade and smuggling co-exist, the consumption point will be intermediate between E_S^0 and E'_C and therefore, in this case also, welfare will certainly be less than that represented by I'. The assumption of equality between the relative price of smuggling and the legal relative price is nevertheless purely hypothetical.

obtainable of the two commodities A and B are no longer definable independently of the total amount of smuggling (and therefore of the third commodity, C).

On the other hand, the fact that, by modifying the initial assumptions, we obtain a different result should come as no surprise: as usual in economic theory, by changing the structure of the model, the results may change, and the problem we are examining is no exception to that rule.

So far we have dealt with smuggling in the narrow sense, but in reality there are many other forms of illegal transactions in international trade, which might be defined as "quasi-smuggling". For example, over- and under-invoicing in the course of otherwise legal commercial transactions.

This means not only that legal and illegal trade exist side by side, but that quasi-smuggling is practised by the operators of legal trade themselves. In some countries, for example Indonesia, a great deal of the smuggling that goes on (which, unlike that analysed above, is *export smuggling*), is in fact practised by the legal exporters themselves. Legal export activity therefore provides a cover for illegal export activities: in economic terms, legal trade may be considered as an input into the smuggling activity. This idea has been formalized in some studies (see, for example, Pitt, 1981) from which it has emerged, yet again, that smuggling can both reduce and increase welfare.

It seems therefore necessary to conclude that, in general, smuggling can have either positive or negative effects on social welfare. For a general survey see Bhagwati (1981); see also Martin and Panagariya (1984), Norton (1988), and Fausti (1992).

References

Barone, E. (1908). *Principi di economia politica* (Pt. III, par. 79).

Batra, R. N., & Casas, F. R. (1973). Intermediate products and the pure theory of international trade: A Neo-Heckscher-Ohlin framework.

Beccaria, C. (1764). Tentativo analitico sui contrabbandi.

Bhagwati, J. N. (Ed.). (1974). *Illegal transactions in international trade: Theory and measurement*.

Bhagwati, J. N. (1981). Alternative theories of illegal trade: Economic consequences and statistical detection.

Bhagwati, J. N., & Hansen, B. (1973). A theoretical analysis of smuggling.

Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade*.

Bruno, M., & Sachs, J. (1982). Energy and resource allocation: A dynamic model of the "Dutch Disease".

Casprini, F. (1979). L'incertezza nella teoria pura del commercio internazionale.

Chacholiades, M. (1985). Circulating capital in the theory of international Trade.

Corden, W. M. (1984a). Booming sector and Dutch disease economic: Survey and consolidation.

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialization in a small open economy.

Cunynghame, H. (1904). *A geometrical political economy* (84 ff).

Dixit, A. K. (1981). The export of capital theory.

Djajić, S. (1983). Intermediate inputs and international trade: An analysis of the real and monetary aspects of a change in the price of oil.

Dunning, J. H. (1977). Trade, location of economic activity and the MNE: A search for an eclectic approach.

Dumas, B. (1980). The theorems of international trade under generalized uncertainty.

Ethier, W. J. (1972). Nontraded goods and the Heckscher-Ohlin model.

Ethier, W. J. (1979). The theorem of international trade in time-phased economies.

Ethier, W. J. (1981). A reply to professors Metcalfe and Steedman.

Ethier, W. J., & Svensson, L. E. O. (1986). The theorems of international trade with factor mobility.

Fausti, S. (1992). Smuggling and parallel markets for exports.

Findlay, R. (1995). *Factor proportions, trade, and growth.*

Garegnani, P. (1970). Heterogeneous capital, the production function and the theory of distribution.

Grossman, G.M., & Rossi-Hansberg, E. (2008). A simple theory of offshoring.

Haberler, G. (1936). *The theory of international trade with its applications to commercial policy* (chap. 12).

Hahn, F. (1982). The neo-ricardians.

Hamilton, C., & Svensson, L. E. O. (1983). Should direct or total factor intensities be used in tests of the factor proportions hypothesis?

Harcourt, G. C. (1972). *Some Cambridge controversies in the theory of capital*.

Hazari, B. R., & Sgro, P. M. (2001). *Migration, unemployment and trade*.

Hazari, B. R., Sgro, P. M., & Suh, D. C. (1981). *Non-traded and intermediate goods and the pure theory of international trade*.

Helpman, E., & Razin, A. (1978). *A theory of international trade under uncertainty*.

Hicks, J. R. (1939, 1946). *Value and capital* (chap. II, sect. 4; Mathematical Appendix, sect. 10).

Hill, J. K., & Méndez, J. A. (1983). Factor mobility and the general equilibrium model of production.

Hoff, K. (1994). A reexamination of the neoclassical trade model under uncertainty.

IOM (International Organization for Migration), *World migration report*, published yearly and available at http://www.iom.int/jahia/jsp/index.jsp.

Jones, R. W. (1967). International capital movements and the theory of tariffs and trade.

Jones, R. W. (1971). A three-factor model in theory, trade, and history.

Jones, R. W., & Dei, F. (1983). International trade and foreign investment: A simple model.

Kemp, M. C. (1964). *The pure theory of international trade* (chaps. 7, 9, 10, 13, 14).

Kemp, M. C. (1969b). *The pure theory of international trade and investment* (chaps. 5–7, 9).

Kemp, M. C. (1976). *Three topics in the theory of international trade: distribution, welfare and uncertainty* (Pt. III, chaps. 20–22).

Kemp, M. C., & Long, N. V. (1984). The role of natural resources in trade models.

Laffer, A. B., & Miles, M. A. (1982), *International economics in an integrated world* (chap. 8).

Lindbeck, A. (Ed.). (1979). *Inflation and employment in open economies*.

Long, N. V. (1983). The effects of a booming export industry on the rest of the economy.

Mainwaring, L. (1984). *Value and distribution in capitalist economies*.

Mainwaring, L. (1988). A neo-ricardian analysis of international trade.

Mainwaring, L. (1991). *Dynamics of uneven development*.

Markusen, J. R. (2002). *Multinational firms and the theory of international trade*.

Markusen, J. R., & Svensson, L. E. O. (1985). Trade in goods and factors with international differences in technology.

Martin, L., & Panagariya, A. (1984). Smuggling, trade and price disparity: A crime-theoretic approach.

Metcalfe, J. S., & Steedman, I. (1981). On the transformation of theorems.

Mundell, R. A. (1957a). A geometry of transport costs in international trade theory.

Mundell, R. A. (1957b). International trade and factor mobility.

Mundell, R. A. (1968). *International economics* (chaps. 5 and 6).

Neary, J. P. (1995). Factor mobility and international trade.

Norman, V. D., & Venables, A. J. (1995). International trade, factor mobility, and trade costs.

Norton, D. A. G. (1988). On the economic theory of smuggling.

Ohlin, B., Hesselborn, P., & Wijkman, P. M. (Eds.). (1977). *The international allocation of economic activity*.
Padoan, P. C. (1977). Una nota sulla distinzione fra beni “commerciabili” e beni “non commerciabili”.
Parrinello, S. (1970). Introduzione a una teoria neoricardiana del commercio internazionale.
Parrinello, S. (1988). On foreign trade and the Ricardian model of trade.
Pasinetti, L. L. (1977). *Lectures on the theory of production*.
Pasinetti, L. L. (1981). *Structural change and economic growth—A theoretical essay on the dynamics of the wealth of nations*.
Pitt, M. A. (1981). Smuggling and price disparity.
Pomery, J. (1979). Uncertainty and international trade.
Riedel, J. (1976). Intermediate products and the theory of international trade: A generalization of the pure intermediate goods case.
Robinson, J. (1954). The production function and the theory of capital.
Ruffin, R. J. (1984). International factor movements.
Salter, W. E. G. (1959). Internal and external balance: The role of price and expenditure effects.
Samuelson, P. A. (1954). The transfer problem and transport costs, II: Analysis of effects of trade impediments.
Samuelson, P. A. (1962). Parable and realism in capital theory: The surrogate production function.
Samuelson, P. A. (1965). Equalization by trade of the interest rate along with the real wage (sect. VI).
Samuelson, P. A. (1971). Ohlin was right.
Schefold, B. (1985). Cambridge price theory: Special model or general theory of value?
Sheikh, M. A. (1974). Smuggling, production and welfare.
Smith, M. A. M. (1984). Capital theory and trade theory.
Snape, R. H. (1977). Effects of mineral development on the economy.
Spaventa, L. (1968). Realism without parables in capital theory.
Sraffa, P. (1960). *Production of commodities by means of commodities*.
Steedman, I. (1979). *Trade amongst growing economies*.
Steedman, I. (Ed.). (1979). *Fundamental issues in trade theory*.
Svensson, L. E. O. (1984). Factor trade and goods trade.
von Thünen, J. H. 1966 (1826), *Von Thünen’s isolated state*.
UNCTAD (United Nations Conference on Trade and Development), World investment report, published yearly and available at http://www.unctad.org/en/Pages/Publications.aspx
Vanek, J. (1963). Variable factor proportions and interindustry flows in the theory of international trade.
van Wijnbergen, S. (1984). The “Dutch Disease”: A disease after all?
Various Authors, (1966). Paradoxes in capital theory: A symposium.
Various Authors, (1983). International factor mobility: A symposium.
Wong, K.-Y. (1995). *International trade in goods and factor mobility*.

Part III
New Explanations for International Trade

Chapter 7
An Overview

7.1 Introduction

The paradigms treated in the previous chapters make up a consistent doctrine in which from certain basic premises various theorems are deduced, concerning both positive and normative economics. This is the doctrinal body with which the "traditional" theory of international trade is nowadays identified.

Leaving aside the assumptions specific to each model, the fundamental assumptions of this theory are:

(i) Perfect competition obtains;
(ii) The commodities which are internationally traded are homogeneous, and identical in the various countries. This means that the homogeneous commodity A produced in country 1 is identical to the homogeneous commodity A produced in any other country, and so on for all commodities.

However, even a casual observation of reality shows that:

1. Market forms different from perfect competition (such as monopolistic competition and oligopoly) are the norm rather than the exception;
2. Product differentiation is much more frequent than product homogeneity.

Although these aspects had already been examined in isolated pioneering contributions, it was only in the late 1970s and in the 1980s that they received due attention and were tackled with an analytical apparatus (partly drawn from industrial economics) comparable to that used in the traditional theory. Thus, the models of the 'new' theories of international trade (also called the *industrial organisation approach* to international trade) were born. We use the plural, because—unlike the traditional theory—there is not *one* new theory but several, with different assumptions and results. Although the adjective 'new', quite appropriate at the time, may now—after over three decades—appear incongruous, we shall continue using it for convenience's sake.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_7,

Box 7.1 Measuring International Specialization and IIT

The concept of *revealed comparative advantage* (RCA) is widely used in practice to determine a country's weak and strong sectors. The most frequently used index in this respect is called the Balassa index (Balassa, 1965). This measure captures to what extent a country exports more of a product than the average country. Given a group of reference countries the Balassa index basically compares the share of the product category in that country's exports to the share of that product category in the reference group (for example the overall world exports). In particular, if X_j^i is country i's export value of industry j, X_j^{ref} is industry j's export value for the reference countries, X^i are the total exports of country j, and X^{ref} the total exports of the group of reference countries, then country i's Balassa index of RCA for industry j, BI_j^i can be written as follows:

$$BI_j^i = \frac{X_j^i / X^i}{X_j^{ref} / X^{ref}}.$$

A value of $BI_j^i > 1$ (< 1) suggests that country i has a comparative advantage (disadvantage) in industry j. The larger the BI value, the higher the degree of comparative advantage.

However, BI turns out to produce values which are asymmetric around 1, because the index ranges from 0 to 1 (if a country is said not to be specialized in a given sector), while it ranges from 1 to infinity (if a country is said to be specialized in that sector). To obtain symmetric values an adjusted (or normalized) index is calculated as

$$(BI - 1)/\,(BI + 1)\,,$$

that ranges from -1 to $+1$. Similar to the export pattern, the structure of a country's imports may likewise contain useful information about a country's comparative-disadvantage situation. Therefore it can be calculated a similar index for a country's import side, the *revealed comparative disadvantage* (RCDA).

The *degree of intra-industry trade* (IIT) is commonly measured by Grubel and Lloyd's index. Grubel and Lloyd (1975) defined IIT as the value of exports in an industry which is exactly matched by imports in the same industry. Its value is measured by:

$$G_i = (X_i + M_i) - |X_i - M_i|\,,$$

where G_i is the value of intra-industry trade and X_i and M_i are the values of exports and imports of industry i, or a given country for a given period. To perform easy comparisons across countries and industries, the values of the index can be expressed as a percentage of each industry's (or country's) combined exports and imports:

$$G_i = \frac{(X_i + M_i) - |X_i - M_i|}{X_i + M_i} \times 100.$$

This measure ranges from 0 to 100, with higher values representing higher levels of IIT.

The common feature of these theories is that they drop the assumption of perfect competition and/or of product homogeneity.

Two additional features are often stressed as peculiar to the new trade theories: the explanation of intra-industry trade and the use of increasing returns to scale.

The first amounts to saying that the new theories can explain *intra-industry* trade while the traditional theory cannot. *Intra-industry* trade (also called *horizontal trade*, *two-way trade*, *cross-hauling*) is defined as the simultaneous import and export of commodities belonging to the same industry. For example, country 1 simultaneously exports and imports commodity A or, more precisely, similar goods belonging to the same category defined as A (see Sect. 8.5). Now, so the conventional opinion continues, the kind of international trade considered by the traditional theory can only be of the *inter-industry* type, i.e., exchange of products of different industries. In our 2×2 setting, this means that country 1 imports one commodity, say commodity A, and exports the other (commodity B), while country 2 imports B and exports A.

In fact, according to the traditional theory, a country cannot export and import the same good at the same time (see Sect. 3.2.3, Eq. 3.17). Therefore, this theory cannot explain international trade of the intra-industry type, which is a huge limitation because intra-industry trade is a an important part of international trade (the greater part at the European level).

This opinion, however, does not seem to be acceptable. We already know the case of transport costs as determinants of intra-industry trade (see Sect. 6.3). Other explanations of intra-industry trade can be given in the context of the traditional theory (see below, Sect. 8.5), and it has even been claimed (Davis, 1995, and Sect. 8.5) that by putting together the Ricardo and Heckscher-Ohlin approaches it is possible to give a general explanation of intra-industry trade in the context of the traditional theory.

As regards the second feature, it is claimed that the new theories can accommodate increasing returns to scale while the traditional theory cannot. This is certainly not true if we consider increasing returns to scale due to external economies, which are perfectly compatible with the traditional theory (see Sect. 3.5). Only increasing returns to scale due to internal economies are incompatible with perfect competition and hence with the traditional theory. Besides, the identification between increasing returns to scale and the new theories is wrong for an additional reason: as we shall see, there are new trade theories that take production as occurring under constant returns to scale.

Be it as it may, these points do not touch the main innovation of the new trade theories, which is their focus on differentiated products and/or imperfectly competitive markets (these theories can of course also explain intra-industry trade, as we shall see in Chap. 9).

It is important to point out, to conclude this introduction, that the new trade theories have significant consequences on all aspects of our discipline. These are briefly summarized below, before going on to a detailed treatment in Chap. 9.

Table 7.1 Traditional theory and the new theories of international trade

	Markets		
Products	Perfect competition	Monopolistic competition	Oligopoly
Homogeneous	Traditional theory	—	Brander (1981)
Vertically differentiated	Neo Heckscher-Ohlin theories (Falvey, 1981)	—	Shaked and Sutton (1984)
Horizontally differentiated	—	Demand for variety (Krugman, 1979, 1980); Demand for characteristics (Lancaster, 1980)	Eaton and Kierzkowski (1984)

7.2 Theory

We have stressed that there is not *one* new theory but several, with different assumptions and results. Table 7.1 gives an overview of the field.

In this table—taking the traditional theory as the reference point—we have classified all the new theories according to two main elements: the type of good and the market form. The names of the authors are merely exemplificative, given the host of contributions now existing (many of which are collected in Grossman Ed., 1992).

About the market form it is sufficient to remark that in the "oligopoly" heading we include not only duopoly but also, as a limiting case, monopoly. About the differentiation of the product, it is instead as well to clarify the terminology.

Vertical differentiation refers to products that differ only in the *quality*. For example, woollen suits that are identical except for the quality of the wool.

Horizontal differentiation refers to products of the same quality that differ in their (real or presumed) characteristics. For example, woollen suits made of the same quality of wool but of different cut and colour.

In the case of vertical differentiation, it is incontrovertible that all consumers prefer higher-quality to lower-quality goods. This, of course, presupposes the existence of universally accepted criteria for evaluating the quality. Hence, in the absence of budget constraints, all consumers would demand the highest-quality good (the assumption is that the price of a commodity increases as its quality increases). It follows that the demand for different commodities, i.e. commodities of different quality, is related to different income levels of consumers.

In the case of horizontal differentiation, the various characteristics are valued differently by different consumers (there are those who prefer a colour and those who prefer another; those who prefer a cut and those who prefer another, etc.). In any case, consumers generally love *variety* (even the person who prefers a certain colour will usually own suits of different colours rather than all of the same colour). It follows that the demand for different commodities, i.e. commodities having different characteristics, is related to love for variety and/or to different subjective evaluations of the characteristics, as we shall show in Sect 9.2. Actually, most commodities can differ in both quality and characteristics, but for analytical convenience we keep the two cases distinct.

Given the greater realism of the assumptions underlying the new theories, shouldn't we drop the traditional theory as irrelevant? The answer is given by Paul Krugman, one of the founders of the new theories. If one were asked to give an actual example of the new theory of international trade with respect to the traditional theory, one could say that "conventional theory views world trade as taking place entirely in goods like wheat; new trade theory sees it as being largely in goods like aircraft. Since a good part of world trade *is* in goods like wheat, and since even trade in aircraft is subject to some of the same influences that bear on trade in wheat, traditional theory has by no means been disposed of completely. Yet the new theory introduces a whole range of possibilities and concerns" (Krugman, 1990, pp. 1–2).

We have mentioned above the existence of precursors. These authors, though not giving a rigorous analytical treatment of the problems, set forth the basic ideas. Ideas that were later taken up, explicitly or implicitly, by most models classified in Table 7.1. We shall first examine these pioneering contributions (see Chap. 8), and then treat the models of the table.

7.3 Policy

The policy consequences of the new explanations of international trade will be examined in Sect. 10.8. See also Baldwin (1992), Guerrieri and Padoan (1996), Haberler (1990), Markusen and Venables (1988), Markusen et al. (1995), Pomfret (1992), Puga and Venables (1997).

7.4 Growth

Growth in the basic neoclassical model is exogenous (see Chap. 15). On the contrary, the new growth theory (for a general treatment see Aghion and Howitt (1998), Barro and Sala-i-Martin (2004), Long and Wong (1998), Romer (1994), and Solow (1992)) stresses the *endogenous* determination of technical progress, which actually means an endogenous determination of the main source of growth (hence the name of endogenous growth theory). The basic ideas were already present in the traditional neoclassical growth theory, but in endogenous growth theory they are at the centre of the stage.

Another point emphasized by endogenous growth theory is the absence of decreasing returns to capital. Hence from the point of view of the interrelations with international trade, endogenous growth is often associated with the new trade theories, that usually take increasing returns and imperfect competition as their points of departure. This topic will be examined in Sect. 15.3.

7.5 Location Theory and Trade

We shall deal with the relations between location theory and trade in the context of the traditional theory in Sect. 16.2. By adding economies of scale to the picture, a richer 'story' of geographical concentration and core-periphery relations can be proposed (Krugman, 1991). This topic will be examined in Sect. 16.3.

References

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*.

Baldwin, Robert E. (1992). Are economists' traditional trade policy views still valid?

Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth*.

Brander, J. A. (1981). Intra-industry trade in identical commodities.

Davis, D. R. (1995). Intra-industry trade: A Heckscher-Ohlin-Ricardo approach.

Eaton, J., & Kierzkowski, H. (1984). Oligopolistic competition, product variety and international trade.

Falvey, R. E. (1981). *Commercial policy and intra-industry trade*.

Grossman, G. M. (Ed.). (1992). Imperfect competition and international trade.

Guerrieri, P., & Padoan, P. C. (1996). Strategic trade policy: A review of the theoretical debate.

Haberler, G. (1990). Strategic trade policy and the new international economics: A critical analysis.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade.

Krugman, P. R. (1990). *Rethinking international trade.*

Krugman, P. R. (1991). *Geography and trade*.

Lancaster, K. (1980). Intra-industry trade under perfect monopolistic competition.

Long, N. V., & Wong, K.-Y. (1998). Endogenous growth and international trade: A survey.

Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: theory and evidence* (chap. 17).

Markusen, J. R., & Venables, A. J. (1988). Trade policy with increasing returns and imperfect competition: Contradictory results from competing assumptions.

Pomfret, R. (1992). International trade policy with imperfect competition.

Puga, D., & Venables, A. J. (1997). Preferential trading arrangements and industrial location.

Romer, P. M. (1994). The origins of endogenous growth.

Shaked, A., & Sutton, J. (1984). Natural oligopolies and international trade.

Solow, R. M. (1992). *Siena lectures on endogenous growth theory.*

Chapter 8
The Precursors

In this chapter we give a brief treatment of the precursors to the new explanations for international trade.

8.1 Availability

According this approach, due to Kravis (1956), international trade is explained by the fact that each country imports the goods that are not available at home. This unavailability may be due to lack of natural resources (oil, gold, etc.: this is *absolute* unavailability) or to the fact that the goods cannot be produced domestically, or could only be produced at prohibitive costs (for technological or other reasons): this is *relative* unavailability. On the other hand, each country exports the goods that are available at home.

Now, as regards the presence or absence of natural resources, this aspect could easily be fitted into the Heckscher-Ohlin model that, as we know, stresses the differences in relative factor endowments. We only have to add a factor *natural resources* (and, indeed, this has been done: see the discussion of Leontief's paradox in Sect. 4.6; for a general treatment of natural resources see Kemp & Long, 1984) and use the generalized version of the model (see Sect. 20.4)

The originality of this approach lies in its second aspect, that is, in the reasons put forward to explain international differences in relative availability. Essentially there are two reasons: *technical progress* and *product differentiation.*

As regards the first reason, Kravis observes that the stimulus to exports provided by *technological change* is not confined to the reduction in costs (in which case we remain in the context of the traditional theory) but also includes the advantages deriving from the possession of completely new products and of the most recent improvements of existing types of goods. In such cases the operation of the *demonstration effect* of Duesenberry (1949) creates an almost instantaneous demand

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_8,

abroad for the products of the innovating country and thus generates international trade.

As regards *product differentiation,* the idea of Kravis is to extend to international trade the results of the theory of monopolistic competition. Different countries produce similar commodities or, more exactly, commodities that are not substantially different from the point of view of their intended purpose (clothes, automobiles, watches, cameras, cigarettes, liqueurs, etc.). These commodities, however, due to different industrial designs, past excellence, advertising, real or imaginary secondary characteristics and so on and so forth, are *considered* different by consumers. This creates, on the one hand, a more or less limited monopolistic power of the single producing countries, and on the other a consumers' demand for foreign commodities that they believe different from similar domestic commodities, the result being to create international trade.

8.2 Technology Gaps

The advantage enjoyed by the country that introduces new goods, already considered by Kravis as one of the elements of his availability approach, is focused on by other authors (Hufbauer, 1966; Posner, 1961). As a consequence of research activity (especially of the Research & Development type) and entrepreneurship, new goods are produced and the innovating country enjoys a monopoly until the other countries learn to produce these goods: in the meantime they have to import them. Thus, international trade is created for the time necessary to imitate the new goods (*imitation lag*).

This lag has several components, that Posner (1961) classifies (from the point of view of the importing country) in the following categories:

(a) *Foreign reaction lag*. This is the time between the successful utilization of the innovation by entrepreneurs in the innovating country and the new goods becoming regarded, by some firms in the importing country, as a likely competitor for their products.
(b) *Domestic reaction lag*, which is the time required for all firms in the importing country to become aware of the competition from the new good.
(c) *Learning period*, which is the time required for the importing country's firms to learn to produce the new good, and actually produce and begin selling it on the domestic market.

According to Posner, to get the total net lag, one should subtract from the imitation lag a *demand lag*, that is, the time elapsing between the introduction of the new good in the innovating country and the appearance of a demand for it in other countries (some time elapses before the other countries' consumers come to know of the new good and acquire a taste for it). Imports of the new good will therefore take place only in the period of time resulting from the difference between the imitation lag and the demand lag. Contrary to this subtraction one may argue that it is consumers demand that stimulates the imitation by the firms of the importing country, so that the imitation lag starts from the moment in which demand appears; it is therefore incorrect to subtract the demand lag. This is however a unessential problem.

Once the imitation has been successfully performed, imports by the imitating country tend to cease, but as there is a *flow of innovations* through time, this aspect of international trade perpetuates itself. Besides, the imitation lag has not the same length in all countries, so that even if one or more countries successfully imitate new good, the innovating country will have an advantage in other countries, where the imitation lag is longer thanks to its greater experience in producing the good.

With reference to this, Posner defines the *dynamism* of a country as a function of its flow of innovations (that is, the number of new goods that it successfully introduces per unit of time) and of the speed with which it imitates foreign innovations.

When, in a two-country model, one is much more dynamic than the other, the less dynamic country will have to pay for its imports of new goods by exports of traditional goods at less and less favourable prices, and thus will not be able to carry out the massive investment (so as to modernise plants, etc.) required to increase its own dynamism. In other words, the less dynamic country remains trapped in its low level of dynamism.

On the contrary, when various countries have a very similar dynamism, international trade can stimulate a general growth process thanks to the fact that the innovations introduced in any country are rapidly imitated by the others. According to some authors, this is the phenomenon that occurred in the "golden age" of the European Economic Community (now European Union).

8.3 The Product Cycle

According to this theory, due to Hirsch (1967, 1975) and Vernon (1966, 1979), in the life cycle of a product it is possible to distinguish various phases: the *introduction* of the new good, its *maturation,* and its *standardization*, which together constitute the product cycle, with important effects on international trade.

The starting point is that equal *access* to scientific principles in all the advanced countries does *not mean equal probability of the application* of these principles in the production of new goods. According to Vernon (1966, p. 192), in fact, there are good reasons (for example, information costs) to believe that entrepreneurs' ability to get to know of new opportunities and to respond to them is a function of ease of communication with the market, which in turn depends on geographical proximity. As a consequence, firms generally introduce new products which are likely to satisfy the demand of the national market in which they sell. In the first phase, then, the production of the new good will be located in the country where the innovating firm operates, and the domestic market will be served.

When the new product has gained a hold upon the domestic market, the producer will begin to get into foreign markets, initially by exporting the good to them. In this phase of *maturation* the motives underlying the initial location disappear, and the firm will begin to examine the best way of serving foreign demand. On the one hand, the firm can continue to produce all the output at home and export the amount demanded abroad. On the other, the firm can licence foreign producers, or directly engage in producing the good in plants located in foreign countries where a demand exists; in this phase, the countries concerned will usually be advanced countries.

According to Vernon, in the case of new goods the licensing alternative is an inferior choice due to the inefficiencies and imperfections in the international market for technology (patents, licences, etc.). The firm having a monopolistic power thanks to the introduction of the new good will try to exploit this power also by way of price discrimination. As it is usually impossible to satisfy the conditions for optimal discrimination by using licences, to produce on one's own (either domestically or abroad) is a superior choice. To choose rationally between producing for exports at home or setting up producing subsidiaries abroad, the firm will compare the marginal cost of producing for exports at home, augmented by transport costs and tariffs (if any) levied by importing countries, with the unit cost of producing in a foreign subsidiary. A possible triggering event that induces the firm to set up a subsidiary abroad is the appearance in the foreign importing countries of local producers of the good. Another important element is the danger that the governments of importing countries, to protect their industries, may impose rigid restrictions such as quotas on the imports of the new product.

In the second phase, therefore, it is likely that the innovating firm will set up producing subsidiaries abroad, in developed countries. Thus, the export from the innovating country to these countries will dwindle away to zero, whilst it will continue to export to developing countries.

Finally, in the third phase of the cycle, we have advanced *standardization* of the good, hence the central, if not exclusive, importance of the cost of production in determining profitability. In this phase it may become advantageous to locate production units in less-developed countries because of the low cost of labour there. It may seem strange that this advantage makes itself felt also in the case of capital-intensive goods, but a less-developed country may offer competitive advantages as a location for the production of these goods, because the cost of capital may be less important than other factors (e.g., the marketing of the product, or such a low cost of labour to more than offset the greater capital intensity).

In the third phase, according to Vernon, in the country where the commodity originated, production dwindles whilst demand keeps increasing, so that this country gradually becomes an importer of the commodity, from other industrialised countries to begin with, then from less-developed countries.

The product cycle model also implicitly offers an explanation of the localization of production in different parts of the world and of the changes in this localization, hence it can also be considered as a precursor of the 'economic geography' models (see Sects. 16.3–16.5).

8.4 Income Effects

We examine here the theories which first focused on demand and income, among which are the theories of Linder (1961) and Barker (1977).

8.4.1 Linder's Theory

According to this theory, while the Heckscher-Ohlin theory is well suited to explain the pattern of trade in primary goods and, generally, in products intensive in natural resources, it is inadequate to explain the pattern of trade in manufactures. The alternative theory that he suggests starts from the concept of *potential trade* (potential exports and potential imports) of a country.

Potential exports are determined by domestic demand. More precisely, Linder's basic proposition (Linder, 1961, p. 87) is this: a necessary (albeit not a sufficient) condition for a product to be a potential export is that this product should be used as a consumption or an investment good in the home country. i.e. that a "representative" domestic demand for the product exists. Representative means that the product should be generally demanded. For example, although there is a demand for Ferraris, Rolls and Cadillacs in Saudi Arabia, this is not a representative demand and so it cannot turn luxury cars into potential export goods for Saudi Arabia.

Three main reasons are given by Linder to support his proposition:

1. It is unlikely for entrepreneurs to undertake the production of goods for which there is no domestic need;
2. Even if the existence for such a need abroad were perceived by entrepreneurs, they may be unable to conceive the product that will suit this need;
3. Even if this product were conceived, it is unlikely that it could be adapted to unfamiliar conditions without additional prohibitive costs.

This amounts to saying that, contrary to Heckscher-Ohlin theory, production functions are not internationally identical but that, for the entrepreneurs of a country, the production functions of commodities domestically demanded are the most advantageous. In other words, all this amounts to what businessmen call "the support of the domestic market".

As regards *potential imports*, it is domestic demand that determines which commodities may be imported (obviously this demand need not be representative). It follows that the range of potential exports coincides with, or is a subset of, the range of potential imports.

From this basic proposition it follows that the more similar the demand structures of two countries are, the more intense the potential trade between them will be. As an index of this similarity Linder takes the *similarity of per capita income levels*, since, in his opinion, there is a strong relationship between per-capita income and the types of commodities that demanded: for example, as per-capita income increases, higher-quality consumer goods will be demanded.

So far we have dealt with potential trade; we must now examine the forces that cause *actual* trade. Let us begin with an extreme case, in which two countries have identical per-capita income and so identical potential trade, for the potential exportables and importables are the same in both countries. Why then should there be (actual) trade between these countries? The answer is simple. When entrepreneurs broaden their horizons to the international market, they discover that they can expand into each other's country thanks to *product differentiation*. As Linder (1961, p. 102) remarks, "the almost unlimited scope for product differentiation—real or advertised—could, in combination with the seemingly unrestricted buyer idiosyncrasies, make possible flourishing trade in what is virtually the same commodity".

As regards countries with different per capita income, it is plausible to think that the same forces are at work, with the difference that the number of commodities for which the demands overlap will be lower and so actual trade will also be lower.

It goes without saying that growth induces increases in the per capita income of a country and so the structure of demand changes. As a consequence, the range of potential exports (and so of actual exports) is changing through time in a gradual and predictable way: "If Japan has been an importer of cars and exporter of bicycles, she might, within a decade, export cars and import bicycles" (Linder, 1961, p. 106). This is a prediction that hit the nail on the head.

Side by side with the forces that foster actual trade, there are forces that put a brake on it, for example distance (which comes into play in the form not only of transport costs, but also of other elements such as the imperfect knowledge of faraway markets), tariffs and other impediments to trade. Therefore, the braking forces will make actual trade—which, in their absence, would coincide with potential trade—smaller than the latter.

It is important to stress, in conclusion, that in Linder's theory it is the *similarity in demand* that generates trade (in similar but differentiated products): the greater the similarity the more trade there is, contrary to the traditional theory where one of the causes of trade is the *difference* in preferences (see Chap. 3), and the volume of trade increases as the economies become more dissimilar. For a re-examination of this idea see Economides (1984); for empirical tests of Linder's theory see the survey by Deardoff (1984), Eltis (1983), Kleinman and Kop (1984), and Hanink (1990).

8.4.2 Barker's Variety Hypothesis

Barker (1977) puts demand at the centre of the picture and acknowledges the contributions of Linder and other authors, but observes that these do not come to grips with the fact that trade grows *more than proportionally* to income. He therefore formulates the *variety hypothesis*, according to which consumers love variety, and so "as real incomes increase, purchasers are enabled to buy more varieties of a product; and since a greater number of these extra varieties is available from abroad rather than at home, the share of imports in demand tends to increase. Taking imports as a whole the quantity of imports in demand tends to increase more than proportionally with real income *per capita*" (Barker, 1977, p. 155).

The variety hypothesis starts from the theory of demand based on the *characteristics* of goods. As we cannot fully explain this theory here, we shall only recall its general principles, and refer the reader to its author (Lancaster, 1966, 1971).

According to this theory, the consumer actually desires the characteristics of the goods available, rather than the goods themselves. Characteristics are defined as "those objective properties that are relevant to choice by people" (Lancaster, 1971, p. 6). Thus the consumer purchases the good to obtain the characteristics embodied in them. To make an example, the consumer does not desire the commodity "automobile" as such, but desires a set of characteristics such as safety, fuel consumption, comfort, colour, acceleration, braking, steering, prestige, speed, etc., embodied in varying degrees in the various automobiles available on the market.

The first step in the consumer's choice is to find the *efficient* set of goods, that is, the set of goods which are not dominated by any other good. A good is dominated by another when, at the same price, it contains a lower amount of at least one characteristic and no higher amount of any characteristic.

The choice within the efficient set of goods will then be made on the basis of the budget constraint and of the utility function of the consumer; the arguments of this function are, as we said, the characteristics not the goods.

This said, Barker adds a series of assumptions (the goods are produced and can be purchased in several countries; there are transport costs, etc.) and demonstrates various propositions, amongst which (Barker, 1977, p. 160):

(a) There will be international trade in any tradeable good, since foreign goods will contain combinations of characteristics preferred by some buyers.
(b) The volume of trade in a set of goods having similar combinations of characteristics increases as per-capita real income increases, because the higher spending possibilities (relaxation of the budget constraint) enable consumers to buy more of the available goods.
(c) Up to the point of saturation, as per-capita real income increases, the purchase of imported goods increases by more than the purchase of analogous goods produced at home.

From these propositions, in particular from (c), Barker shows, by aggregating, the validity of the variety hypothesis formulated at the beginning. For empirical tests of the variety hypothesis see Barker (1977) and Vori (1984).

Table 8.1 Example of SITC classification

Digits	Items
8	Miscellaneous manufactured articles
...	...
85	Footwear
851	Footwear
...	...
851.01	Footwear with outer soles and uppers of rubber or artificial plastic material
851.02	Footwear with outer soles of leather or composition leather; footwear (other than footwear falling within heading 851.01) with outer soles of rubber or artificial plastic material

8.5 Intra-industry Trade or, the Traditional Theory Strikes Back

Let us recall from Sect. 7.1 that intra-industry trade is defined as the simultaneous export and import of products belonging to the same industry, which gives rise to an exchange of goods within, rather than between, industries. The empirical studies (see, e.g., Kol & Tharakan, 1989; Tharakan, 1983; these studies also describe the indexes used to measure intra-industry trade) show an increasing quantitative importance of this phenomenon.

Now, it is often alleged that the traditional theory cannot explain intra-industry trade which, on the contrary, is the normal outcome of the new trade theories. Hence an alleged superior explanatory power of the new theories. This claim is ill-founded, since—as shown both by precursors (Grubel & Lloyd, 1975) and by more recent writers (Davis, 1995)—intra-industry trade can be accounted for by the traditional theory.

To begin with, it should be observed that—apart from problems of physical homogeneity, which will be dealt with presently—internationally traded goods are usually classified in categories according to the Standard International Trade Classification (SITC) issued by United Nations (1975). This classification starts from a limited number of very broad basic classes, distinguished by *one* digit: for example, Sect. 1 is “Beverages and Tobacco”, Sect. 8 is “Miscellaneous Manufactured Articles”. Within each of these, more detailed categories are distinguished by *two* digits; each two-digit category is in turn disaggregated into various *three*-digit categories, and so on up to *five* digits. It should be noted that SITC, as internationally adopted, arrives at five digits; for further disaggregation, the individual countries are free to choose their own description and coverage. In practice the maximum disaggregation used arrives at seven digits (an example is 851.02.07 – Sand shoes, rubber-soled—see Grubel & Lloyd, 1975, pp. 19–20). It is clear that the higher the number of digits of an item the more precisely defined the set of similar goods included in that item. In Table 8.1 we give an example of the SITC classification, in which we have considered only a few disaggregations.

Obviously, if one considers the two-digit items only, the phenomenon of intra-industry trade is not a surprise, for we are dealing with classes so broad as to include heterogeneous goods.

Intra-industry trade would then be a spurious phenomenon, due to statistical aggregation. But since intra-industry trade is also observed in higher-digit items (Vona, 1990), even going as far as the seven-digit ones, it cannot be neglected from the theoretical point of view. Grubel and Lloyd were among the first systematically to examine the problem (Grubel, 1967; Grubel & Lloyd, 1975). From the theoretical point of view we must distinguish between the case of *identical* goods and the case of non-identical (though belonging to the same industry) goods.

8.5.1 Perfectly Homogeneous Goods

In the case of identical goods the traditional theory can supply various explanations, the oldest being that of transport costs (see Sect. 6.3). A second explanation is given by what Grubel and Lloyd call *periodic trade*, which can be due to:

(i) Seasonal factors. For example, country 1 and country 2 both produce the same summer fruit, but they lie at the antipodes, so that when it is summer in country 1 this country will export summer fruit to country 2 where it is winter, and vice versa. Thus we shall observe intra-industry trade on a yearly basis. This can be easily fitted in the traditional theory, by assuming transformation curves that periodically change their position.
(ii) Varying conditions of demand. For example, it is normal that neighbouring countries exchange electrical power with one another to meet demand peaks in one or another country. This can also be fitted into the traditional theory, by assuming demand curves that periodically change their position.

A third explanation refers to the import and export of goods after mere storage and wholesaling (*entrepôt trade)* or after simple manipulations (such as packaging, bottling, cleaning, sorting, etc.) which leave the goods essentially unchanged (*re-export trade*) . Even in the case of re-export trade the manipulations are usually not sufficient to warrant the reclassification of the goods in a different SITC class, so that intra-industry trade is observed.

A fourth explanation refers to the effects of government intervention. Let us assume, for example, that in a three-country world countries 1 and 2 join a free trade area and country 2 levies higher duties against country 3 than country 1 does. It may then be advantageous for country 3, in order to export a good to country 2, first to export the good to country 1 and so pay a lower tariff, and then re-export it to country 2 as coming from country 1, thus paying no further duties. Country 1 will then appear as an importer and exporter of the same commodity.

To conclude: intra-industry trade in perfectly homogeneous goods can be quite well accommodated by the traditional theory. But what about differentiated products?

8.5.2 Differentiated Products

As soon as we drop the traditional assumption of product homogeneity, the presence of intra-industry trade in products which are sufficiently similar to belong to the same SITC category but have some degree of differentiation, becomes a necessary consequence. As a matter of fact, *all the theories treated in* Sects. 8.1–8.4, *which consider product differentiation an essential element of trade, can be used to explain intra-industry trade.*

To examine intra-industry trade in differentiated products, it is convenient to follow a classification introduced by Grubel and Lloyd (1975) , based on similarity of input requirements and substitutability in use.

The *first group* contains commodities with similar input requirements but low substitutability in use, such as bars and sheets of iron.

The *second group* includes commodities with low similarity in input requirements but high substitutability in use, such as wood and plastic chairs.

The *third group* contains commodities with similarity in input requirements and high substitutability in use, such as cars with similar characteristics, but manufactured by different producers.

It goes without saying that the group of commodities with low similarity in input requirements and low substitutability in use does not come into consideration, for these commodities belong to different SITC classes and no intra-industry trade will be observed.

Intra-industry trade in commodities belonging to the first group can be explained by the traditional theory, for their low substitutability in use makes them *different* commodities from the point of view of demand. Intra-industry trade is simply a phenomenon due to statistical aggregation.

Intra-industry trade in commodities belonging to the second group can also be explained by the traditional theory, for the dissimilarity in their input requirements means that they have to be considered as *different* commodities from viewpoint of production: intra-industry trade is, again, a phenomenon due to statistical aggregation.

We are left with the third group in which we may further distinguish two cases.

The first one is when the commodities are so similar (as regards both input requirements and substitutability in use) that they can be considered as homogeneous for all practical purposes, and we are back in the situation examined in the previous section.

The second case is the relevant one: the commodities, though very similar, have to be considered different from the economic point of view, because of technological differences in production and/or because consumers believe them to be different (for reasons of brand, design, advertising, etc.) even if they are perfectly substitutable in use and with identical inputs (toothpastes or medicines with the identical chemical composition are an example). At this point the market form becomes essential.

If the differentiated goods are produced under constant returns to scale and the market remains perfectly competitive, then the traditional theory can again be invoked. In fact, a commodity which differs, however slightly, from another commodity from the point of view of production and/or demand can be formally treated as a different commodity (for example, two commodities that have identical factor proportions but even a slight Hicks-neutral productivity difference have to be classified as different commodities, though belonging to the same industry). And, as long as markets are perfectly competitive, we can apply the traditional theory in its generalization to n commodities (see Sects. 2.4, 3.7, and 20.4).

It should be emphasized that by "traditional theory" we do not mean solely the Heckscher-Ohlin model, but—as clearly stated in Sect. 7.1—all the theories examined in Part II, hence also the Ricardian theory. As we know (see Sect. 1.2), the Ricardian model emphasizes technical differences while the Heckscher-Ohlin model emphasizes factor endowments; both are firmly rooted in the perfectly competitive framework with constant returns to scale.

Now, commodity differentiation from the production point of view can arise from different factor proportions (Heckscher-Ohlin) and/or from different technologies (Ricardo), as we pointed out above. This is, in fact, the approach followed by Davis (1995) who, after defining "perfectly-intraindustry goods" as those goods that for all factor price ratios are produced under identical factor intensity (hence they are Heckscher-Ohlin identical), assumes that two such goods have a small Hicks-neutral productivity difference across the two trading countries (hence they are Ricardo different).

It is then no surprise that intra-industry trade can take place, which can coexist with inter-industry trade in a model (called by Davis a Ricardo-Heckscher-Ohlin model) in which there also are perfectly homogeneous goods (i.e., goods with absolutely identical production functions).

This result reinforces what we have repeatedly noted in this section on intra-industry trade, namely that this phenomenon can quite well be accommodated in the context of the traditional theory.

To conclude: increasing returns to scale (which are typically associated with imperfect competition and hence with the new theories of international trade, see Sect. 7.1) are not necessary to account for intra-industry trade. A conclusion that does not detract from the merits of the new trade theories, but puts the entire question into proper perspective: the traditional theory cannot be attacked (and the new theories cannot be praised) just on the basis of the inability or ability to explain intra-industry trade. The focus must shift on whether we are dealing with perfectly or imperfectly competitive markets, which is a factual rather than a theoretical question. It is comforting to know that international trade theory (both old and new) gives us the tools for coping with all market forms.

References

Barker, J. (1977). International trade and economic growth: An alternative to the neoclassical approach.

Davis, D. R. (1995). Intra-industry trade: A Heckscher-Ohlin-Ricardo approach.

Deardoff, A. V. (1984). Testing trade theories and predicting trade flows.

Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior*.

Economides, N. (1984). Do increases in preference similarity (across countries) induce increases in trade? An affirmative example.

Eltis, C. M. (1983). An alternative interpretation and empirical test of the Linder hypothesis.

Grubel, H. G. (1967). Intra-industry specialization and the pattern of trade.

Grubel, H. G., & Lloyd, P. J. (1975). *Intra-industry trade: The theory and measurement of international trade in differentiated products*.

Hanink, D. M. (1990). Linder, again.

Hirsch, S. (1967). *Location of industry and international competitiveness*.

Hirsch, S. (1975). The product cycle model of international trade: A multi-country cross section analysis.

Hufbauer, G. C. (1966). *Synthetic materials and the theory of international trade*.

Kemp, M. C., & Long, N. V. (1984). The role of natural resources in trade models.

Kleinman, E., & Kop, Y. (1984). Who trades with whom – The income pattern of international trade.

Kol, J., & Tharakan, P. K. M. (1989). *Intra-industry trade: Theory, evidence, and extensions*.

Kravis, I. B. (1956). "Availability" and other influences on the commodity composition of trade.

Lancaster, K. (1966). A new approach to consumer theory.

Lancaster, K. (1971). *Consumer demand: A new approach.*

Linder, S. B. (1961). *An essay on trade and transformation*.

Posner, M. V. (1961). International trade and technical change.

Tharakan, P. K. M. (Ed.). (1983). *Intra-industry trade: Empirical and methodological aspects.*

United Nations, (and subsequent revisions). (1975). *Standard international trade classification*. New York: UN.

Vernon, R. (1966). International investment and international trade in the product cycle.

Vernon, R. (1979). The product cycle hypothesis in a new international environment.

Vona, S. (1990). Intra-industry trade: A statistical artifact or a real phenomenon?

Vori, S. (1984). L'ipotesi della varietà nelle nuove teorie del commercio internazionale: il modello e la sua verifica empirica.

Chapter 9
The Models

In this chapter we examine the main models of the new theories of international trade, as classified in Table 7.1. These theories also introduce new arguments in the debate on free trade versus protectionism (e.g. Baldwin, 1992, Brander and Spencer, 1984, 1985, Flam and Helpman, 1987, Gabel and Neven, 1988, Grossman and Richardson, 1985, Haberler, 1990, Pomfret, 1992, Venables, 1985, 1987), but these arguments are better studied in the context of trade policy (see Chap. 10, Sect. 10.8).

9.1 Neo-Heckscher-Ohlin Theories

This designation derives from the fact that in these theories (also called neo factor proportions) the departure from the traditional theory is kept to a minimum (in particular, the assumption of perfect competition is maintained), and the conclusion is obtained that intra-industry trade conforms (with due modifications) to the traditional statement of the Heckscher-Ohlin theorem.

The model that we examine is due to Falvey (1981), who starts from the idea that each industry does no longer produce a single homogeneous output, but instead can produce a range of products differentiated by quality (each quality is produced by many competing firms). Thus, according to the terminology introduced in Sect. 7.1, we are in the case of *vertical* differentiation. The second point of departure from the traditional theory is the nature of capital: the capital stock is no longer homogeneous, but consists of capital equipment *specific* to each industry. Because of its specificity the capital stock is immobile among industries, but of course freely mobile in the production of the various qualities within each industry. The labour force is—like in the traditional theory—homogeneous and hence mobile also among industries.

For simplicity, the analysis is limited to a single industry (hence we are in a partial equilibrium context). This industry owns a certain amount of specific capital

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_9,

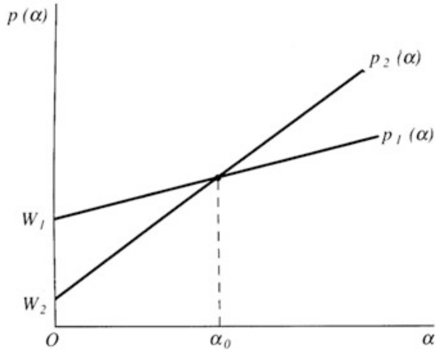


Fig. 9.1 Vertical differentiation and international trade

(whose rate of return, R, adjusts so as to maintain the full employment of the capital stock) and can employ any amount of labour at the current wage rate W. The industry under consideration produces a continuum of different qualities of the product (the assumption of the production of a continuum of qualities is made for mathematical convenience), with a constant-returns-to-scale technology. The problem now arises of defining the quality. For this purpose Falvey introduces a numerical index α such that greater values of α correspond to higher qualities, and assumes that the production of higher-quality goods requires a correspondingly higher quantity of capital per unit of labour. It is now possible to define the measurement units in such a way that the production of a good of quality α requires the input of one unit of labour and α units of capital. Given the assumption of perfect competition, for any quality the price equals the unit cost of production, namely

$$\begin{aligned} p_1(\alpha) &= W_1 + \alpha R_1, \\ p_2(\alpha) &= W_2 + \alpha R_2, \end{aligned} \tag{9.1}$$

where the subscripts 1 and 2 refer as usual to the two countries, whose technology is assumed identical (again in agreement with the Heckscher-Ohlin framework).

Without loss of generality we can assume that $W_1 > W_2$. It follows that international trade requires $R_1 < R_2$: in the opposite case, in fact, we see from Eqs. (9.1) that country 2 could produce any quality of the commodity at a cost (and hence price) which is lower than in country 1, so that there would be no scope for international trade. Assuming then $R_1 < R_2$, it follows that a certain subset of qualities will be produced in country 1 at a lower cost than in country 2, and vice versa for the other subset. In order to identify these two subsets, let us use a diagram (Fig. 9.1), where we have drawn the two linear price-cost relationships given in Eqs. (9.1). Let us note that $R_i (i = 1, 2)$ is the slope of line p_i, hence the p_2 line is steeper than p_1, since $R_2 > R_1$. We see from the diagram that prices are equal in the two countries in correspondence to the "marginal" quality α_0, while country 2 has a comparative cost advantage over country 1 for lower-quality products ($\alpha < \alpha_0$); conversely, country 1 has a comparative cost advantage for higher-quality products ($\alpha > \alpha_0$).

If we now make the plausible assumption that in both countries there is a demand for both lower-quality and higher-quality products, it follows that, in the typical situation of free trade with no transport costs, there will be international trade in the products of the industry considered: country 1 will export higher-quality products to (and import lower-quality products from) country 2. Since we are dealing with products of the same industry, what has taken place is indeed *intra-industrial* trade.

What is more, such a trade follows the lines of the Heckscher-Ohlin theorem, as can be easily shown. Given the assumptions made on the returns to the factors of production, we have $R_1/W_1 < R_2/W_2$, which means that country 1 is capital abundant relative to country 2 according to the price definition of relative factor abundance (see Sect. 4.2). Now, since higher values of α mean both higher qualities and higher values of the capital/labour ratio, we observe that country 1, the capital-abundant country, exports capital-intensive products (conversely country 2, the labour-abundant country, exports labour-intensive products).

In a subsequent model (Falvey & Kierzkowski, 1987) two industrial sectors have been introduced, one of the type treated above and the other traditional, namely producing a single homogeneous commodity. This model is able simultaneously to generate inter-industrial and intra-industrial trade along the lines of Heckscher-Ohlin theorem, in a context of perfect competition and very similar to the traditional one.

It is finally worthwhile emphasizing the fact, mentioned at the beginning of this section, that a plausible model of intra-industry trade has been produced with a minimum of departure from the traditional theory: apart from product differentiation, it has not been necessary to introduce economies of scale or monopolistic competition as other models do. This does *not* mean that these features are unimportant or uninteresting, it simply stresses that the phenomenon of intra-industry trade can be made to fit into the traditional theory, with results similar to those of the Heckscher-Ohlin model.

9.2 Monopolistic Competition and International Trade

In this section we shall present the foundations of a new theory of international trade as developed in the seminal papers by Krugman (1979, 1980; see also Gabszewicz et al., 1981, Grossman, 1992, Harrigan, 1994, 1996, Helpman, 1990, Helpman and Krugman, 1989, Markusen et al., 1995). The new theory features monopolistic competition and posits that the market structure, regardless of comparative advantage, gives rise to international trade. The section continues with the discussion of a number of developments such as a Heckscher-Ohlin-Krugman synthesis model, the home market effects, the gravity equation and the heterogeneity in firms performance.

9.2.1 The Krugman Model

9.2.1.1 Introduction

The two fundamental elements of this theory are the economies of scale internal to the firm and the demand for differentiated products. As we shall see below these two simple elements give rise to international trade. Trade is of the intra-indutry type and takes place even in the absence of comparative advantage.

9.2.1.2 The Demand Side

Two alternative theoretical foundations for the demand for varieties were proposed in the 1970s. These, gave a rigorous foundation to the treatment of demand under monopolistic competition and made it possible to extend the analytical apparatus of monopolistic competition to international trade theory. We review them both briefly.

The first is due to Dixit and Stiglitz (1977) and Spence (1976). They argue that behind the demand for differentiated goods there is simply the desirability of variety as such, which is implicit in the traditional indifference curves that are *convex* to the origin. If a consumer is indifferent between two goods—namely if the combinations (1,0) and (0,1) of these goods lie on the same indifference curve—then an intermediate combination like $(1/2, 1/2)$ is preferred to both extremes. This is because the intermediate combination lies on the straight line segment which joins the two extreme combinations, hence this combination will lie on a higher indifference curve. This can easily be formalized introducing a utility function such that the utility index increases, *ceteris paribus*, as the number of varieties consumed increases. Therefore each consumer demands all the existing varieties of a differentiated good.

It is convenient to present here the utility function used in Dixit-Stiglitz, then adopted in Krugman's works and vastly utilized in the international trade literature. The utility function takes the following functional form:

$$u = \left(\sum_{k=1}^{n} (D_k)^{\alpha} \right)^{\frac{1}{\alpha}}, \qquad 0 < \alpha < 1, \tag{9.2}$$

where D_k is the quantity consumed of the variety k and n is the number of varieties available to the consumer. In the appendix we expound the various properties of this utility function; our purpose here is just to show how it gives rise to the demand for variety. We begin by observing that each variety is equally liked, since each contributes to total utility in the same way. Therefore, in equilibrium (and if production costs are the same), each variety will have the same price. Since prices are identical, each variety will be demanded in the same amount as any other. Imagine then a consumer whose total consumption D is to be spread equally over a number of varieties. He will consume a quantity $D_k = D/n$ of each variety. Now replace $D_k = D/n$ into the utility function (9.2) so as to obtain $n^{\frac{1-\alpha}{\alpha}} D$. It is clear that utility increases with total consumption (D) and with the number of varieties (n). It is also clear that the consumer desires to spread a given total consumption (D) over the maximum possible number of varieties since utility increases in n for any given D. Thus, the consumer does indeed demand all existing varieties, and if new varieties become available he will demand them too.

The preferences *à la S-D-S* (Spence-Dixit-Stiglitz) have been used by Krugman in several works in which he builds a theory of international trade in differentiated goods based on monopolistic competition. This has been called *neo-Chamberlinian* monopolistic competition, because it is nearer to the original vision of Chamberlin himself (see Kierzkowski, 1985).

A second line of analysis of the demand side has been taken by Lancaster (1980), who observes that for all the varieties of a differentiated product to be demanded at the aggregate level, it is not necessary that such a demand also exists at the individual level: it is, in fact, sufficient that each consumer (or group of consumers) has different tastes and so demands a different variety of the product. He starts from an intuition of Hotelling (hence the name of *neo-Hotelling* monopolistic competition given to Lancaster's approach: see Kierzkowski (1985, p. 17)) and applies his own goods-characteristics approach to demand, arriving at a model of monopolistic competition that he extends to international trade. The Lancaster approach (already mentioned in Sect. 8.4.2) starts from the assumption that the consumer does not want the commodities as such, but the characteristics embodied in the commodities. It follows that the demand for the commodities is an indirect or derived demand that depends on the preferences with respect to the characteristics and on the technical properties that determine how the characteristics are embodied in the different commodities. The different individual reactions of different consumers with respect to the same commodity are then seen to be the result of *different individual preferences* with respect to the characteristics (which are perceived in the same way by all consumers) embodied in that commodity rather than the result of different individual perceptions of the characteristics of that commodity.

Lancaster's demand theory is more sophisticated and flexible than the *S-D-S* preferences, but to explain international trade (and intra-industry trade in particular) the reason why at the aggregate level all the varieties of a horizontally differentiated good are demanded does not make much difference. As Krugman observes (1990, p. 75), both approaches lead to a monopolistically competitive equilibrium in which several differentiated goods are produced by different firms all of which have monopolistic power but none of which earns monopolistic profits. Thus we shall follow Krugman (1980) and adopt the S-D-S preferences.

We now turn to the supply side of the model.

9.2.1.3 Technology and Production

As mentioned above, the main objective of the Krugman model is to show how the market structure generates international trade in the absence of comparative advantage. We therefore eliminate from the model any possible source of comparative advantage by assuming that the technology of production is identical between countries and so are factor proportions. Since we rule out any role for endowments it is convenient to assume that there is only one factor of production, namely labour. Further, since there is no comparative advantage, it is unnecessary to have two goods in the model. After all, the model wants to explain intra-industry trade and one good therefore suffices. The model is simplified to the utmost so that we can focus on its two essential elements: the desire for variety and internal economies of scale.

The technology of production is assumed to be identical for all firms and to be characterized by the presence of fixed and variable inputs, both in terms of labour. The production function may be conveniently written in terms of labour input, l, per q units of output: $l = F + cq$, where F is the fixed labour input and cq is the variable labour input. With w denoting wage, $w(F + cq)$ is the total cost. It is immediate that the average cost, $AC = w(F/q + c)$, is declining with the output of the firm while the marginal cost, cw, is constant. The presence of fixed cost makes it optimal for the firm to produce all its output in one plant. Indeed, if it produced in two or more plants it would incur a fixed cost Fw for each plant, which would increase the overall average cost.

Given the desire for variety, each firm will find it optimal to produce a variety different from that of every other firm. The reason is simple: by producing a different variety the firm will be the sole producer of that variety (a monopolist in the market for that variety) whereas if it produced an existing variety it would find itself in direct competition (duopoly) with the other firm producing that variety. Since monopoly profit is larger than duopoly profit, each firm will choose to differentiate its product. This differentiation is profitable since consumers like variety per se and are always happy to consume any existing variety and any new variety introduced in the market. Firms maximize profits by applying the general rule of profit maximization: *marginal revenue = marginal cost*. Since each producer is a monopolist in the market for its variety the profit maximization rule yields a price larger than the marginal cost. The price/marginal-cost ratio is called the mark-up and reflects the market power of the producer. Let $\mu > 1$ denote the mark up, the profit maximization condition is:

$$p^* = \mu cw. \tag{9.3}$$

The technology is identical across firms and all firms face identical demand because consumers like all varieties with the same intensity. Therefore, the profit-maximizing price is the same for all firms, this is why p^* has no index referring to any particular variety.

Free entry does not let any positive profit to remain. If profits were positive new firms (producing new varieties) would enter the market until profits were driven to zero. The free entry assumption therefore gives zero profit as an equilibrium

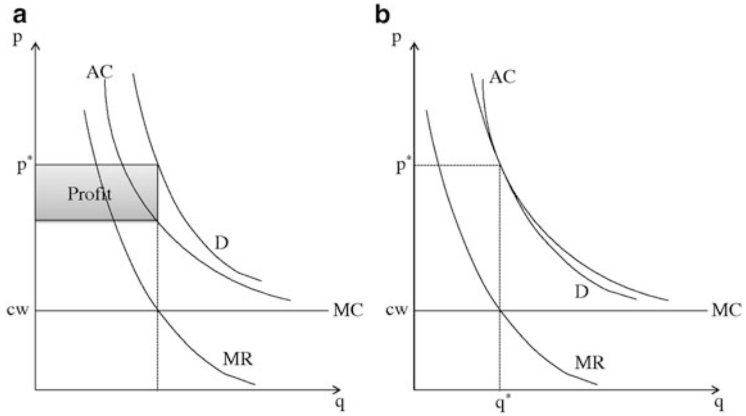


Fig. 9.2 Profit maximization

condition. Let $\pi = pq - w(F + cq)$ be the profit of a firm. Replacing p^* into the expression for profits the zero profit condition is $\mu cq - (F + cq) = 0$, which, solved for q, gives the equilibrium output of each firm:

$$q^* = \frac{F}{c\,(\mu - 1)} \tag{9.4}$$

The profit-maximization condition is represented in Fig. 9.2a, which depicts the situation of any firm. The curves labelled AC, MC, D, and MR represent average cost, marginal cost, demand for the variety in question, and marginal revenue respectively. The producer maximizes profits by choosing a price such that the corresponding marginal revenue is equal to the marginal cost.[1] Profits are represented by the gray area. But positive profits induce the entry of new firms, so Fig. 9.2a does not represent the equilibrium in the market. As new firms enter the market, total expenditure on the industry will be spread over a larger number of varieties. Consequently, the demand curve for any variety will shift to the left until it is tangent to the average cost curve. When tangency occurs, profit is zero because the average cost is equal to the price. The situation when the profit-maximization condition and the zero-profit condition are both satisfied is depicted in Fig. 9.2b. The demand-reducing effect for any existing firm due to the entry of new firms is called the *market crowding* effect. We shall refer to it in a number of occasions below in this chapter. In our discussion of the zero profit condition, we have referred to the profit-erosion (or demand-reducing) effect of entry to help the intuition. It should be clear, however, that in the model there is no entry or exit dynamics and the market settles immediately in the equilibrium depicted in Fig. 9.2b.

[1]In this model it is immaterial whether the producer maximizes profit by setting the price or the quantity.

Firms make zero profits, but they make an operating profit. The latter is defined as the difference between revenue and variable costs. More precisely, firms make an operating profit per unit sold given by the unit price minus the marginal cost: $p^* - wc$. The operating profit is therefore $(p^* - wc)\, q^*$. Since profits are zero, the operating profit is exactly equal to the fixed costs Fw, as can be easily verified by replacing p^* and q^* into the expression for operating profit.

Autarky Equilibrium

We first consider the autarky equilibrium and then a free-trade equilibrium between two identical countries. We do not use any country index in the discussion of the autarky equilibrium because countries are identical. There are four equilibrium conditions in each of the two economies. The first two are profit maximization and zero profit, which have been discussed above. The third, is the demand-equal-supply condition in the market for any variety (there are as many such conditions as there are varieties but these conditions are all identical and therefore reduce to just one condition). The fourth, is the demand-equal-supply condition in the labour market. The third and fourth condition are identical in autarky and we therefore need consider only one of them. Consider the labour market equilibrium. A firm's demand for labour is $F + cq^*$. Let n be the number of firms in the market, itself an endogenous variable. Total demand for labour is $n\,(F + cq^*)$. Equilibrium in the labour market requires that $L = n\,(F + cq^*)$, where L is labour endowment. Solving this equation for n and using expression (9.4), we obtain the equilibrium number of varieties (firms) in the economy, expression (9.4), we have the equilibrium number of varieties (firms) in the economy:

$$n^* = \frac{L}{F}\frac{\mu - 1}{\mu} \tag{9.5}$$

It is instructive at this point to discuss the role played by μ and F in the results obtained in expressions (9.3)–(9.5). We recall that μ represents the market power of producers which, clearly, is increasing with the rigidity of the demand curve. Demand rigidity depends on the importance that consumers attach to variety per se. If the taste for variety per se is very high, the demand for any variety is very rigid (consumers are reluctant to substitute one variety for another) and the mark-up is therefore very large. Conversely, if the taste for variety is low, the demand curve is very elastic and the mark-up very small. With this in mind, it is clear why a stronger taste for variety (high μ) makes p^* higher, reduces the size of firms q^*, and increases the number of varieties, n^*. The fixed cost F plays no role in the determination of p^* since neither marginal revenue nor marginal cost depend on F. However, F plays a role in the determination of q^* and n^*. To understand this, consider the effect of an increase in F. With higher F, firms could only survive if they made higher operating profits, but this requires an increase in output, since the mark-up is constant (see expression (9.4)). Furthermore, since total demand for the good in question, wL, is constant, larger firm size is only possible if the number of firms declines (see expression (9.5)).

Free Trade Equilibrium

Consider now a world composed of two countries indexed by $i = 1, 2$. Countries have identical technology, identical preferences and endowments equal to L_1 and L_2. The profit-maximization condition gives $p_i^* = \mu c w_i$ for any i and the zero-profit condition gives $q_1^* = q_2^* = q^*$. Equilibrium conditions for the two labour markets give the number of varieties produced in each country:

$$n_1^* = \frac{L_1}{F}\frac{\mu - 1}{\mu} \tag{9.6}$$

$$n_2^* = \frac{L_2}{F}\frac{\mu - 1}{\mu} \tag{9.7}$$

Consumers demand all varieties, both domestic and foreign. Therefore, each variety is not only sold domestically but also exported. There is no comparative advantage, and yet there is international trade. International trade is the result of the market structure: the desirability of variety and single-plant production being the key elements. All trade is intra-industry, i.e., the exchange of different varieties of the same good.

In moving from autarky to free trade, the real wage in terms of any variety, w_i/p_i^*, remains unchanged; therefore there is no gain from trade resulting from changes in prices. Nevertheless, welfare increases because consumers can spread their expenditure over a larger number of varieties. The number of varieties available to consumers increases from n^* in autarky to $\left(n_1^* + n_2^*\right)$ in free trade.

Lastly, we need to establish the trade flows. We begin by noting that given the twin structure of technology and preferences, countries will have the same wage $w_1 = w_2 = w$. Thus, consumers in country i will spend a fraction $n_j^*/\left(n_1^* + n_2^*\right)$ of their expenditure on foreign varieties (with $j \neq i$). The value of country 1's imports, IMP_1, is therefore:

$$IMP_1 = \left[n_2^*/\left(n_1^* + n_2^*\right)\right] wL_1 = wL_1L_2/\left(L_1 + L_2\right) \tag{9.8}$$

and equals the value of country 2's imports, which is:

$$IMP_2 = \left[n_1^*/\left(n_1^* + n_2^*\right)\right] wL_2 = wL_1L_2/\left(L_1 + L_2\right) \tag{9.9}$$

which confirms that the trade balance is in equilibrium with wage equalization.[2] We need not worry about the equilibrium condition in the market for any variety, as this is identical to the trade balance equilibrium condition. Expressions (9.8) and (9.9) also show that the more similar countries are in size, the larger the volume of trade. For any given size of the world labour force, $L_1 + L_2$, the volume of trade is larger the nearer L_1 is to L_2. The reason is that the more alike countries are, the more evenly-distributed the number of firms between them, and these countries will therefore have a lot more to exchange with one another. While the volume of trade is determined, there is indeterminacy as to which variety is produced and exported by each country. This indeterminacy of the direction of trade is, however, irrelevant, since nothing hinges on who produces what.

The model is extremely simple and extremely powerful. Its simplicity comes at the cost of missing some important aspects of reality, such as the sensitivity of the mark-up to the intensification of competition or the presence of multiproduct firms. We shall address some of these aspects below, but here we anticipate that taking account of these elements does not change the fundamental result that market structure alone can generate international trade between identical countries. To fully appreciate the power of this model, we should recall that it solved one of the major puzzles in the field of international trade at the end of the 1970s, i.e., the large volume of intra-industry trade among very similar countries and the lack of a theory to explain that phenomenon.

The monopolistic competition model offers an entirely new explanation for international trade, but it would be a mistake to see it as incompatible with the Heckscher-Ohlin model. The two models can be combined in a single model where comparative advantage determines the specialization and direction of trade in goods and the market structure explains intra-industry trade. This is what we shall discuss in the next section.

9.2.2 A Simple Synthesis Model

The stratagem to simplify the analysis, already used in Sect. 4.3.2 to examine the factor price equalization set, is to start from an integrated world economy, which will be subsequently divided into two countries. So at the beginning we have a closed economic system (the world), producing two commodities: a differentiated commodity, say a manufactured good (henceforth called good A) and a homogeneous commodity, say food (henceforth called good B). In industry A there are increasing returns to scale and monopolistic competition, while in industry

[2] Here we have taken an innocuous shortcut. Wage equalization results from the equilibrium condition for the trade balance. But this would take us into some unnecessary technicalities. To simplify the exposition, we have "guessed" that wages equalize and have verified that the trade balance equilibrium condition is satisfied under wage equalization.

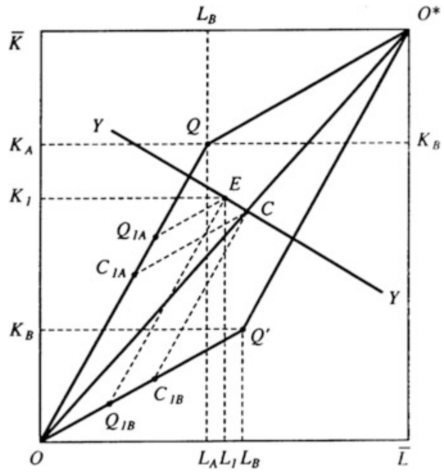


Fig. 9.3 Monopolistic competition and international trade

B there are constant returns to scale and perfect competition. Both industries use homogeneous capital and labour as factors of production; both factors are freely mobile between industries and fully employed. Given the prevailing set of factor prices and goods prices, there will be a certain factor allocation between the two sectors. Let us consider Fig. 9.3 (adapted from Krugman (1990, p. 76)), which can be considered as an extension to two countries of the box diagram explained in Sect. 3.1. In the diagram, the length of the sides of the box represents the total quantities of labour and capital existing in the economy ($\overline{L}$ and $\overline{K}$). Point Q (the end-point of vector OQ) gives the allocation of resources to sector A in the integrated economy. Thus sector A employs OK_A of capital and OL_A of labour. Similarly point Q', end-point of the vector OQ', represents the allocation of factors to sector B, which employs OK_B of capital and OL_B of labour. We see that good A is more capital-intensive than B, but this is not important.

Since both factors are fully employed, we have $OK_A + OK_B = O\overline{K}$, hence $OK_B = K_A\overline{K} = O^*K_B$. Similarly, $OL_A + OL_B = O\overline{L}$, hence $OL_B = L_A\overline{L} = O^*L_B$. Thus by construction we have $O^*Q = OQ'$, i.e. vector O^*Q has the same length and slope as vector OQ', and vector O^*Q' has the same length and slope as vector OQ. Hence OQO^*Q' is a parallelogram.

The next step in the analysis is to imagine the world divided into two countries, say country 1 and country 2, which are identical to the integrated economy as regards tastes (consumers also have the same structure of demand), technology, and market forms. The prices (of goods and factors) are also the same as in the integrated economy. The only difference is in factor endowments (note, incidentally, the analogy with the standard Heckscher-Ohlin assumptions). If we measure country 1's endowment starting from O and country's 2 endowment starting from O^*, the subdivision of $\overline{K}$ and $\overline{L}$ between the two countries can be represented by a point in the box. Let us suppose that such a subdivision is given by point E, so that country 1 has an endowment of OK_1 of capital and OL_1 of labour: the rest is the endowment of country 2.

Since the prices of goods and factors are assumed to be same as in the integrated economy, country 1 will produce the two goods with the same techniques used in the integrated economy, namely it will produce good A with the capital/labour ratio given by the slope of OQ and good B with the capital/labour ratio given by the slope of OQ'. With reference to the box OK_1EL_1, the allocation of K_1 and L_1 to the production of A and B can thus be determined (similarly to what we did in the integrated economy) by drawing the parallelogram $OQ_{1A}EQ_{1B}$, where the side EQ_{1A} is parallel to OQ_{1B}, and EQ_{1B} is parallel to OQ_{1A}.

Let us now draw through E the straight line YY having the same slope as the factor-price ratio p_L/p_K. We recall that profits disappear not only under perfect competition but also under monopolistic competition, so that the price coincides with the average cost of production, namely the cost of factors. Since national income coincides with the value of total factor rewards, which in turn coincides with national product, all the points along YY represent a value of the national income of country 1 equal to that existing at point E. If we then look at YY from the point of view of origin O^*, we can conclude that along this line, the value of country 2's national income (which can of course differ from that of country 1) is also constant. Therefore OC/OO^* measures country 1's share of world income (output).

Given the initial assumption of identical structures of demand in the two countries, it follows that both countries demand the goods—and hence consume the factor services embodied in them (see Sect. 3.1 for the transition from the space of goods to the space of factors)—in the same proportion. The consumption point will thus be along the diagonal OO^*. Since all national income is consumed, the consumption point is C. To determine the composition (in terms of the two goods) of the consumption basket represented by C we draw the usual parallelogram, obtaining points C_{1A} and C_{1B}. Since C_{1A} is nearer than Q_{1A} to the origin O, it contains a smaller quantity of good A. It follows that there will be *net* exports of A (we shall presently see why *net*), since consumption is smaller than output. Similarly we can see that point C_{1B} represents a consumption of good B greater than domestic output (point Q_{1B}): thus country 1 imports B.

We have reached the conclusion that country 1, the relatively capital-abundant country, exports the relatively capital-intensive good A, and exports the relatively labour-intensive good B. These results are perfectly in line with the conventional Heckscher-Ohlin theorem. But there is more to it than that: while international trade in good B will be of the conventional inter-industry type, trade in good A will be of the *intra-industry type*. We have in fact just seen that the exports of A are *net* exports: this means that country 1 will simultaneously export and import goods belonging to industry A, the exports being however greater than the imports. To show this, we must recall that—as a consequence of economies of scale in the production of each variety of commodity A—no country can produce the entire range of varieties of this commodity, but only part of it. Therefore, even if both countries produce manufactured goods, each will produce different varieties; which country produces which varieties cannot be determined, but this is not important for our analysis. In fact—independently of the hypothesis made on preferences (see Sect. 9.2)—consumers in each country are assumed to demand all varieties. Thus, to satisfy domestic demand, country 1 will import from country 2 the varieties that it does not produce, and export to country 2 the varieties that it produces, to meet country 2's domestic demand. There is, consequently, intra-industry trade (which in the aggregate, as we have seen above, gives rise to net exports of A from country 1), that will coexist with inter-industry trade.

This result is independent of the kind of preferences (S-D-S or Lancaster) assumed: these, however, come back into the picture when we go on to examine the gains from international trade. These gains are the availability of a greater number of varieties and an increased scale of production of the single varieties, giving rise to a lower unit cost of production thanks to scale economies. As Krugman shows (1990, p. 79), only the first type of benefit is possible with S-D-S preferences, while both types are possible with Lancaster preferences.

The welfare effects of international trade in the synthesis model come from comparative advantage and from the expansion of the number of varieties. Therefore, trade is certainly beneficial to both countries. There is, however, a new result with respect to the Heckscher-Ohlin model, concerning income distribution. We have seen in Sect. 4.3.1 that any trade liberalization in the Heckscher-Ohlin model hurts the relatively scarce factor and benefits the relatively abundant factor, in the sense that it reduces the real wage of the former and increases that of the latter. This notwithstanding, in the synthesis model it is possible that the welfare gain deriving from the increased number of varieties available to consumers may outweigh the loss for the relatively scarce factor coming from the loss of purchasing power in terms of any variety. Thus, as shown in Krugman (1981), it is possible that both factors gain from trade. The more alike countries are in terms of relative factor endowments and the stronger the taste for variety, the more likely this is to happen.

9.2.3 Monopolistic Competition and Welfare Effects of Trade Opening

In monopolistic competition there is a new source of welfare gain from international trade. This source is represented by the expansion of the number of varieties available to consumers when passing from autarky to free trade. Since consumers like variety per se, such expansion brings about an increase in welfare.

This source was the only one in the Krugman model (Sect. 9.2.1). The welfare effects of international trade in the synthesis model (Sect. 9.2.2) instead come from comparative advantage and from the expansion of the number of varieties. As we have seen when studying the Heckscher-Ohlin model, the presence of comparative advantage suffices for trade to be beneficial to all countries. In the synthesis model welfare is a fortiori beneficial to all countries since the welfare gain coming from the expansion of the number of varieties is added to the welfare gain obtained from comparative advantage. There is, however, a new result in the synthesis model. While in the Heckscher-Ohlin model the welfare of the relatively abundant factor increases and that of the relatively scarce factor declines, in the synthesis model it is possible that all factors gain from trade. As shown in Krugman (1981), this occurs when the welfare gain deriving from the increased number of varieties available to consumers outweighs the loss for the relatively scarce factor. The more alike countries are in terms of relative factor endowments and the stronger the taste for variety, the more likely this is to happen.

9.2.4 The Home Market Effect

In the presence of trade costs or other form of market segmentation, the size of expenditure in a country relative to the other country has an impact on wages and specialization. This impact is known as the "home market effect" and is the subject of this section. More precisely, the home market effect refers to either of these phenomena: (1) a positive relationship between a country's relative wage and relative size of expenditure; (2) a more than proportional relationship between the relative size of output of a good in a country and the relative size of that country's expenditure.[3]

To simplify matters, we shall assume that all international trade costs may be modeled as international transport costs. We shall adopt the iceberg transport costs already introduced in Sect. 6.3. Thus, for each unit of a variety sent from country i to country j, only a fraction $\tau \in (0, 1)$ of it arrives at its destination, the remaining $(1 - \tau)$ being lost in transit.

9.2.4.1 Demand and Wages

We consider again the model in Sect. 9.2.1 to which we add trade costs. The price charged by a firm to domestic and foreign consumers cannot be the same, since the latter includes trade costs. The mark-up is the same in all markets but the firm takes account of the fact that the marginal cost of producing for the foreign market includes the fraction of the variety lost in transit. The marginal cost of producing for the domestic market is still cw_i. The marginal cost of producing for the foreign market is instead $\frac{1}{\tau}cw_i$, since in order to sell one unit of output in the foreign

[3]The terminology "home market effect" appears for the first time in Helpman and Krugman (1985, chap. 10), where it refers to the second phenomenon mentioned in the text. Later it became clear that the two phenomena are just two different manifestations of the same economic mechanism. See Head and Mayer (2004) for a critical and comprehensive appraisal on the literature referring to either of these two phenomena.

market the firm has to produce $1/\tau$ units. Let p_{ii}^* and p_{ij}^* be, respectively, the profit-maximizing price in i and j of a variety produced in i. These prices are:

$$p_{ii}^* = \mu c w_i \tag{9.10}$$

$$p_{ij}^* = \frac{1}{\tau}\mu c w_i \tag{9.11}$$

Consider two identical countries. Since countries are identical, equilibrium will be such that all endogenous variables will be identical between countries, notably, $w_1 = w_2$, and $p_{11} = p_{22}$. Consider a symmetric demand shock by which demand increases in country 1 and decreases by the same magnitude in country 2, thus leaving world demand unchanged. Since there is only one good, the change in demand can only originate from a change in country size. With populations being initially $L_1 = L_2$, the shock is of the type $\Delta L_1 = -\Delta L_2 > 0$. Such a symmetric demand shock results in an excess demand for any variety produced in country 1 and in an excess supply for any variety produced in country 2. The reason is due to the presence of transport costs which make $p_{12} > p_{11}$. Since $p_{12} > p_{11}$, foreign demand for any variety produced in 1 is smaller than domestic demand. Therefore, the increase in demand originating from country 1 dominates the fall in demand originating from country 2 and overall demand for any variety produced in 1 increases.[4] Obviously, if we had assumed $\Delta L_1 = -\Delta L_2 < 0$ then we would have an excess supply for any domestic variety. In sum, at home, the home market shock dominates on the foreign market shock. We refer to this dominance as to the "home market dominance" in the demand shocks. The magnitude of the home market dominance (the excess demand or excess supply) depends on trade costs and on the taste for variety. The higher the trade costs, the bigger the excess demand or supply. In the extreme case of autarky, the excess demand or supply reaches its maximum value, since there is no fall in foreign demand. Second, the weaker the taste for variety, the greater the excess demand or supply. To understand this, recall that a weak taste for variety means that varieties are highly substitutable for one another. Therefore, any given price difference between a domestic and a foreign variety will induce a larger reduction in demand for the latter and the excess demand or supply generated by the shock will therefore be larger.

For clarity of exposition let us continue with the case where $\Delta L_1 = -\Delta L_2 > 0$. The excess demand resulting from the expenditure shock will have to be absorbed by a change in output and/or a change in prices. In the model we are using, there will

---

[4] **As an example, assume that** $L_1 = L_2 = 10$ **and** that the other model parameters are such that in the initial equilibrium $w_1 = w_2 = 1$ and that the expenditure on any domestic variety emanating from country 1's residents is 10 % of income while the expenditure on any domestic variety emanating from country 2's residents is 8 % of income. Initial national income is 10 in both countries. Now consider a shock $\Delta L_1 = -\Delta L_2 > 1$. The excess demand for any country 1's variety is $0.1 - 0.08 = 0.02 > 0$.

be changes in output and prices. Let us see why. The shock $\Delta L_1 = -\Delta L_2 > 0$ is in itself a shock to the labour force, causing an expansion in total industry output in country 1 and a reduction in total industry output in country 2. Since output per firm is constant, changes in total industry output occur via entry of firms in country 1 and exit of firms in country 2. The change in relative industry outputs induced by the change in relative labour forces is perfectly proportional to the latter. Indeed, from Eqs. (9.6) and (9.7) we obtain $n_1/n_2 = L_1/L_2$ which shows the perfect proportionality. So far we have established that an increase in L_1/L_2 causes an excess demand for varieties produced in country 1 and, via the labour market, an increase in the relative number of varieties produced in country 1. Interestingly, the increase in the relative number of varieties produced in country 1 is not sufficient to clear the excess demand for varieties produced in country 1 (likewise for the reduction of varieties in country 2 and the excess supply there). To understand this, consider the effect that the entry of a new firm has on other firms' profits. It is convenient to begin with the case of a country in autarky. Imagine that this country experiences an increase in demand of a given magnitude. If there were no entry by new firms, the excess demand would be distributed evenly overall firms, thus giving rise to positive profits. This induces the entry of new firms. The entry of new firms will subtract demand from existing firms by exactly the amount that brings them back to the initial level of demand and to zero profit. This is quite obvious: in autarky, the expenditure on every variety is wL/n^*, which shows that an equiproportional change in L and n^* leaves the expenditure per variety unchanged. This means that any increase in demand induced by ΔL is entirely absorbed by the corresponding increase in the number of varieties induced by ΔL itself. The fact that a new firm subtracts demand from other firms is called the market crowding effect, as we have already mentioned above. In autarky, the market crowding effect is perfect in the sense that the entry of a new firm reduces total demand for the aggregate of all varieties one for one. Let us now return to the situation of two countries and assume that country 1 experiences an excess demand of the same magnitude as the autarkic country in the previous example. Now the entry of new firms in country 1 induced by ΔL will subtract demand not only from domestic firms but also from foreign firms: it will therefore absorb only a fraction of the excess demand for domestic varieties. Likewise, the excess supply for country 2 varieties will not be cleared by the exit of firms induced by ΔL_2, precisely because part of the expenditure freed by the disappearance of those firms is reallocated to all firms, not only to those in country 2. Overall, the entry of firms in country 1 and the exit of firms in country 2 shifts demand towards the aggregate of varieties produced in country 1, thus generating further excess demand. Therefore, after proportional entry and exit there is still a residual excess demand for varieties produced in country 1.

The residual excess demand can only be absorbed by an increase in the relative price of varieties produced in country 1 (further entry is not possible since all labour is already employed). Since prices and wage are in constant proportion the increase in the relative price of domestic varieties brings about an increase in the relative wage of country 1, w_1/w_2. This result may be summarized as follows.

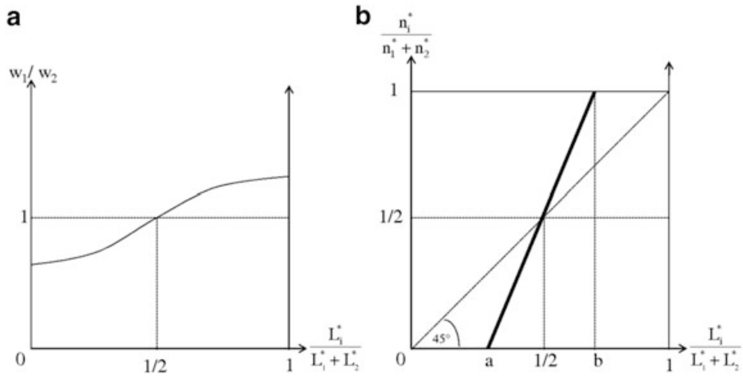


Fig. 9.4 The home market effect: (**a**) demand and wages (**b**) demand and specialization

Proposition 9.1. *In the model described in Sect. 9.2.1 and in the presence of international trade cost, the wage is higher in the country where demand is greater.*

Interestingly, the wage difference is due neither to different technologies nor to different factor proportions; it is only due to different market sizes. Figure 9.4a shows the relationship between relative wages and relative size of demand resulting from this model.

Note that the excess demand generated by the demand shock is simply due to the presence of trade costs and is not particularly related to monopolistic competition. It would occur, albeit with different intensity, even if the goods were homogenous. The fact that the excess demand is not entirely absorbed by proportional entry is instead due to product differentiation. This distinction has been used in some empirical studies to which we shall return in Sect. 9.4.

9.2.4.2 Demand and Specialization

In Sect. 9.2.4.1, the residual excess demand had to be absorbed by changes in wages, since the entry of additional firms was not possible due to the resource constraint. But this need not be the case in general. In this section we modify the model so that the excess demand will be absorbed uniquely by the entry of firms. This extension is based on Helpman and Krugman (1985, sect. 10.4), although the economic mechanism and results had already been presented in Krugman (1980). The only difference with respect to the model of Sect. 9.2 is that there are two goods, A and M. Consumers spend a fraction γ of total expenditure on good M and the remaining $(1-\gamma)$ on good A. Good A is produced using a constant-return-to-scale technology under perfect competition. Specifically, the production function is $A = L$, which means that one unit of labour input produces one unit of output. Since there is perfect competition in A, the price of A will be equal to the marginal cost, i.e., $p_{Ai} = w_i$. Good M is differentiated and the market structure is monopolistic competition as described in Sect. 9.2. International trade is free in good A while it is subject to iceberg costs in good M. Since there is free trade in A, the price of A must be the same in both countries, i.e., $p_{A1} = p_{A2} \equiv p_A$, which implies $w_1 = w_2 \equiv w$. The profit-maximizing price for any variety of good M is given by $p_{ii}^* = \mu c w$ and $p_{ij}^* = \frac{1}{\tau}\mu c w$ which differ from expressions (9.10) and (9.11) only in that the wage is the same in the two countries. We recall that the first subscript refers to the country where the good is produced and the second subscript refers to the country where the good is sold.

Consider now a symmetric expenditure shock. Since there are two goods, the expenditure shock may have two sources: (1) a symmetric shock to preferences such that γ changes in opposite directions in the two countries or (2) a symmetric shock to country size as in Sect. 9.2.4.1. The source of the shock is irrelevant and we shall choose the second. So let us consider a shock $\Delta L_1 = -\Delta L_2 > 0$. As in Sect. 9.2.4.1, the home market dominance makes that the symmetric expenditure shock gives rise to an excess demand for varieties produced in country 1. For the same reasons as in Sect. 9.2.4.1 this excess demand is not entirely absorbed by a proportional change in the number of varieties. Differently from the model in Sect. 9.2.4.1, in this section countries can specialize. Thus, the excess demand not absorbed by a proportional increase in the number of varieties will be absorbed by further entry of firms in country 1 and further exit in 2. Thus, industry M will expand more than proportionally with respect to changes in the relative labour force of country 1. This is possible because in country 1, the labour needed for a more than proportional expansion of the M industry may be taken from industry A. Likewise, the labour released by industry M in country 2 will be employed in industry A. This result may be stated as follows:

Proposition 9.2. *In the model described in this Sect. 9.2.4.2 there is a more than proportional relationship between the relative size of demand and the relative size of output in good M. Thus, the country with a relatively larger demand for M will specialize in the production of M.*

Interestingly, the source of international specialization is not comparative advantage but market size. Figure 9.4b shows the relationship between the relative size of output and the relative size of demand resulting from this model. Note that the relative size of output is $n_i p_{ii}^* q^* / \left(n_1 p_{11}^* q^* + n_2 p_{22}^* q^*\right)$ which equals $n_i / (n_1 + n_2)$ by virtue of expression (9.4) and since $p_{ii}^* = \mu c w = p_{jj}^*$. The relative size of demand is $L_i / (L_1 + L_2)$. The thick line in Fig. 9.4b represents the more than proportional relationship between the share of output and the share of demand. Naturally, the more than proportional relationship holds until one country is completely specialized. Such a situation is represented in Fig. 9.4b at point a (where country i has completely specialized in A) and at point b (where country j has completely specialized in A).

9.2.4.3 Robustness of the HME

A number of works have studied the robustness of the HME to reasonable model modifications. Head, Mayer, and Ries (2002) consider alternative forms of market structure. They find that the HME is pervasive and may emerge even in the presence of an oligopoly with homogeneous goods, as long as markets are segmented either by trade costs or by the demand structure. Brülhart and Trionfetti (2009) find that demand influences specialization even in the absence of trade costs as long as it gives rise to some form of market segmentation. Davis (1998), using a model similar to that in Sect. 9.2.4.2, notes that the relationship between the share of output and the share of demand in the differentiated good can only be proportional if there are prohibitive trade costs in A. Crozet and Trionfetti (2008) show that with trade costs in all goods (including good A) and product differentiation by country of origin, the more than proportional relationship becomes non-linear, being weaker for similar countries and stronger for countries of very different size. Behrens, Lamorgese, Ottaviano, and Tabuchi (2009) develop a multicountry and asymmetric trade cost model and show that the relationship between share of output and share of demand is not necessarily more than proportional in this setting. Head and Mayer (2004), in their critical and comprehensive appraisal of this literature, show that the HME can disappear when the intersectoral mobility of labour is less than perfect.

Other papers have studied the home market effect predominantly from the empirical point of view and we shall review them below in Sect. 9.4

9.2.5 Adding Some Realism to the Monopolistic Competition Model

In this section, we discuss two aspects of the monopolistic competition model described in Sect. 9.2 that seem particularly unsatisfactory. The first is that the mark-up is constant and the second is that firms only produce in one country.

9.2.5.1 Variable Mark-Up

Constant mark-up represents a convenient simplification when the objective is to show that the market structure generates international trade between identical countries, but it sacrifices too much realism when the objective is to study how firms adjust to trade opening. After all, the mark-up reflects the market power and it seems reasonable to think that in moving from autarky to free trade the market power of each firm declines because of fiercer competition. We shall refer to the equations of the model in Sect. 9.2.1 in discussing this matter. To be precise, however, we should change the demand structure (typically not that of S-D-S preferences) and specify some additional aspects of firms' behaviour. These modifications would bring us into a tedious taxonomy of cases without adding substantial matter to the understanding of the economic mechanisms. We therefore stay with the equations already obtained above since they approximate the equations obtained from alternative specifications of the model. Let μ^A and μ^T denote the mark-up in autarky and in free trade with $\mu^T < \mu^A$. First, we see from (9.3) and (9.4) that a decline in the mark-up reduces the price and expands the output of each firm. This is often referred to as the *pro-competitive effect* of international trade which results in lower mark-ups and lower average costs. Second, some firms will succumb to fiercer competition. Replacing μ^A in (9.5) and μ^T in either (9.6) or (9.7) shows that $n^* > n_i^*$, which means that the increased competition pushes some firms out of the market.[5] This is sometimes referred to as the *firm exit* effect. Furthermore, under some conditions on the demand functions (which we omit in order to avoid unnecessary technicalities), it is possible to show that the decline in mark-up is such that $n_1^* + n_2^* > n^*$. This means that the number of varieties available to consumers is larger in free trade than in autarky although each country will produce a smaller number of varieties. It should be clear that the pro-competitive effect and the firm exit effect are not specifically related to the presence of product differentiation. As a matter of fact, these effects are typical of oligopoly models with or without product differentiation; see Markusen (1981) for a deeper treatment. In monopolistic competition models these effects appear when the perceived elasticity of demand is not constant; see, e.g., Krugman (1979) and Ottaviano, Tabuchi, and Thisse (2002).

9.2.5.2 Multiproduct Firms

In the monopolistic competition model used above, firms are single-product (they only produce one variety) and national (they only produce in one country). This result seems at odds with reality. Firms often produce more than one variety and typically in different countries. This aspect of reality can easily be taken into account in the monopolistic competitive model if we assume, as is reasonable, that there are costs of trading between countries. The presence of trade costs entails that each national market is partially protected from foreign competition. Then, as argued in Baldwin and Ottaviano (2001), a firm would find it optimal to set up another production plant abroad, producing a variety different from that produced at home. Thanks to the market segmentation caused by trade costs, this strategy allows the firm to gain market share abroad without generating too much competition with its own home-produced variety. This does not affect the existence of intra-industry trade, however. Indeed, the variety produced abroad by the national firm is sold abroad and domestically like any other variety.

In conclusion, it is clear that taking into account multiproduct multinationals and variable mark-ups would make the model more realistic but would leave unchanged

[5]Firms are identical, so the model does not indicate which firms will succumb. This is an issue that we shall discuss in Sect. 9.2.7 below.

the fundamental result that the market structure gives rise to international trade between identical countries.

9.2.6 The Gravity Equation

Data on international trade flows show a remarkably stable empirical regularity known as the gravity equation. The gravity equation posits that the trade flow between two countries is increasing in the 'mass' of goods the exporter has to offer, increasing in the 'mass' of demand emanating from the importing country, and decreasing in trade costs. In the early specifications of the gravity equation, the mass of supply and demand were represented by GDP. This relationship was named gravity because of the analogy with the gravity between planets (stars, etc.), increasing in the planets' masses (GDPs) and decreasing in distance (trade costs). The relationship posited by the gravity equation has been confirmed over several decades of empirical studies. However, for long it lacked a neat theoretical foundation. The monopolistic competition model offers a very direct foundation for it which can be grasped by inspection of expressions (9.8) and (9.9) above. These expressions show that exports between countries are, ceteris paribus, increasing with the size of the exporter, L_1, with the size of the importer, L_2, and with the similarity in the size of the countries. Noting that the size of the labour force is, essentially, the GDP of the country, one can formulate a relationship between exports from one country to another as depending positively on the exporting country GDP, on the importing country's GDP and on the similarity of GDPs. Furthermore, declining trade flows with increasing distance are easily derived by enriching the model with iceberg trade costs (see Sect. 23.2.3.1 for a formal derivation). The theoretical foundation provided by the monopolistic competition model gave a lot more meaning to the gravity equation and stimulated further research which continues to date. The theoretical and empirical advancements since the pioneering study by Anderson (1979) are thoroughly discussed in Head and Mayer (2013).

9.2.7 Heterogeneous Firms

Firms are a major actor in international trade. Exporting is undertaken by firms in response to demand emanating by foreign firms and/or foreign consumers. Yet firms remain in the backstage in Ricardian and Heckscher-Ohlin theories of international trade. This is due to the perfect-competition and representative-firm assumptions adopted in these models. These assumptions have made it possible to focus on country/industry characteristics (comparative advantage) as determinants of international trade and specialization. Imperfect competition, and especially the Krugman model, has brought to light the importance of market structure. In this model firms play a more active role and their decisions are crucial in determining international trade. Yet, while the assumption of perfect competition is dropped, that of the representative firm is maintained. This assumption does not allow to examine how firms and the industry as a whole reorganize themselves when economies open up to international trade. Consider, for instance, the specialization induced by international trade in comparative advantage models: the industry with the comparative advantage expands and the other one contracts. These changes in the size of industries probably do not affect all firms in the same way, but comparative advantage models are silent on this matter since the assumption of identical atomistic firms rules out any scrutiny of what happens to them. Consider the Krugman model: in this model all firms export, and yet even a cursory inspection of data shows that only a very small fraction of firms are actually engaged in international trade. This diversified reality about firms and their response to trade opening found a theoretical collocation in the work of Melitz (2003).

Melitz developed a general equilibrium model where firms differ in their productivity levels. We can grasp the crucial mechanisms of this model by applying some modifications to the monopolistic competition model studied in Sect. 9.2.1. In the present context we assume the presence of iceberg trade costs. Let $\phi \equiv 1/c$ be the marginal productivity of labour (recall that c are the units of labour needed to produce one unit of output). We have already seen in Sect. 9.2.1 that in monopolistic competition with S-D-S demand, the profit-maximizing price for any firm is a multiple μ of the marginal cost:

$$p = \frac{\mu}{\phi} w \tag{9.12}$$

Unlike the model in Sect. 9.2.1, here the marginal productivity ϕ varies across firms. To simplify matters, assume that firms *draw* their marginal productivity ϕ from a probability distribution which has support $(0, \infty)$.[6] Once ϕ is drawn, the firm can compute its domestic and foreign profit. The profit of a firm is given by revenue minus variable costs minus fixed costs and the operating profit is given by revenue minus variable costs. Firms face a fixed costs of production, wF, and a fixed costs of exporting wF_x. Therefore domestic and foreign profits for a firm in country i are:

$$\pi_{ii} = \pi_{ii}^{o}\left(\overset{+}{\phi}\right) - wF \tag{9.13}$$

$$\pi_{ij} = \pi_{ij}^{o}\left(\overset{+}{\phi}\right) - wF_x \tag{9.14}$$

where the notation $\pi_{ii}^{o}(\phi)$ and $\pi_{ij}^{o}(\phi)$ indicates that domestic and foreign operating profits depend on productivity, ϕ. As indicated by the algebraic sign above ϕ, a rise in productivity increases operating (and total) profits. Since domestic and foreign

[6]The draw is not free. Firms have to pay a fixed cost equal to F_e units of labour in order to draw the marginal productivity. This stylized mechanism may reflect, for instance, the cost incurred in acquiring the relevant information about the expected costs and benefits of operating a business.

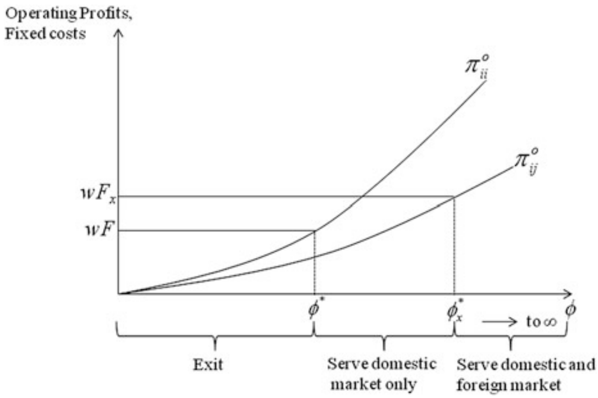


Fig. 9.5 Zero profit and zero foreign profit conditions

profits are independent, the firm takes two separate decisions after drawing ϕ: to stay or not to stay, to export or not to export. The firm will stay in the market if it draws a high enough ϕ for the profits on the domestic market to be non-negative. The firm will decide to export if it draws a high enough ϕ for profits in the foreign market to be non-negative.[7] These decisions give two separate zero-profit conditions:

$$\pi^o_{ii}(\phi) = wF \quad \text{Zero Domestic Profit Condition} \tag{9.15}$$

$$\pi^o_{ij}(\phi) = wF_x \quad \text{Zero Foreign Profit Condition} \tag{9.16}$$

Equation (9.15) is the zero-profit condition on the domestic market. The value of ϕ determined by this equation is in fact the smallest value of ϕ such that profit in the domestic market is non-negative. We refer to the value of ϕ determined by Eq. (9.15) as the zero profit productivity cutoff and denote it ϕ^*. A firm will stay in the market if it draws a value of ϕ larger or equal to ϕ^* and will exit otherwise. Equation (9.16) is the zero-profit condition on the foreign market. The value of ϕ determined by this equation is the smallest value of ϕ such that profit in the foreign market is non-negative. We refer to the value of ϕ determined by Eq. (9.16) as the zero exporting profit productivity cutoff and denote it ϕ^*_x. A firm will export if it draws a value of ϕ larger or equal to ϕ^*_x and will not export otherwise. Firms' decisions with respect to staying and exporting are depicted in Fig. 9.5.

ϕ is plotted on the abscissa and operating profits and fixed costs on the ordinate. The curves emanating from the origin represent domestic and foreign operating profits as functions of productivity, ϕ. The horizontal lines represent fixed

[7]Unlike the model studied in Sect. 9.2.1, the free entry condition and the zero profit condition are disjoint. The free entry condition requires that the expected profit from running a business should equal the entry cost wF_e. The expected profit from running the business depends on the expected value of ϕ, which in turn depends on the probability distribution of ϕ.

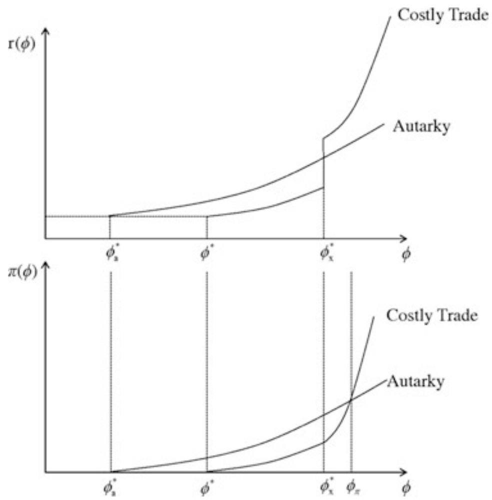


Fig. 9.6 From autarky to costly trade

production costs and fixed exporting costs. The intersection between each operating profit line and its corresponding fixed cost line gives the cutoff values ϕ_x^* and ϕ^*. In this figure, $\phi_x^* > \phi^*$. This ranking is consistent with the fact that while some firms do not export, exporting firms typically also supply the domestic market. It can be seen from Fig. 9.5 that this ranking of cutoff values obtains thanks to appropriate restrictions on the relative size of fixed costs, which we assume to be satisfied. Unlike the monopolistic competition model studied in Sect. 9.2.2, when firms are heterogeneous not all of them make zero profit. Only the firms that have drawn a productivity level equal to ϕ^* will make zero profit. Their situation is the same as that depicted in panel (b) of Fig. 9.2. We refer to these firms as the *cutoff firms*. Firms drawing a higher productivity level will make positive profits. Quite intuitively, individual market share and profits increase with productivity.

In autarky, no firm exports and only the zero-profit productivity cut-off is defined. Let ϕ_a^* be the zero-profit productivity cut-off in autarky. Free trade is characterized by $\tau = 1$ and $F_x = 0$. In free trade all firms export. Costly trade is characterized by $\tau \leq 1$ and $F_x > 0$. In costly trade, in general, not all firms export, since not all of them can afford to pay F_x and still make non-negative profits on the foreign market. The presence of fixed export costs (and not the presence of iceberg costs) endogenously generates the partition of all firms into exporting and non-exporting firms. If there were no fixed export costs, the model would simplify to the model developed in Sect. 9.2 with only minor differences. The effect of moving from autarky to costly trade is conveniently represented in Fig. 9.6 drawn from Melitz (2003). On the abscissa of the upper panel we measure revenues, r, and on the abscissa of the lower panel we measure profit, π. The notation $r(\phi)$ and $\pi(\phi)$ recalls that revenues and profits positively depend on the productivity level ϕ drawn by a firm. Firms drawing a productivity level $\phi < \phi_a^*$ will exit immediately without engaging in any production. Firms drawing a productivity level $\phi > \phi_a^*$ will stay in the market and produce. Sales and profits increase smoothly with productivity, thus firms having drawn a higher ϕ will sell more and make higher profits, as shown by the lines labelled "Autarky".

In moving from autarky to costly trade the *market crowding* effect pushes some firms out of the market. Clearly, it will be the least efficient firms that succumb. Therefore, the zero-profit productivity cut-off moves to $\phi^* > \phi_a^*$. Now, firms with productivity ϕ between ϕ_a^* and ϕ^* exit. But that is not all. Firms with productivity ϕ between ϕ^* and ϕ_x^* will find it profitable to produce only for the domestic market, whereas firms with productivity $\phi > \phi_x^*$ will sell in the domestic and foreign market. Furthermore, trade causes a reallocation of market share. Comparing the lines "Autarky" and "Costly Trade" in the upper panel we see that firms with productivity ϕ between ϕ_a^* and ϕ_x^* have lost market share, while firms with productivity $\phi > \phi_x^*$ have gained market share. Profits are reallocated too. In the lower panel we see that firms with productivity $\phi < \phi_\pi$ lose part or all of their profits, while firms with productivity $\phi > \phi_\pi$ expand their profits. Interestingly, firms with productivity ϕ between ϕ_x^* and ϕ_π gain market share but lose profits. Any further decline in trade costs will cause a shift of ϕ^* further to the right and a shift of ϕ_x^* further to the left, thereby causing the exit of more firms and increasing the number of firms able to export.

Clearly there are consequences on average productivity. Since the zero-profit productivity cut-off moves to the right, the average productivity of the industry increases with any decline in trade costs. The number of varieties available to consumers may increase or decrease; it tends to decline because of the exit of some domestic firms but to increase because of the increase in the number of foreign exporters. It can be shown that welfare increases when iceberg trade costs decline.

This model gives a richer picture of what happens inside an industry when a country moves from autarky to costly trade. The heterogenous firm model may be combined with a Heckscher-Ohlin model as proposed in Bernard, Redding, and Schott (2007). The resulting synthesis model exhibits inter-industry and intra-industry trade in a way analogous to the synthesis model studied in Sect. 9.2.2. But there are a number of additional results. One of the most interesting new results is that under some conditions on the probability distribution, the ex-ante probability of exporting is higher in the industry of comparative advantage. Another result is that, ceteris paribus, the zero-profit productivity cut-off, ϕ^*, is higher in the industry of comparative advantage, implying that the industry with the comparative advantage will have a higher average productivity, ceteris paribus, than the other industry. This, in turn, adds a sort of endogenously-determined Ricardian comparative advantage to an otherwise identical-technology Heckscher-Ohlin structure.

Many other models extensions have been developed in the literature after Melitz's work. We shall study some of them in Sect. 17.5 with particular attentions to the implications for the labour market. For a comprehensive discussion of theoretical developments see Redding (2011).

9.3 Oligopoly and International Trade

9.3.1 Introduction

In the previous sections, we have considered models based on market forms that might be called "structurally competitive", namely where the number of firms is sufficiently high for no firm influencing, with its own decisions, the decisions of the other firms. On the contrary, we consider here models based on oligopolistic markets, where the problems of strategic interdependence among a limited number of firms become essential.

As we know from microeconomics (see, e.g., Friedman, 1977, Varian, 1992), there does not exist a general model of oligopoly. Oligopolistic firms can act in collusion, tacit or explicit (as in cartels) or in a non-cooperative manner. When they do not cooperate, the result of their interaction depends on several factors: the decision variable of the firm (price or quantity), the nature of the firms' *conjectural variations* (i.e., of the assumptions that each firm makes as regards the other firms' reactions to its price or quantity changes), the specification of the product, the nature of the market (i.e., whether it is segmented or not), etc. Thus it not possible to give a *general* analysis of the effect of oligopoly on international trade. It is however possible—through the study of specific cases—to obtain interesting results especially as regards intra-industrial trade. In what follows we have set up our treatment according to the product type, in agreement with the classification in Table 7.1.

9.3.2 Homogeneous Commodities

International trade in homogeneous goods in a context of oligopolistic markets was examined by Markusen (1981), who assumed integrated markets, and by Brander (1981) and Brander and Krugman (1983), who assumed segmented markets. Here we shall follow the latter approach, because it can explain intra-industry trade in homogeneous goods. For a synthesis of the two approaches see Venables (1990).

Intra-industry trade in homogeneous goods, that we have already treated in Sect. 8.5, is explained by Brander as the result of the interaction among oligopolistic firms in different countries. Let us consider the simplest case of duopoly: one firm in country 1 and one in country 2, both producing the same homogeneous commodity. The decision variable is assumed to be the quantity, so that each firm has to decide how much of its output to sell at home and how much abroad (the whole output is produced domestically). Transport costs are modelled according to the iceberg assumption (see Sect. 6.3), are borne by the producers, and are assumed to be symmetrical—that is to say, the unit transport cost of the output of firm 1 to (the market in) country 2 is equal to the unit transport cost of firm 2's output to country 1. To make the model as simple as possible the technology is assumed internationally identical with identical production costs (marginal costs are constant); the demand

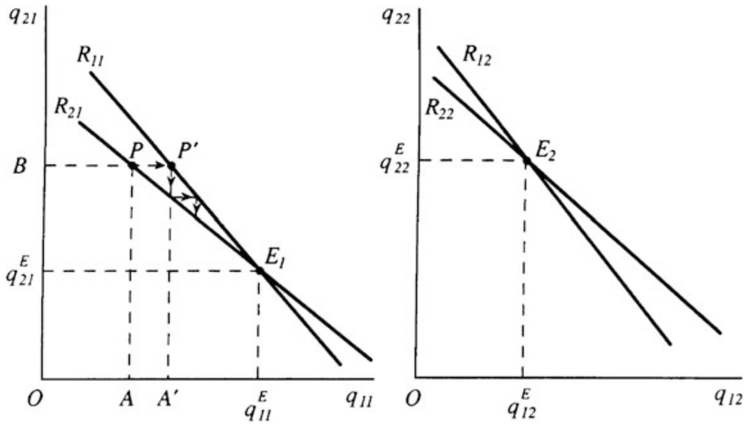


Fig. 9.7 Homogeneous duopoly and reciprocal dumping

functions are also internationally identical. The two markets are assumed to be segmented, so that firm i can sell at a different price at home and abroad. Naturally, since the product is homogeneous, in a given market the price will be identical for both the domestically produced and the imported good.

The strategic interaction between the firms is modelled following the Cournot hypothesis, according to which each firm maximises profit choosing its decision variable (the quantity) on the assumption that the quantity supplied by the other firm remains the same. The only (but important) difference between the conventional Cournot duopoly and the case under examination is that here each firm acts in two different markets, in each of which it employs a Cournot strategy as regards the other firm's supply to the same market. To be precise, if we denote by q_{ij} $(i, j = 1, 2)$ the quantity offered by firm i on market j, we have that firm 1 chooses q_{11} and q_{12} so as to maximise profit, assuming that q_{21} and q_{22} remain the same; similarly firm 2 will choose q_{21} and q_{22} so as to maximize profit, taking q_{11}, q_{12} as constant. In calculating its profit each firm must take account of transport costs on the part of its total output sold abroad, namely q_{12} for firm 1 and q_{21} for firm 2.

As we know from microeconomics (see, for example, Kreps (1990, sect. 10.1)) the equilibrium point in Cournot's duopoly can be determined employing the *reaction curves* (or *best-reply functions*, as they are sometimes called). A reaction curve shows the optimal quantity supplied by a duopolist for any given quantity supplied by the other one. In our case, we have two couples of such curves, namely one couple in each market, that we indicate by R_{ij} (reaction curve of firm i on market j). In Fig. 9.7 we have drawn the two couples of reaction curves, that for simplicity's sake we have assumed linear. They are also assumed to be *separable*, namely the reaction curve of a firm in a market only depends on the quantities being supplied (by the firm under consideration and by the rival) in that market, and not on the quantities being supplied in the other market. Thus, for example, R_{11}

does not shift as q_{22} and q_{12} change. This very convenient property depends on the assumption that the marginal cost is constant (see the Appendix, Sect. 23.3.2).

It can now be shown, through the usual dynamic mechanism underlying Cournot reaction curves, that the equilibrium point is stable in both markets. Consider for example market 1, and take an arbitrary initial situation in which the local firm offers OA. The foreign firm, given its reaction curve R_{21}, will offer OB (the ordinate of point P). The domestic firm, given the supply OB from the foreign firm, will then offer OA' (the abscissa of point P' on the domestic firm's reaction curve), and so forth. The dynamic path clearly converges to the equilibrium point E_1. A similar reasoning can be applied to market 2 to show that the equilibrium point E_2 is stable.

Given the assumptions made (identical size of the two markets, identical production costs, identical demand, identical transport costs), the two equilibrium solutions are symmetrical, i.e. $q_{11}^E = q_{22}^E$ and $q_{12}^E = q_{21}^E$. Furthermore, owing to the presence of transport costs, $q_{ii} > q_{ji}$, namely in each country the share of demand satisfied by the domestic firm is greater than the share satisfied by the foreign firm.

This form of intra-industry trade due to oligopolistic interaction can be seen as a form of dumping or *reciprocal dumping*, as Brander and Krugman (1983) called it. To show this, let us begin by observing that, due to the symmetry property, the overall quantity supplied to each market will be the same in both markets and hence, since the demand functions have been assumed identical, the price also will be identical in the two markets. It follows that, due to transport costs, for each firm the f.o.b. price of exports is lower than the domestic price of the same commodity, and therefore there is a kind of reciprocal dumping.

9.3.3 Vertically Differentiated Goods

Let us recall that we are in the case in which goods differ only in quality. In the neo-Heckscher-Ohlin model of Sect. 9.1, quality was assumed to be an increasing function of capital intensity; here, we assume that it is the expenditure on R&D (Research and Development) to enable firms to produce a better good. An additional important consideration is why in this section we assume an oligopolistic market rather than a competitive one like in the neo-H-O model. The reason is that when the burden of quality improvement falls on high fixed costs such as R&D expenditure, there is an upper limit to the number of firms that can profitably operate (for simplicity's sake we assume that each firm produces only one quality). Such a situation—i.e., very high fixed costs with respect to variable cost—is called *natural oligopoly* by Shaked and Sutton (1983) and other authors that have examined it. These studies, initially referred to a closed economy, were then extended to open economies (Shaked and Sutton, 1983, Motta, 1992).

On the demand side, we assume consumers with identical tastes but with different incomes: those with a higher income are willing to pay more for a higher-quality product. Thus the market is divided in a fairly simple manner: the highest quality supplied is bought by all consumers with an income above a certain critical level; the next to highest quality is bought by all consumers in the immediately lower income bracket, and so forth.

In studying international trade, the authors start from initially closed economies, amongst which trade is subsequently opened, and distinguish between the short and the long run. In the short run, given the upper bound to the number of firms that can coexist, the opening of international trade will in any case bring about a reduction in the number of firms existing in the combined economy (countries 1 and 2 form now a single world market). If we examine for example the extreme case of two equal countries, let B denote the maximum number of firms (and so of goods) that can coexist in each of them separately considered. In the combined market still B firms at most can coexist, which means that some firms will be eliminated from the markets through price competition (the assumption is that the oligopolistic interaction does not take place through the quantity, like in the Cournot model used in Sect. 9.3.2, but through prices, like in the Bertrand-Edgeworth oligopoly model). Hence, in the post-trade situation consumers will be better off thanks to lower prices, and intra-industry trade will occur because consumers will continue demanding the B varieties of the commodity, which are now produced partly in country 1 and partly in country 2.

When the two autarkic economies are different (the diversity being measured by a different income distribution), a greater number of firms can coexist in the combined world economy when trade is opened up; but this number becomes smaller as the income distributions get nearer.

Let us now come to the long run, always starting from two initially autarkic economies. The Shaked and Sutton model shows (see Sect. 23.4.2) that the number of firms that can survive in each country is only two, and that other firms that tried to enter the market would suffer losses (hence they do not enter). What happens when international trade is opened? We must as before distinguish two cases, that in which the two economies are identical, including income distribution, and that in which they are different as regards income distribution. In the former case the same result as in the two autarkic economies will continue to hold for the combined world economy, namely no more than two firms producing two different qualities will survive. The model cannot however forecast which are these firms, so that it might happen that the two surviving firms belong to the same country. In this case there would be one-way trade, for the other country would have to import both commodities; of course there will have to exist other sectors in which such country can export, because in the context of the pure theory no country can be only an importer. When, on the contrary, the two surviving firms belong to different countries, since the consumers in both countries demand both commodities, there will be intra-industry trade with the simultaneous import and export of different qualities of the commodity. Finally observe that, since each firm will serve not only the domestic but also the foreign market, the economies of scale will allow a price reduction, hence an increase in consumers' welfare (the gains from trade).

If income distribution is different in the two autarkic countries, the number of firms that can coexist in the world economy is greater; but for our purposes it is sufficient to observe that the result will be in any case the creation of intra-industry trade to satisfy consumers' demands in both countries.

9.3.4 Horizontally Differentiated Goods

Eaton and Kierzkowski (1984) considered the case of an economic system where two goods are produced: a homogeneous commodity (good A, produced under constant returns to scale) and a horizontally differentiated commodity (good B, produced under increasing returns to scale). While the market for good A is perfectly competitive, market B is oligopolistic.

The firms in sector B first choose the variety of the good to be produced (each firm is assumed to produce only one variety) and then decide the price. More precisely, the assumption here is that a firm incurs the fixed cost when it chooses a variety to produce, before it decides on the level of output and price. Thus, the decisions concerning entry and price are taken sequentially rather than simultaneously. According to the authors, this is consistent with the views of Linder (see Sect. 8.4.1), who holds that production is typically first developed for the domestic market; international trade takes place only later, when firms have already selected their models and incurred fixed costs.

Oligopolistic interaction takes place through prices, according to a modified Bertrand assumption. More precisely, when a firm contemplates price reductions it assumes that the other firms will not change their price, while when it considers price increases it anticipates that the competitors will lower their price.

The demand for the differentiated commodity follows Lancaster's approach based on characteristics (see Sect. 9.2). We must add that consumers will be willing to demand the differentiated good provided that the price of the variety they desire is not higher than a certain critical level, above which they will demand the homogeneous good only.

The opening of trade between such economies will give rise to a vast number of short-run and long-run effects, partly depending on the number of firms existing in the two countries before and after trade. Thus the authors are compelled to adopt a taxonomic approach. Among the several cases they examine there is that in which free trade is not the best situation for a country, which, on the contrary, can improve its welfare levying a tariff on imports of the differentiated good. To show this let us assume that in the pre-trade situation commodity B is not produced in country 2, for example because its price would be higher than the critical level, so that consumers do not demand it and spend all their income on the homogeneous commodity. In country 1, on the contrary, consumers demand both the homogeneous commodity and commodity B (only one variety, produced by a single firm, is assumed to exist) because their critical price is higher than that of country 2's consumers. Let us limit ourselves to the short-run effects, so that the productive situation remains unchanged. With the opening of trade country 1's producer of good B will try to sell also in country 2's market by lowering the price. But since no market discrimination is assumed to exist, this producer will have to lower the price also in the domestic market. Country 1's consumers will benefit, and the producer will get higher profits. It is in fact obvious that the producer under consideration, who already earned monopoly profits in country 1's market before the opening of trade, will decide to sell also in market 2 by reducing the price only if the elasticity of the two countries's combined demand shows this decision to be the superior alternative.

Let us now ask what happens to country 2. Local consumers will have no benefit, because the monopolist producer of commodity B will be able to charge a price that in the margin will leave country 2's consumers indifferent between consuming the homogeneous commodity only (like in autarky) or consuming both the homogeneous and the differentiated commodity. Thus we conclude that free trade benefits country 1 but leaves unchanged the welfare of country 2, contrary to the result of the traditional theory, according to which, as we know, free trade is beneficial to both countries.

Under the heading of intra-industry trade in horizontally differentiated goods produced by oligopolistic firms we also must mention the so-called "biological" model of trade (Bhagwati, 1982). In biology the same set of genetic traits, or genotype, interacts with different environments and gives rise to different actual biological forms, or phenotypes. In economics, the same set of know-how and technological capabilities (the genotype) will interact with different local historical and cultural environments (including different tastes) to give rise to different varieties of a horizontally differentiated good (the phenotypes). In other words, each country in autarky tends to specialize in the production of those varieties of a differentiated good that best suit the tastes of the domestic consumers. When trade is opened, consumers will be better off by consuming more varieties of the commodity, and intra-industry trade will result. For a formalisation of this approach see Dinopoulos (1988).

9.4 Empirical Studies in the Light of Theory

Most of the empirical studies on the non-traditional theories of international trade concern the monopolistic competition model in its several variants. These studies are the subject of the present section.

At the end of the 1970s there was a rather visible discrepancy between international trade theory and international trade facts. The theoretical paradigm based on comparative advantage was elegant, profound, and intellectually appealing but spectacularly at odds with the observed patterns of trade. As Deardoff (1984, p. 499) notes in his *Handbook* chapter, "The Ricardian and Heckscher-Ohlin theories are thought by many to provide a less than complete explanation of world trade. The reason for this dissatisfaction lies only partly in the somewhat ambiguous support that tests of the theories have provided. Rather, many authors have noted a number of empirical regularities in the data of international trade that seem, on the surface at least, to be unexplainable in terms of these dominant theories." In particular, three empirical regularities constituted a puzzle for the comparative advantage theories while standing strongly in support of the monopolistic competition model of international trade. The first was represented by the dominant presence of intra-industry trade. Many studies interpreted the large volume of intra-industry trade with respect to inter-industry trade as evidence in support of the monopolistic competition model and against the comparative advantage model (see Leamer & Levinsohn, 1995, for a critical appraisal). The second was represented by the excellent empirical performance of the gravity equation and the fact that the latter can be derived directly from the monopolistic competition model (as we have seen in Sect. 9.2.6). A prominent contribution based on this fact is Helpman (1987), which carried out an extensive empirical analysis of the monopolistic competition model on OECD data from 1956 to 1981. He tested both a model in which all trade is intra-industry and a model in which intra-industry and inter-industry trade coexist. His conclusion was that the theory finds some support in the data. The third was the gigantic volume of trade among developed countries (countries with similar technology and factor endowments) relative to the volume of trade between developed and developing countries (countries with different technology and factor endowments). This fact is precisely the contrary of what the Ricardian and Heckscher-Ohlin theory predicted. However, by the beginning of the 1990s, these views were challenged on theoretical and empirical grounds. Studies, such as Davis (1995), Deardorff (1998), Eaton and Kortum (2002), and Evenett and Keller (2002), showed that intra industry trade and gravity-type predictions may be derived from a variety of other models, not only in monopolistic competition. Furthermore, Hummels and Levinsohn (1995) found that the gravity equations also fitted excellently with a data set for non-OECD countries, a piece of evidence that they plausibly interpret as being at odds with the assumptions of the monopolistic competition model of trade. Davis (1996) showed that large volumes of trade between countries with similar endowments and technologies and small volumes between countries with different endowments and technologies do not require monopolistic competition and are perfectly consistent with the Ricardian and Heckscher-Ohlin theories. These studies, combined with the new evidence in favor of an amended version of the Heckscher-Ohlin-Vanek model (discussed in Sect. 4.4), made it clear that further investigation was needed to assess the empirical merits of the monopolistic competition model.

An innovative approach was proposed by Davis and Weinstein (1999, 2003). The novelty of their approach is that they identify a discriminating criterion that allows to distinguish between the monopolistic competition and the perfect competition models of Heckscher-Ohlin inspiration. The discriminating criterion is based on the demand-specialization manifestation of the home market effect (HME) that we have already encountered in Sect. 9.2.4.2. They argue that there is a more than proportional relationship between the share of output and the share of demand in monopolistic competitive sectors, while there is a less than proportional relationship in perfectly competitive sectors. They regress the share of output on the share of demand on a data set comprising a large number of countries and industries. The estimated coefficient of such regression indicates whether there is an HME. An estimated coefficient statistically larger than one is consistent with the HME and therefore constitutes evidence in favor of the monopolistic competition model of international trade. Conversely, an estimated coefficient smaller than one is inconsistent with the HME and therefore constitutes evidence in favour of the perfect competition model. Their results show evidence of the existence of the HME when using aggregate expenditure but only mild evidence of the existence of the HME at the sector level. The work of Davis and Weinstein has stimulated a lively research programme. Head and Ries (2001) use the sensitivity of the HME to changes in trade costs as a discriminating criterion. They find that the size of the relationship between share of output and share of demand decreases with trade costs in constant returns to scale and perfectly competitive sectors, while it increases with trade costs in increasing return and monopolistic competitive sectors. They use this feature as a discriminating criterion. They find evidence in support of both models depending on whether parameter identification comes from the cross section or from the time series, but the perfect competition model seems to be supported more strongly. Hanson and Xiang (2004) have tested a different version of the HME, namely that larger countries tend to export relatively more of high-transport-cost, strong-scale-economy goods and relatively less of low-transport-cost, weak-scale-economy goods. They tested this prediction on country pairs' exports to third markets and found evidence of HME in high transport-cost, strong-scale-economy industries, as predicted by the theory. Davis (1998) was the first to note that most of the theoretical and empirical studies on the HME assume the existence of an outside good (a freely-traded good produced with constant returns to scale and in perfect competition like good A in Sect. 9.2.4.2). He shows that prohibitive trade costs in the outside good eliminate the HME. Crozet and Trionfetti (2008) follow up on Davis' work. They find that in a slightly more general theoretical setting, the HME is attenuated and becomes non-linear. They also pursue an empirical investigation and find evidence of a pervasive but quantitatively mild presence of the HME in its non-linear shape. Brülhart and Trionfetti (2009) develop a new discriminating criterion using home-biased expenditure. The criterion predicts that countries' relative output and their relative home biases are positively correlated in differentiated-goods sectors (the "home-bias effect"), while no such relationship exists in homogeneous-goods sectors. Their empirical results suggest that the monopolistic competition model fits particularly well for a number of sectors that account for some 40 % of sample manufacturing output. Other works, such as Redding and Venables (2004) and Head and Mayer (2006), find evidence of the existence of the demand-wages manifestation of the HME discussed in Sect. 9.2.4.1.

Strong support in favour the synthesis model discussed in Sect. 9.2.2 is found in Romalis (2004). He examines how factor proportions determine the structure of commodity trade in a many-country version of the synthesis model to which he adds iceberg trade costs. The commodity structure of production and bilateral trade is fully determined thanks to trade costs and monopolistic competition. He finds two important results. The firs is that countries capture larger shares of world production and trade in commodities that make more intensive use of their abundant factors (the Heckscher-Ohlin theorem). The second is that countries that rapidly accumulate a factor see their production and export structures systematically shift towards industries that use that factor intensively (the Rybczynski theorem).

We have seen in Sect. 9.2 that welfare is higher in free trade than autarky because the number of varieties available to consumers expands. For a long time this source of gain from trade remained empirically unexplored. Broda and Weinstein (2006) were the first to measure this gain from trade. They estimate that the gain from trade for US consumers between 1971 and 2001 was 2.6 % of GDP. Expressed differently, they find that consumers in the US would be willing to pay 2.6 % of their income to access the wider set of varieties available in 2001 rather than those available in 1972.

Coming to heterogenous firm models we note that some empirical regularities, such that not all firms exports, that exporters are larger than non-exporters, and that trade liberalization leads to the reallocation of market shares, not only are explained by heterogenous firm models but constitute one of the motivations for the development of such models. This is only one the merits of this family of models. As Bernard, Jensen, Redding, and Schott (2012) note in their comprehensive and instructive appraisal of the empirical literature, heterogenous firm models also paved the way to new explorations on the relationship between trade liberalization and aggregate economic variables, such as the composition of intra-industry trade flows or the implications for the labour markets. Furthermore, they contributed to the understanding of the relationship between trade liberalization and the internal organization of the firm, of its decisions concerning offshoring, of the modalities of procuring inputs and of choosing the strategies of international expansion.

References

Anderson, J. E. (1979). A theoretical foundation for the gravity equation.

Baldwin, Robert E. (1992). Are economists' traditional trade policy views still valid?

Baldwin, Richard E., & Ottaviano, G. I. P. (2001). Multiproduct multinationals and reciprocal FDI dumping.

Behrens, K., Lamorgese, A., Ottaviano, G. I. P., & Tabuchi, T. (2009). Beyond the home market effect: Market size and specialization in a multi-country world.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2007). Comparative advantage and heterogenous firms.

Bhagwati, J. N. (1982) Shifting comparative advantage, protectionist demands, and policy response.

Brander, J. A. (1981) Intra-industry trade in identical commodities.

Brander, J. A., & Krugman, P. R. (1983). A reciprocal dumping model of international trade.

Brander, J. A., & Spencer, B. J. (1984). Tariff protection and imperfect competition.

Brander, J. A., & Spencer, B. J. (1985). Export subsidies and international market share rivalry.

Broda, C., & Weinstein, D. E. (2006). Globalization and the gains from variety.

Brülhart, M., & Trionfetti, F. (2009). A test of trade theories when expenditure is home biased.

Crozet, M., & Trionfetti, F. (2008). Trade costs and the home market effect.

Davis, D. R. (1995). Intraindustry trade: A Heckscher-Ohlin-Ricardo approach.

Davis, D. R. (1996). Critical evidence on comparative advantage? North-north trade in a multilateral world.

Davis, D. R. (1998). The home market, trade and industrial structure.

Davis, D. R., & Weinstein, D. E. (1999). Economic geography and regional production structure: An empirical investigation.

Davis, D. R., & Weinstein, D. E. (2003). Market access, economic geography and comparative advantage: An empirical test.

Deardoff, A. V. (1984). Testing trade theories and predicting trade flows.

Deardorff, A. V. (1998). Determinants of bilateral trade: Does gravity work in a neoclassical world?

Dinopoulos, E. (1988). A formalization of the "biological" model of trade in similar products.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity.

Eaton, J., & Kierzkowski, H. (1984). Oligopolistic competition, product variety and international trade.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade.

Evenett, S. J., & Keller, W. (2002). On theories explaining the success of the Gravity equation.

Falvey, R. E. (1981). Commercial policy and intra-industry trade.

Falvey, R. E., & Kierzkowski, H. (1987). Product quality, intra-industry trade and (Im)perfect competition.

Flam, H., & Helpman, E. (1987). Industrial trade policy under monopolistic competition.

Friedman, J. W. (1977). *Oligopoly and the theory of games.*

Gabel, H. L., & Neven, D. (1988). *Fair trade in commercial aircraft: The case of boeing vs. airbus industry.*

Gabszewicz, J. J., Shaked, A., Sutton, J., & Thisse, J.-F. (1981). International trade in differentiated products.

Grossman, G. M. (Ed.). (1992) *Imperfect competition and international trade.*

Grossman, G. M., & Richardson, J. D. (1985). *Strategic trade policy: A survey of the issues and early analysis.*

Haberler, G. (1990). Strategic trade policy and the new international economics: A critical analysis.

Hanson, G. H., & Xiang, C. (2004). The home market effect and bilateral trade patterns.

Harrigan, J. (1994). Scale economies and the volume of trade.

Harrigan, J. (1996). Openness to trade in manufactures in the OECD.

Head, K., & Mayer, T. (2004). The empirics of agglomeration and trade, chap. 59. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of urban and regional economics* (Vol. 4). North Holland.

Head, K., & Mayer, T. (2006). Regional wages and employment responses to market potential in the EU.

Head, K., & Mayer, T. (2013). Gravity equations: Workhorse, toolkit, and cookbook.

Head, K., & Ries, J. (2001). Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade.

Head K., Mayer, T., & Ries, J. (2002). On the pervasiveness of home market effects. *Economica, 69*, 357–370.

Helpman, E. (1987). Imperfect competition and international trade: Evidence from fourteen industrialized countries.

Helpman, E. (1990). *Monopolistic competition in trade theory.*

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade.*

Helpman, E., & Krugman, P. R. (1989). *Trade policy and market structure.*

Hummels, D., & Levinsohn, J. A. (1995). Monopolistic competition and international trade: Reconsidering the evidence.

Kierzkowski, H. (1985). Models of international trade in differentiated goods.

Kreps, D. M. (1990). *Microeconomic theory.*

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade.

Krugman, P. R. (1980). Scale economies, product differentiation, and the pattern of trade.

Krugman, P. R. (1981). Intraindustry specialization and the gains from trade.

Krugman, P. R. (1990). *Rethinking international trade.*

Leamer, E. E., & Levinsohn, J. (1995). International trade theory: The evidence.

Markusen, J. R. (1981). Trade and the gains from trade with imperfect competition.

Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: Theory and evidence* (chap. 17).

Melitz, M. (2003). The impact of trade on intra-industry reallocation and aggregate industry productivity.

Motta, M. (1992). Sunk costs and trade liberalization.

Ottaviano, G. I. P., Tabuchi, T., & Thisse, J.-F. (2002). Agglomeration and trade revisited.

Pomfret, R. (1992). International trade policy with imperfect competition.

Redding, S. J. (2011). Theories of heterogeneous firms and trade.

Redding, S., & Venables, A. J. (2004). Economic geography and international inequality.

Romalis, J. (2004). Factor proportions and the structure of commodity trade.

Shaked, A., & Sutton, J. (1983). Natural oligopolies.

Shaked, A., & Sutton, J. (1984). Natural oligopolies and international trade.

Spence, A. M. (1976). Product selection, fixed costs, and monopolistic competition.

Varian, H. R. (1992). *Microeconomic analysis.*

Venables, A. J. (1985). Trade and trade policy with imperfect competition: The case of identical products and free entry.

Venables, A. J. (1987). Trade and trade policy with differentiated products: A Chamberlinian-Ricardian model.

Venables, A. J. (1990). International capacity choice and national market games.

Part IV
Trade Policy

Chapter 10
Tariff and Non-tariff Barriers

10.1 Introduction

This chapter is concerned with what is called the *theory of commercial policy* in the broad sense. The traditional theory focused on tariffs, starting from two principles generally accepted until the first world war. These were: (a) that impediments to international trade for protectionist purposes should be limited to tariffs, and (b) that no commercial discrimination between supplier countries should be instituted, in the sense that, if a tariff is levied on some imported commodity, it should be applied at the same rate and to all imports of that commodity independently of the supplying country.

Notwithstanding the fact that in the inter-war period, and especially during the Great Depression, these principles were systematically violated, they were taken up again and made the foundation of the international agreement that, it was hoped, was to rule international trade after the second world war: GATT (the General Agreement on Tariffs and Trade). Several international meetings for the purpose of negotiating multilateral tariff reductions (the various "rounds", the last being the Uruguay round ended in December 1993) took place under the aegis of GATT (now replaced by WTO, the World Trade Organization, on which see Sect. 10.2) which, however, had to take a permissive attitude towards the violations of the above principles. The last few decades have seen an expansion of both non-tariff barriers to trade and discriminatory commercial policies (preferential trading agreements etc.), so that the traditional theory (see, e.g., Balassa, 1965, Bhagwati, 1965, 1971, Bhagwati et al., 1998, Corden, 1974, 1984b, El-Agraa, 1984, Greenaway, 1983, Johnson, 1969, Lloyd, 1974, Meade, 1952, 1955, Pearce, 1970, Stern, 1973, Takayama, 1972, Vanek, 1962, Vousden, 1990) has had to be broadened to make the rigorous analysis of these phenomena possible. The emergence of a "new" protectionism, including administered protection, lobbying for protection, and so on, will be dealt with in Chap. 12.

It is usual to distinguish a *positive* and a *normative* (or welfare) theory of commercial policy. The former examines the various effects (on the pattern of

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_10,

consumption, on the allocation of resources etc.) of the imposition of tariffs and of other measures of intervention on trade, but is not concerned with evaluating their desirability and even less with defining the properties of a set of optimal measures: these are the concern of the normative theory. Naturally, in practice it is difficult to separate the positive from the welfare aspects, so that—though the present chapter concentrates on the former and Chap. 11 on the latter—both aspects will be present throughout our treatment.

Before going on, it is advisable to say a few words on the institutional setting.

10.2 GATT and WTO

GATT (General Agreement on Tariffs and Trade) was established in 1947 on a provisional basis with the aim of providing an international forum for negotiating tariff reductions, agreeing on world trade disciplines, solving trade disputes. Provisional because GATT was meant to pave the way for a specialized agency of the United Nations, the ITO (International Trade Organization), to be established shortly afterwards. This did not take place because the national ratification of the ITO charter proved impossible in some countries (amongst which the United States). Thus provisionality lasted for 47 years, until WTO (World Trade Organization) was established.

GATT has promoted international trade liberalization in several ways. It has outlawed the use in general of import quotas, and established the extension to all members of the MFN (Most Favoured Nation) treatment. Under Article I of GATT (also called the MFN clause), members have committed themselves to give to the products of other members a treatment no less favourable than that granted to the products of any other country. Thus, no country can give special advantages to another country or discriminate against it. GATT has also provided a negotiating framework for tariff reductions through multilateral trade negotiations or "trade rounds", the last and most extensive being the Uruguay round (1986–1993). These negotiations have involved not only tariffs, but also subsidies and countervailing measures, anti-dumping, technical barriers to trade, government procurement, and so on.

The original agreement (called GATT 1947) was amended and updated in 1994 (GATT 1994). GATT 1994 is an integral part of WTO, which was established on 1st January 1995.

As the names say, WTO is an organization (see WTO, 1995), while GATT 1947 was an agreement. This is not only a semantic difference or a juridical subtlety: an *agreement* is simply a set of rules with no legal institutional foundation; a (permanent) organization is an *institution* with legal personality and its own secretariat and powers. This implies, amongst other, that the WTO dispute settlement system is faster and more automatic, and the implementation of its decisions on disputes is more easily assured.

Box 10.1 Multilateral Trade Rounds

Since GATT's creation in 1947–1948, there have been eight rounds of trade negotiations, whilst a ninth round, under the Doha Development Agenda, is now underway and expected to end by 1 January 2005 (see table below). The first GATT trade rounds concentrated on further reducing tariffs. With the Kennedy Round, over the 1960s, an Anti-Dumping Agreement and a section on development were brought into the GATT, while the Tokyo Round was the first major attempt to tackle also non-tariff trade barriers. The eighth Round, the Uruguay Round lasted for 8 years and led to the creation of the WTO and to a new set of agreements, such as the General Agreement on Trade in Services (GATS) and on Trade-Related aspects of Intellectual Property (TRIPS).

Year	Place/name	Subjects covered	No. of parties
1947	Geneva	Tariffs	23
1949	Annecy	Tariffs	13
1951	Torquay	Tariffs	38
1956	Geneva	Tariffs	26
1960–1961	Geneva (Dillon Round)	Tariffs	26
1964–1967	Geneva (Kennedy Round)	Tariffs and anti-dumping	62
1973–1979	Geneva (Tokyo Round)	Tariffs, non tariff measures, framework agreements"	102
1986–1994	Geneva (Uruguay Round)	Tariffs, non-tariff measures, rules, services, intellectual property, dispute settlement, textiles, agriculture, creation of WTO	123

While the Singapore ministerial conference (1996) defined the WTO work plan, the Geneva ministerial meeting, held in 1998, provided the mandate to launch a new round of negotiations at its next summit, in Seattle (1999). As the Seattle ministerial meeting turned out to be a complete failure, with critical issues separating industrialised and developing countries, the next negotiating round was launched in Doha, in 2001. The Doha Round delivered the Doha Development Agenda, which, recognising the major role that international trade plays in promoting economic development and poverty alleviation, comprises further market opening and additional rule making, strengthened by commitments to increase assistance to build capacity in developing countries. It also added negotiations and other work on, among others, non-agricultural tariffs, trade and environment and WTO rules such as anti-dumping and subsidies. With the end of the Cancùn ministerial conference (September 2003) without consensus, decisions related to the implementation of the Doha agreement were further postponed. In August 2004 the so-called July Package was approved. which established a number of objectives concerning principally the three main themes of the confrontation on the agricultural sector (internal support, export subsidies, access to markets) and fixed the conclusion of the negotiations at the ministerial conference to be held in Hong Kong the following year. However, at the Hong Kong meeting no step forward was done and since then the negotiations are at a deadlock. However, it has been fixed the date of 2013 for the dismantling of agricultural export subsidies by developed ountries, and measures have been taken to favour the access of developing countries to advanced international markets. The irreconcilability of the various positions of the participants in the negotiations led the director-general of WTO to suspend the Doha round in July 2006. The negotiations were taken up again in 2007, but without much success: the deadlock remains, and the prospects of the Doha round remain very uncertain.

From the economic point of view, WTO has a greater scope than GATT, for GATT rules applied solely to trade in merchandise, while WTO in addition to goods also covers trade in services as well as trade-related aspects of intellectual property.

GATT, and now WTO, are sometimes described as free-trade institutions. This is not entirely correct, if only because tariffs (and, in limited circumstances, other forms of protection) are permitted. The basic aim of GATT and WTO rules is to secure open, fair and undistorted competition in international trade. Rules on non-discrimination, as well as those on dumping and subsidies (governments are allowed to impose compensating duties on these forms of unfair competition), are designed to bring about fair conditions of trade.

10.3 Partial Equilibrium Effects of a Tariff

We begin with the traditional study of the effects of a tariff; henceforth the tariff is assumed to have the form of an *ad valorem* tax on imports (so that, if p is the pre-tariff price, the cum-tariff price will be $(1+d)p$, where d[1] is the tariff rate) and not of a specific tariff (so many dollars per unit of the commodity).

The effects of a tariff can be examined either in a partial or a general equilibrium context. In the former case one considers solely the market for the commodity on which the tariff is imposed and neglects—by a *ceteris paribus* clause—the repercussions on and from the rest of the system; these, on the contrary, are explicitly brought into the analysis in the latter case (see Sect. 10.5).

In Fig. 10.1a we have drawn the domestic demand and supply curves—for simplicity's sake they are assumed linear and normal—for the commodity being examined. If we assume that its world price is p, this will also be its domestic price given the usual assumptions (perfect competition, no transport costs, no tariffs). At this price the imports of the commodity are FH, equal to the domestic excess demand. If a tariff is now levied, say d_1, the domestic price will increase to $p(1+d_1)$ at the same world price p. This implies the assumption that the country levying the tariff is small, so that the variation in its import demand due to the tariff has negligible effects on the world market of the commodity, and the world price remains constant. This assumption will subsequently be dropped.

The consequence is that demand decreases, domestic output (supply) increases and imports decrease from FH to F_1H_1. As an extreme case, it is possible to conceive a tariff—d_2 in Fig. 10.1a—so high that the increase in the domestic price brings this to the level at which domestic demand and supply are equal and imports cease: such a tariff is called a *prohibitive* tariff.

In these brief considerations all the effects of the tariff are included, and can be made explicit as follows:

[1]The symbol generally used for the tariff rate is t. However, since in this book we have used the symbol t to denote time, another symbol (d, from duty) has been used to indicate the tariff rate.

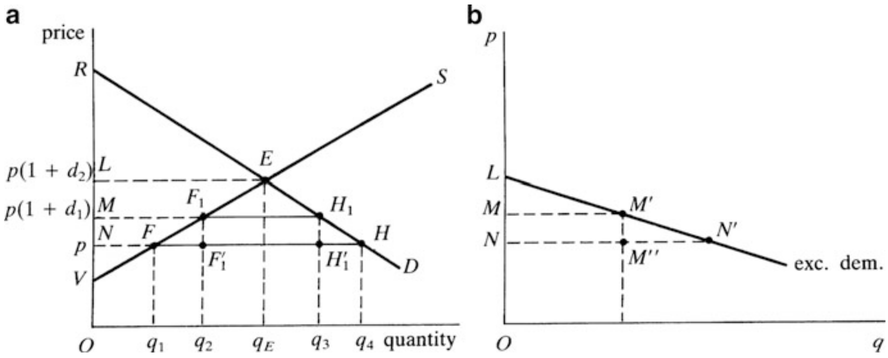


Fig. 10.1 Partial equilibrium effects of a tariff

1. *Consumption effect.* Domestic consumption of the commodity decreases by $q_3q_4 = HH_1'$.
2. *Production* (or protective) *effect.* Domestic output increases by $q_1q_2 = FF_1'$.
3. *Import effect.* Imports decrease by an amount equal to the sum of the two previous effects, as $q_2q_3 = q_1q_4 - (q_3q_4 + q_1q_2)$.
4. *Fiscal revenue effect.* The tariff represents a fiscal revenue for the government of the levying country. To calculate total tariff revenue, note that it is given by the absolute value of the tariff per unit of the commodity multiplied by the quantity imported. The former is $p\,(1 + d_1) - p = d_1 p = MN = F_1F_1'$, the latter is $q_2q_3 = F_1H_1$. Therefore total tariff revenue is $F_1F_1' \times F_1H_1$, that is, the area of the rectangle $F_1F_1'H_1'H_1$.
5. *Redistribution effect.* Since the price has increased, there is a redistribution of income from consumers to producers. This point needs to be gone into a little further.

Actually, it can be said that consumers subsidize the domestic production of the commodity by an amount *MN* per unit, so that the total subsidy is $MNF_1'F_1$. This is also called the *subsidy-equivalent* of the tariff; in other words, if the government directly subsidized the domestic production, instead of imposing a tariff, the total cost of the subsidy to obtain the same amount of protection would be exactly equal to the subsidy-equivalent. In fact, to induce domestic firms to produce the quantity Oq_2 and sell it at unit price *ON* instead of *OM* (in the absence of the tariff the price would remain at *ON*), it is necessary to give them a subsidy equal to the revenue loss, which is exactly $MNF_1'F_1$.

But consumers do not only pay out the subsidy-equivalent: they are also taxed by an amount equal to the tariff revenue which accrues to the government, because this amount ultimately comes out of their own pockets. We can therefore define a *consumer tax equivalent* to the tariff as the sum of the subsidy-equivalent and the tariff revenue. In other words, if—instead of the tariff—a consumption tax were imposed, with the aim of reducing consumption by the same amount as would be reduced in consequence of the tariff, then the unit rate of this tax would have to be *MN*, which would give rise to a fiscal revenue equal to $MNH_1'H_1$, in turn equal to $MNF_1'F_1$ (subsidy-equivalent) $+F_1F_1'H_1'H_1$ (tariff revenue). As a matter of fact, the tariff has the same effect as a consumption tax (with the same rate as the tariff), the revenue of which is used by the government partly to subsidize domestic producers and partly to increase its fiscal revenue.

10.4 The Social Costs of a Tariff

We must now investigate whether, account being taken of the various effects, the imposition of a tariff is beneficial or not. The traditional theory proposed to show that a tariff involves a *cost* for society (economic cost of the tariff or cost of protection, as it is also called).

The basis for this demonstration is the concept of *consumers'* surplus,[2] which can be measured as the area under the demand curve included between the line of the price, the price axis and the demand curve itself. For example, in Fig. 10.1a, consumers' surplus is measured—when the price is p and the quantity q_4—by the area of triangle *NHR*.

Now, with the increase in price from p to $p(1 + d_1)$, consumers' surplus decreases by NHH_1M. This is a cost; to compute the *net* cost, if any, we must calculate the benefits. These are the tariff revenue accruing to the government, $F_1F_1'H_1'H_1$, and the increase in producers' surplus[3] $MNFF_1$. It is important to stress that, in order to be able to net out benefits from costs (both are expressed in money, and so are dimensionally comparable) we must assume that *each dollar of gain or loss has the same importance independently of who is gaining or losing*. Without this assumption, in fact, it would *not* be possible to compare the consumers' loss with the producers' and the government's gain.[4]

Given this assumption, it can readily be seen from the diagram that the reduction in consumers' surplus is only partly offset by the tariff revenue and the increase

[2]It is as well to point out that consumers' surplus—defined by Alfred Marshall as the excess of the total price that consumers would be willing to pay rather than go without the commodity, over that which they actually pay—is a much debated concept and a source of much confusion (it has been humorously renamed "confuser surplus" by Morey (1984)). The graphic measure used in the text is only one of the measures possible and hinges on several simplifying assumptions, amongst which the constancy of the marginal utility of money (see, for example, Hicks, 1981). It should also be stressed that *consumption* and *consumer* should be interpreted in the broad sense to mean *purchase* and *purchaser* respectively, for whatever purpose the product is bought.

[3]Unlike consumers' surplus, this is a well-defined concept, as it is a synonym for the firms' profit (difference between total revenue and total cost). If we neglect the fixed cost (which has no consequence on the variations), the total cost of any given quantity, say q_1, is the area under the marginal cost (i.e. the supply) curve from the origin to the ordinate drawn from that quantity ($OVFq_1$). As total revenue is $ONFq_1$, producers' surplus is *VNF*. If we consider an increase in output from q_1 to q_2, the increase in producers' surplus is $VMF_1 - VNF = MNFF_1$.

[4]It should be further noted that without this assumption it would not even be possible to sum the surpluses of the single consumers to obtain the aggregate consumers' surplus, etcetera.

in producers' surplus: we are left with the areas of the two triangles $FF_1'F_1$ and $H_1'HH_1$, which represent the *social costs* of the tariff.

The first one, $FF_1'F_1$, measures the production cost of protection. If the country had imported an additional amount q_1q_2 at the price p, its cost would have been $q_1q_2F_1'F$. Instead the country produces this amount domestically, with an additional cost measured by the increase in the area below the supply curve, $q_1q_2F_1F$. The difference $FF_1'F_1$ represents the cost of the misallocation of resources caused by the tariff: in fact, if the country had used an amount of resources equal in value to $q_1q_2F_1'F$ to increase the output of its export industry (not shown in the diagram), with the consequent increase in exports it could have obtained q_1q_2 more of the imported commodity. When instead it increases the domestic production of this commodity, the country must use a greater amount of resources (equal in value to $q_1q_2F_1F$) to obtain the same additional amount (q_1q_2) of the commodity.

The second one, $H_1'HH_1$, measures the consumption cost of protection, due to the fact that the tariff brings about an increase in the domestic price of the imported commodity relative to the price of the other commodities and so causes a distortion in consumption.

An alternative graphical representation of the cost of protection is contained in Fig. 10.1b, in which the excess demand curve—derived from the D and S curves of Fig. 10.1a—is drawn. This curve, therefore, represents the demand for imports of the commodity by domestic consumers: for example, NN' in Fig. 10.1b is equal to FH in Fig. 10.1a and, similarly, $MM' = F_1H_1$. It follows by construction that the area $MNN'M'$ is equal to the area F_1FHH_1. Now, the area $MNM''M'$ represents the tariff revenue, as it is equal to the absolute unit tariff MN times the quantity imported MM' and is therefore equal to the area $F_1F_1'H_1'H_1$. We thus can see the cost of protection as the area of the triangle $M'M''N'$.

The results obtained above enable us to understand the reason behind the traditional statement that *free trade is better than tariff-ridden trade*: if, in fact, the imposition of a tariff involves a social cost, the statement is immediately proved. This problem will be taken up again in Sect. 11.6.

According to some writers, the cost of protection is actually *greater* than that found above. Among the arguments for this opinion we can mention the *administrative cost* and the *resource displacement cost* of tariffs. To impose tariffs, a country must maintain a special administrative structure (customs, border patrols, etc.) and so bear the relative cost. This cost will have to be deducted from the tariff revenue, so that the net benefit for the government is less than the area $F_1F_1'H_1'H_1$. Besides, as we have seen, a tariff causes an increase in the domestic output of the protected commodity and so a greater use of resources which—assuming full employment—will have to be shifted from other sectors; this shift involves a cost (displacement of the resources).

It goes without saying that the latter cost will not be present if there is underemployment of resources (a case, however, not contemplated by the traditional theory, where full employment is assumed): in such a case, on the contrary, a tariff will have beneficial effects. These are the *employment effects* of the tariff: with less

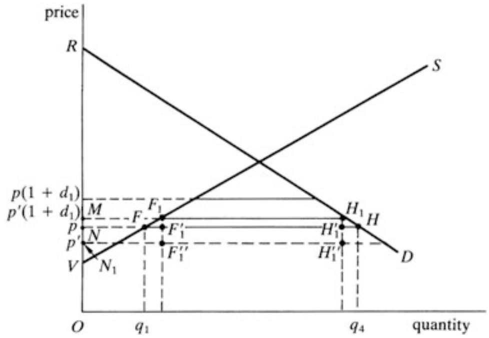


Fig. 10.2 Variations in the world price, and benefits of a tariff

than full employment, the imposition of a tariff, by causing an increase in the domestic output of the imported commodity, will ultimately increase the employment of domestic factors. This effect, however, is certainly present only under the hypothesis that exports remain the same. If, on the contrary, these decrease because foreign countries impose a tariff in retaliation, employment will decrease in the sector of exportables. It is then impossible to determine a priori the net employment effect of the tariff.

The analysis so far carried out assumes—as stated at the beginning of this section—that the domestic price increases by the same amount as the absolute value of the tariff applied to the pre-trade world price of the commodity, owing to the hypothesis that the latter price does not vary. It is however conceivable that the world price decreases in consequence of the tariff: this may be due to the usual demand-supply mechanisms set into motion by the decrease in the demand for the commodity on the world market or to the fact that the foreign country, to offset the tariff and avoid a fall in its exports to the tariff-imposing country, gives a subsidy to its exporters, who reduce the price they charge. This reduces the cost of protection, and it is even possible that an *improvement*, instead of a social cost, takes place in the tariff-imposing country. This possibility is illustrated in Fig. 10.2, which is based on Fig. 10.1a.

As a consequence of the tariff, the world price decreases, for example to p', so that the cum-tariff domestic price is $p'(1 + d_1)$, lower than $p(1 + d_1)$. The decrease in consumers' surplus is measured by NHH_1M. On the side of benefits we count as usual the increase in producers' surplus ($MNFF_1$) and the increase in the government's fiscal revenue, $F_1F''H''_1H_1$. For convenience of analysis let us break this rectangle in two parts: $F_1F''_1H''_1H_1 = F_1F'_1H'_1H_1 + F'_1F''_1H''_1H'_1$. The first of these, added to producers' surplus, leaves the two triangles FF'_1F_1 and H'_1HH_1 (which in the previous case measured the cost of protection) unaccounted for. But now on the side of benefits there is also the area of the rectangle $F'_1F''_1H''_1H'_1$, which is far greater than the sum of the areas of the two aforementioned triangles: the balance between benefits and costs is now positive. *It follows that the tariff has brought about a net benefit to the country that imposes it!*

It can be readily seen that the reason for this benefit lies in the decrease in the world price, which means that foreign exporters have eventually taken part of the burden of the tariff upon themselves. In fact, with respect to the pre-tariff situation, domestic consumers are subjected to an increase in the price of the commodity equal to *MN* only: the remaining part of the absolute amount of the tariff (NN_1) is indirectly paid for by foreign exporters in the form of a price decrease, so that it is as if the amount $F_1'F_1''H_1''H_1'$ had been paid out by these exporters.

If, as has just been shown, it is possible for the tariff-imposing country to improve its welfare, obviously the next question to ask is how to get the *maximum* possible improvement: this leads us to a study of the so-called *optimum tariff* (optimum in the sense that it maximizes the welfare of the country which levies it). However, since this problem can be more rigorously dealt with in the context of a general equilibrium analysis, we shall examine it later (see Sect. 11.1).

We conclude this section by pointing out that the imposition of a tariff has precise effects on factor rewards (Stolper-Samuelson theorem). However, these effects can be analysed only in the context of a general equilibrium model. This will be the subject of Sect. 10.5.

10.5 General Equilibrium Effects of a Tariff

10.5.1 The Production-Possibility Frontier and Tariffs

To examine the general equilibrium effects of a tariff we first consider the "small" country case. The "small" country assumption implies that variations in its demand for imports and supply of exports have negligible effects on the world market, so that the terms of trade do not vary. This assumption will be dropped later: see Sect. 10.5.2.

For our study it is convenient to employ the diagram showing the transformation curve and the social indifference curves explained in Chap. 3—see in particular Fig. 3.14b—and taken up again in Fig. 10.3. In the initial pre-tariff situation, given the terms of trade ($p = p_B/p_A$) represented by the absolute value of the slope of the straight-line segment *RR*, $\tan\alpha$, the country's production and consumption points are E and E_C respectively; imports (of A) are $E_{CA}E_A$ and exports (of B) are $E_{CB}E_B$.

When the country levies a tariff on commodity A, the domestic relative price p_B/p_A is no longer equal to the terms of trade, but lower, equal say to $\tan\beta$ (slope of P_hP_h). Since domestic producers respond to the domestic relative price, the production point shifts from E to H. International exchange, of course, takes place at the given terms of trade (in this context they are assumed to be the same), and so the country can trade by moving from H (where it produces) along the straight line $R'R'$, parallel to *RR*, but, it should be noted, the country will *not* end up at the consumption point H_C (determined by the tangency of an indifference curve, I'', to $R'R'$), because consumers will also respond to the domestic relative price and

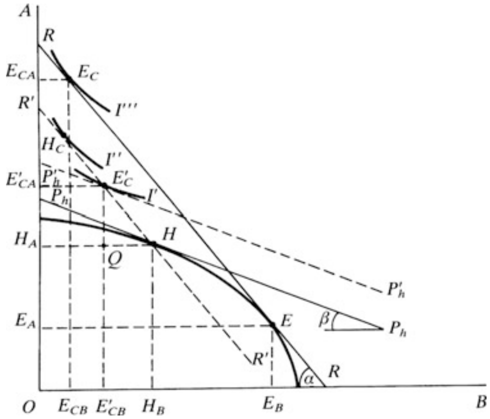


Fig. 10.3 General equilibrium effects of a tariff

so will equalize the marginal rate of substitution to this price. Thus, moving along the straight line $R'R'$ (which, we remember, represents the international exchange possibilities), we must therefore find a point where the marginal rate of substitution (slope of the indifference curve) is equal to the domestic relative price. This point is found to be E'_C, where the indifference curve I' has the same slope as P_hP_h (the straight-line segment $P'_hP'_h$ is, in fact, parallel to P_hP_h).

Let us now consider the various effects of the tariff.

The *production* (or protective) *effect* consists in the passage from E to H: the domestic output of the protected commodity increases by E_AH_A, whilst the output of the other commodity decreases by E_BH_B.

The *consumption effect* consists in the passage from E_C to E'_C: the domestic consumption of the protected commodity decreases by $E_{CA}E'_{CA}$ whilst the consumption of the other commodity increases by $E_{CB}E'_{CB}$.

The effect on the *volume of trade* consists of an *import effect* and an *export effect*. Imports decrease by $E_{CA}E_A - E'_{CA}H_A$, which is equal to the sum of $E_{CA}E'_{CA}$ and E_AH_A, i.e. to the sum of the consumption and production effects. Exports also decrease, by the amount $E_{CB}E_B - E'_{CB}H_B = E_{CB}E'_{CB} + E_BH_B$ (sum of the consumption and production effects). The final effect is a *reduction* in the volume of trade.

The *fiscal revenue effect* can be ascertained by comparing the value of national output (at factor cost) with the value of aggregate consumption expenditure, both evaluated at the new (post-tariff) domestic prices. Since the country produces at H, the value of national output is represented by the position of P_hP_h and, more precisely, national output in real terms, measured for example in terms of commodity A, is given by the intercept of P_hP_h on the vertical axis, that is by OP_h. To show this, we first observe that the value of national output corresponding to point H is $Y_H = p'_AH_A + p_BH_B$, where $p'_A = (1 + d)p_A$. The straight line P_hP_h represents all the combinations of A and B with the same values as the given Y_H, that is $p'_AA + p_{BB} = Y_H$, whence $A = -(p_B/p'_A)B + Y_H/p'_A$ which is the equation of the straight line P_hP_h. The intercept of this line on the A axis is Y_H/p'_A, i.e. the value of national output in terms of A.

Similar reasoning can be made for aggregate consumption expenditure: the value of aggregate consumption expenditure is represented by the position of $P_h' P_h'$ and, measured in terms of A, by the intercept OP_h'. The difference between the value of aggregate consumption expenditure and the value of national output is exactly the tariff revenue, because, in the presence of a tariff, aggregate expenditure exceeds national output by an amount exactly equal to consumers' outlay by way of the tariff.[5] In fact, if we consider the value of aggregate expenditure D (remember that the tariff is ad valorem and applied to commodity A) and the value of national output Y and subtract the latter from the former we get

$$\begin{aligned} D &= (1+d)\, p_A D_A + p_B D_B, \\ Y &= (1+d)\, p_A S_A + p_B S_B, \\ D - Y &= \left[p_A \left(D_A - S_A\right) + p_B \left(D_B - S_B\right)\right] + d p_A \left(D_A - S_A\right), \end{aligned} \tag{10.1}$$

where D (with subscript) and S denote the quantities demanded (consumed) and domestically supplied (produced) respectively, and the subscripts A and B refer to the commodities. Now, the expression in square brackets is the trade balance *evaluated at international prices*, which is always zero as shown in Sect. 3.3. This can be checked in the diagram by considering the triangle $E_C' QH$, where $E_C' Q = QH \cdot \tan Q\hat{H}E_C'$, and noting that: *imports* $= E_C' Q = \left(D_A - S_A\right)$, *exports*$= QH = \left(S_B - D_B\right)$, *international price ratio* $p_B / p_A = \tan \alpha =$ slope of $R'R' = \tan QHE_C'$. Hence

$$D - Y = d p_A \left(D_A - S_A\right), \tag{10.2}$$

which is the total tariff revenue.

10.5.1.1 The Redistributive and Welfare Effects of a Tariff

A tariff also affects *income distribution to the factors of production*. The imposition of a tariff favours (in the sense that it raises the unit real reward of) the factor used intensively in the production of the imported commodity. In fact, a tariff raises the domestic price of the imported commodity, and hence we can immediately apply the Stolper-Samuelson theorem (see Stolper and Samuelson, 1941, and Sect. 5.3).

It should however be pointed out that, in the anomalous (but theoretically possible) cases in which the imposition of a tariff leads to a decrease, instead of an increase, in the domestic price of the imported commodity, then the domestic output of this commodity will decrease and the factor which it uses relatively intensively will suffer a loss (the so-called *Metzler case*): in fact, with the same reasoning

[5] This is true independently of the use that the government will make of the tariff revenue: for example it may use it for public expenditure or redistribute it to consumers in various ways.

followed in the Stolper-Samuelson theorem, if p_B/p_A decreases, p_L/p_K decreases as well, and so on.

Metzler's case will be taken up again in Sect. 10.5.2.1; here we observe that all possible cases (including the anomalous ones) are accounted for by a more general formulation of the Stolper-Samuelson theorem, i.e. that the *imposition of a tariff raises the unit real reward of the factor used intensively in the sector producing the commodity whose relative price increases,* which can be either the importable commodity (in the normal case) or the other one (in Metzler's case).

We shall now examine the effects of a tariff on the welfare of the country that imposes it. In our framework the imposition of a tariff has a *social cost*: it can in fact be seen from Fig. 10.3 that the new consumption point E'_C lies on an indifference curve (I') lower than I''' where E_C was found. An alternative way of showing the cost of protection without having recourse to social indifference curves is to observe that the value of real national output (in terms of A) was OR in the initial free trade situation whilst after the tariff it is $OP_h < OR$ (even if we added, on the side of benefits, the tariff revenue, we would reach OP'_h, still lower than OR). Note also that the value of real national output *at world prices* is lower, for $OR' < OR$. The decrease in the value of real national output gives a quantitative measure of the social cost of protection.

10.5.2 Tariffs and Reciprocal Demand Curves

In this section we shall examine the effects of the imposition of a tariff when the assumption of constant terms of trade is dropped. For this purpose it is convenient to use the graphic apparatus of the reciprocal demand (or offer) curves explained in Sect. 3.4.

In Fig. 10.4 we have drawn the offer curves of the two countries, assumed to have a normal shape. Let us now assume that country 2 (importer, by assumption, of commodity B) imposes a tariff: the consequence—leaving aside the effects of the disposal of the tariff revenue by the government—is a downward shift of this country's offer curve from OG_2 to OG'_2.

In fact, since (by the definition of offer curve) country 2's economic agents are willing, when trade is free, to give up a global amount OE_A of A in exchange for OE_B of B, the consequence of the introduction of a tariff is that the quantity of A that they are now willing to supply as exports in exchange for the *same* quantity of imports is equal to the difference between the quantity of this commodity that they are willing to give up globally and the amount that they have to pay out to the government by way of duty,[6] for example, $E_A E''_A = ES$: this difference is $OE''_A = SE_B$.

[6]This implicitly assumes that the duty is paid out in terms of commodity A (the numéraire). The results would not change if it were paid out in terms of B.

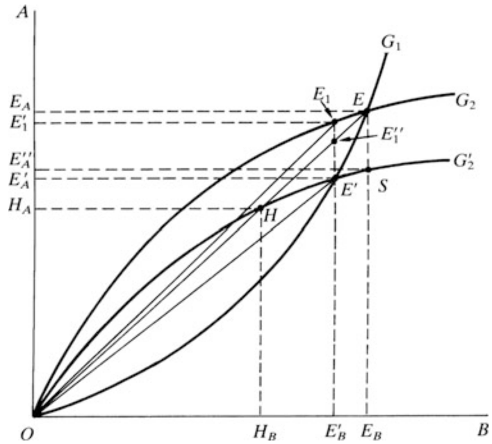


Fig. 10.4 Tariffs and terms of trade

In other words, country 2's agents will now be willing to export OE_A'' of A in exchange for OE_B of B (imports) as they must pay out the amount $E_A E_A'' = ES$ to the government by way of duty. In the diagram we have assumed a tariff rate of 25 %, so that ES is 25 % of SE_B and 20 % of EE_B. In fact, letting $ES = 0.25SE_B$, we have $SE_B = 4ES$. Since $ES + SE_B = EE_B$, by substituting we get $ES + 4ES = EE_B$, whence $ES = 0.20EE_B$.

Since the above reasoning can be applied to any other point of the offer curve OG_2, we conclude that this curve will shift downwards by the same percentage (in our example by 20 %) to position OG_2'.

An alternative way of looking at this shift is to observe that, at the same world prices (terms of trade), the domestic price of imports increases as a consequence of the tariff and this—as we saw in Sect. 10.5.1—reduces both the demand forimports and the supply of exports. Therefore, when the value of the terms of trade is, say, the slope of ray OE, the demand for imports by country 2 will no longer be OE_B but smaller, for example OH_B, and the supply of exports will no longer be OE_A but lower (OH_A). The cum-tariff (or tariff-distorted) offer curve of country 2 must, then, pass through point H. If we repeat this reasoning for all possible terms for trade, we see that the offer curve of country 2 shifts downwards as a consequence of the imposition of a tariff by that country.

The new equilibrium point will be found at the intersection of the OG_1 and OG_2' curves. It is E', where the quantities traded are lower and the terms of trade have shifted in favour of country 2, as can be seen from the fact that ray OE' has a less steep slope than ray OE: in other words, country 2 now gives a smaller amount of commodity A (exports) per unit of B (imports).

To see how the tariff influences country 2's economy we must consider the domestic rather than the world price ratio. To determine the former we must add the absolute value of the unit tariff to the latter. With reference to point E' the total amount of the tariff is $E_1 E'$ as explained above; it follows that consumers' outlay to obtain OE'_B of imports is $OE'_1 = E'_B E_1$ and not OE'_A. Thus the domestic exchange ratio (relative price) will be $E'_B E_1 / OE'_B$, equal to the slope of ray OE_1.

The domestic relative price has increased, *but by a smaller amount than would result from the application of the tariff to the pre-tariff terms of trade*: in fact, the terms of trade have decreased as a consequence of the tariff. The percentage rate of increase of the cum-tariff domestic relative price with respect to the pre-tariff terms of trade can be computed by taking the ratio $E_1 E''_1 / E'_B E''_1$, clearly smaller than the tariff rate $E_1 E' / E'_B E'$.

The increase in the domestic relative price p_B/p_A will make industry B more profitable, so that resources will shift from industry A to industry B and, consequently, the real unit reward of the factor used relatively intensively in industry B will increase (the Stolper-Samuelson theorem).

To sum up, if the offer curves have a normal shape and if we ignore the manner in which the government disposes of the tariff revenue (this will be examined below; see also Sect. 24.1), then:

(a) The imposition of a tariff causes a decrease in the international relative price of the commodity imported, that is, an improvement in the terms of trade of the country that imposes it;
(b) The domestic relative price of imports increases with respect to their after-tariff world relative price;
(c) The improvement in the terms of trade is not such as to offset the tariff, so that the domestic relative price of imports increases with respect to their pre-tariff world relative price, though by a percentage smaller than the tariff rate;
(d) The protected sector becomes more profitable with respect to the sector producing exportables;
(e) Resources will shift towards the protected sector;
(f) The Stolper-Samuelson theorem holds.

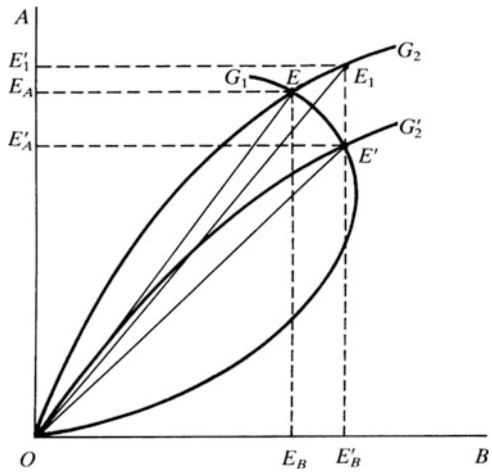
This list does not include the effects of a tariff on social welfare: that we can examine by using social indifference curves; before that (see Sect. 11.1), however, we must complete the foregoing analysis by examining two anomalous cases: *Metzler's case* (already mentioned in Sect. 10.5.1) and *Lerner's case*.

10.5.2.1 The Metzler and Lerner Cases

The *Metzler case* (Metzler, 1949)[7] occurs when the tariff-imposing country's offer curve is normal, whilst that of the other country is anomalous, having a negative (instead of positive) slope in the relevant stretch (as depicted in Fig. 10.5). This means that country 1 is willing to give up decreasing (instead of increasing) amounts of exports in exchange for increasing amounts of imports. It is an anomalous but not impossible case: see Sect. 19.3.1.

[7] Actually, this case is implicitly contained in Lerner (1936).

Fig. 10.5 Tariffs and terms of trade: Metzler's case



When country 2 levies a tariff its offer curve shifts from OG_2 to OG_2' and the point of international equilibrium shifts from E to E'. The terms of trade improve in favour of country 2, but there is more to it than that. Given the anomalous shape of the OG_1 curve, the terms of trade improve so much that the cum-tariff domestic relative price (p_B/p_A) in country 2 is *smaller* than the pre-tariff (international and domestic) relative price. This can be seen from the fact that the slope of ray OE_1, though steeper than the slope of ray OE', is lower than that of OE. The sector producing the importable commodity (B), far from being protected by the tariff, will be harmed by it.

In such a situation it is sector A which becomes more profitable, resources will shift from B to A, and the factor used relatively intensively in A will see an increase in its real unit reward (generalized Stolper-Samuelson theorem).

In the foregoing treatment we have seen that the imposition of a tariff in any case improves the terms of trade in favour of the country which imposes it. This result, however, is by no means generally valid, for cases are possible in which the terms of trade do not change or even move against the tariff-imposing country. These cases can also be attributed to anomalous shapes of the offer curves: if, for example, country 1's offer curve is a straight line through the origin, then any tariff-induced shift in country 2's curve cannot influence the terms of trade, which will in any case coincide with the given slope of country 1's offer curve. More interesting is the case in which the terms of trade move against the tariff-imposing country (*Lerner's case*: see Lerner, 1936).

To examine this case we must first establish how the government disposes of the tariff revenue (in the previous analysis we have explicitly neglected the effects of this). In general the tariff revenue can be disposed of by the government in various ways: it can be redistributed to consumers, or spent entirely on the importable, or spent entirely on the exportable, or spent on a combination of the two commodities. Amongst the various possibilities we consider here the case

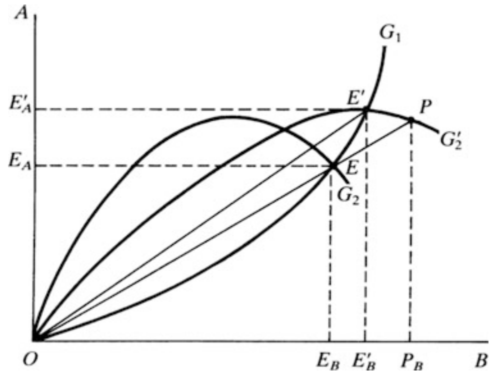


Fig. 10.6 Tariffs and terms of trade: Lerner's case

in which this revenue is spent entirely on imports (the reader interested in the taxonomy of the effects of all possible cases can consult Lerner (1936), Metzler (1949), and Chacholiades (1978, chap. 18); see also Sect. 24.1).

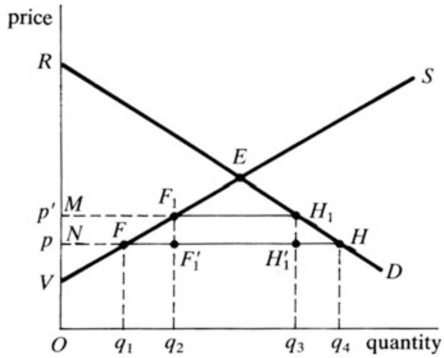
It can be seen intuitively that, if the domestic demand for this commodity is rigid (so that price changes affect it very little or not at all), the additional demand on the international market for the importable will cause an increase in the world price of this commodity, that is a worsening of the terms of trade of the tariff-imposing country.

This is shown in Fig. 10.6, where the offer curve of country 2 is anomalous (this shape is a possible occurrence, as was shown in Sect. 19.3.1), and the initial international equilibrium (point E) lies in its downwards sloping part. The imposition of a tariff causes OG_2 to shift to OG'_2. Observe that, unlike in the previous diagrams, OG'_2 cuts OG_2 at a point to the left of E so as to lie to the right of OG_2 itself along the ray whose slope represents the pre-tariff terms of trade (ray OEP). This follows from the two assumptions made above (rigidity of the domestic demand for the importable and expenditure of the whole tariff revenue on the importable itself), which imply that at the given pre-tariff terms of trade there is a world excess demand for commodity B, so that point P must lie to the right of E along the terms-of-trade ray. Only in this case, in fact, at the given terms of trade, is the demand for commodity B by country 2 greater than the supply of the same commodity by country 1, as $OP_B > OE_B$. After these preliminaries, it can immediately be seen that at the new international equilibrium point E' the world relative price p_B/p_A obtaining after the tariff, has increased (slope of OE' > slope of OE): the cum-tariff terms of trade have moved against (the tariff-imposing) country 2.

10.6 Quotas and Other Non-tariff Barriers

From the theoretical point of view there are numerous impediments to free trade other than tariffs; as stated in Sect. 10.1, these impediments are taking on an ever increasing practical importance, so that they deserve something more than a cursory mention. Some of these impediments have a consolidated theory and practice behind

Fig. 10.7 Effects of a quota



them, but new types, not previously envisaged, are being introduced in practice, so that an exhaustive list would contain dozens. Therefore we shall concentrate on the main traditional types (quotas, export duties, etc.) and give a necessarily brief treatment of some of the others, referring the reader to Baldwin (1971). The relevance of the "new" protectionism, based on non-tariff barriers, will be assessed in Chap. 12.

10.6.1 Quotas

An (import) quota is a quantitative restriction (so many cars of a certain type per unit of time) imposed by the government on the imports of a certain commodity and, therefore, belongs to the category of direct controls on international trade. For this purpose the government usually issues import licences (which it can distribute to importers according to various criteria) but other forms are possible.

The effects of a quota[8] can be analysed by means of a diagram similar to that used in Sect. 10.3 (see Fig. 10.1) to analyse the effects of a tariff. In Fig. 10.7, p is the world price of the commodity, of which a quantity q_1q_4 is imported under free trade. The government now decides that imports have to be reduced, for example from q_1q_4 to q_2q_3 and, accordingly, decrees a quota. The domestic price of the commodity will rise to p', since the (unsatisfied) excess demand by domestic consumers will drive it up from p to the level at which the actual excess demand is exactly equal to the given quota, $F_1H_1 = q_2q_3$.

The effects of a quota on domestic price, output, consumption, and on imports, are the same as those which would occur if a tariff were imposed such as to cause an increase in the domestic price from p to p': this can be readily seen by comparing Fig. 10.7 with Fig. 10.1. The *equivalent tariff rate* can be computed from

[8]For brevity's sake we shall examine the effects of a quota exclusively in a partial equilibrium context and under the assumption that the world price does not change.

the equation $p' = (1 + d_1)p$. Some authors (for instance Corden, 1971a, p. 213) call this the implicit tariff rate. However, since this term is also used in the sense of effective rate of protection (see below, Sect. 10.7), to avoid confusion we do not use it here.

There is, however, a difference between a quota and an equivalent tariff: whilst in the case of a tariff the government collects a fiscal revenue ($F_1F_1'H_1'H_1$ in Fig. 10.1), it now collects nothing and the quota gives rise to a gain of equal size ($F_1F_1'H_1'H_1$ in Fig. 10.7) accruing to the quota holders (this is true under assumption that the country is small and that there is perfect competition among the foreign exporters. In the opposite case, these could avail themselves of the occasion of the quota to raise the price charged to domestic importers, thus depriving them of part of the gain under consideration). Now, why should the government deprive itself of a fiscal revenue if the same quantitative restriction and the same effects of a quota can be obtained by a tariff?

Let us first observe that, in principle, the government could *sell the import licences by auction*[9]: with a perfect auction in a perfectly competitive market, the revenue of the auction would be exactly the same as that of the equivalent tariff. This is so because competition between importers to get hold of the licences will induce them to make higher and higher bids until extra profits (which are equal to $F_1F_1'H_1'H_1$) disappear in favour of the government. But this is a theoretical possibility difficult to realize in practice.

The answer to the above question can be found in the fact that only a quota gives the certainty of the desired quantitative restriction on imports, which is lacking in the case of a tariff for various theoretical and practical reasons, among which (for a complete treatment see Takacs, 1978):

1. The equivalence of the effects on imports depends on the existence of perfectly competitive conditions at home and abroad: in the opposite case, in fact, the effects of a tariff and of a quota can be very different. For example, if foreign exporters do not operate under perfect competition, they may reduce the price in order not to lose market shares when the home country imposes a tariff, so that the increase in the domestic price will be smaller than that required to achieve the desired reduction in imports.
2. A quota, unlike a tariff, can have important effects on the market structure of the country which imposes it, for it can convert a *potential* into an *actual monopoly*, that is, enable the domestic industry, fully protected from foreign competition by the quota, to establish a monopoly. In fact, let us assume that in the country there is a potentially monopolistic industry. In the presence of a tariff, this industry cannot raise the price above the world price plus tariff, for its sales would drop to zero (domestic consumers will buy solely imported goods if the domestic

[9]The auction is only one method of issuing licences to importers. Another is the first-come, first-served basis, still another is the subdivision of the licences among importers in proportion to the quantities imported by each before the introduction of the quota. But it is clear that only by a perfect auction the government's revenue will be the same as that of an equivalent tariff.

importable has a price higher than the world price plus tariff). If instead of the tariff the country decrees a quota, the potential monopoly can become an actual one, because the domestic industry can now raise the price without danger of its sales dropping to zero, as imports cannot exceed the quota.

3. The computation of the tariff (d_1 in our example) which brings about exactly the desired reduction in imports can be made only if the curves D and S are known exactly and do not shift unpredictably. Notwithstanding the advances in econometrics, these curves can be determined only within a (usually large) confidence interval. Furthermore, the possibility of (large though predictable) shifts in these curves (because the underlying exogenous factors change in a known way) compels the government to compute, levy and enforce changing tariff rates.

10.6.2 Export Duties

In addition to taxing imports, it is also possible to tax exports (customs duties on exports). For brevity's sake we limit the analysis to the partial equilibrium context and assume that the export duty has no consequence at all on the world price of the commodity.

In Fig. 10.8 the usual partial equilibrium demand and supply curves are drawn. As we are dealing with an exportable commodity, we must consider the part of the diagram above the autarky equilibrium point, where excess supply is present. Let us assume that the initial free trade price is OM: the supply of exports (domestic excess supply of the commodity) is $FH = q_1q_4$. The levying of an export duty, say MN, causes a decrease in the domestic price from OM to ON. Domestic producers, in fact, by selling the commodity abroad at the given world price OM, eventually receive only ON per unit of the commodity, as they must pay out MN to the government by way of duty. Therefore the price on which domestic firms base their output calculations is ON. From the dynamic point of view the imposition of an export duty induces domestic firms to shift their supply from the foreign to the domestic market, where in the moments immediately after the levying of the duty, the price is the same as before. But this greater supply on the domestic market causes a decrease in the domestic price; the decrease will continue until the price has fallen to ON. When the domestic price is ON, the domestic supply is lower whilst demand is higher with respect to OM: the result is a contraction of exports from FH to $F_1H_1 = q_2q_3$.

Since the domestic price is lower, domestic consumers will benefit, whilst domestic producers will lose. Benefits and costs can be calculated by using the concepts explained in Sect. 10.3 in the case of an import duty.

Consumers' surplus here increases by the area MNF_1F and *producers' surplus* decreases by the area MNH_{1H}; the government collects a fiscal revenue (by way of export duty) measured by the area $F_1'F_1H_1H_1'$. Therefore the area of the triangles FF_1F_1' and $H_1'H_{1H}$ remains unaccounted for, and represents the *social cost* of the

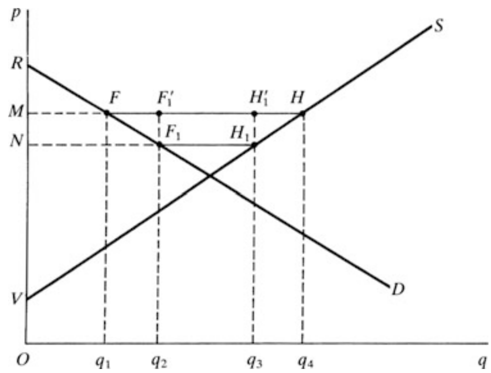


Fig. 10.8 Effects of an export duty

duty. A further symmetry can be found in the relation between the pre-duty and post-duty relative price at home and abroad.

Let p_m, p_x denote the domestic price of importables and exportables, and p_{mw}, p_{xw} the respective world prices. In the absence of duties, $p_m = p_{mw}$, $p_x = p_{xw}$. Suppose now that an import duty is imposed at the rate d: the domestic price of importables becomes $p_m = (1 + d)p_{mw}$ whilst the domestic price of exportables remains equal to the world price; therefore the domestic relative price becomes $p_x/p_m = p_{xw}/(1 + d)p_{mw}$. In the case of an export duty at the same rate d, the relation between the domestic and world price of exportables is $p_{xw} = (1 + d)p_x$, whilst the domestic price of importables remains equal to the world price; the domestic relative price becomes $p_x/p_m = [p_{xw}/(1 + d)]/p_{mw}$, which is algebraically equal to that found in the case of an import duty.

The analogy of the results concerning the social cost of an import and an export duty induced some authors (for example Lerner, 1936) to talk of a "symmetry" between these two types of duty.

The analysis has so far been based on the assumption that the country under consideration has no monopolistic power (in a broad sense) on the international market. In the opposite case it would be possible to use the export duty to exploit this power to the national advantage, as part of the duty would be charged to the rest of the world by way of an increase in the world price. It is important to note that the monopolistic power can be increased by an agreement among exporting countries which form an *international cartel*. This is the subject of Sect. 10.6.3.

10.6.3 International Cartels

An international cartel (Caves, 1979) consists of a group of producers of a certain commodity located in various countries who agree to restrict competition among themselves (in matters of markets, price, terms of sale etc.).

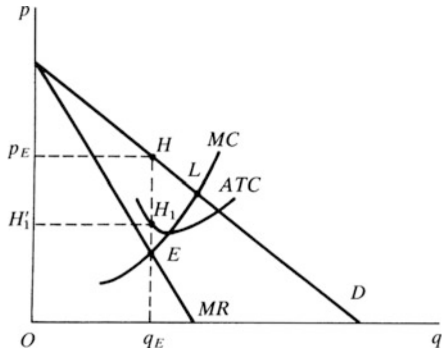


Fig. 10.9 The monopolistic cartel

We shall be mainly concerned with cartels aimed at the control of the world price of the commodity by fixing a common price. The agreement is often at the level of governments (the typical example is OPEC, for which see below), but agreements among private producers are also possible (examples are the agreement among the main international firms trading in tobacco in the 1880s and, in the same period, the cartel concerning the level of railway fares).

If the cartel includes the total number of producers, a full monopoly comes into being, to which the well-known principles of monopoly theory can be applied. In such a situation, given the world demand curve for the cartelized commodity, the price which maximizes the cartel's profits is obtained by reading off the demand curve the price corresponding to the quantity determined by the intersection of the marginal cost curve MC and the marginal revenue curve MR. In Fig. 10.9, the price is p_E and the quantity sold q_E (in a competitive market, on the contrary, in the short run, price and quantity would be determined in correspondence to point L); given the average total cost curve ATC, the profit will be $H_1'H_1Hp_E$. We also recall from microeconomic theory that the monopolist's markup, namely the proportional excess of price over marginal cost, is given by the reciprocal of the price elasticity of demand (η_w):

$$\frac{p_E - MC}{p_E} = \frac{1}{\eta_w}, \tag{10.3}$$

so that the more rigid the world demand, the higher the cartel's markup. So far we have implicitly assumed that the cartel behaves as a single entity, but even in this case the problem arises of apportioning the production of the commodity among the members. In an ideal cartel the various members can be considered as the various plants of a single monopolist, so that we can apply the theory of the multiple-plant monopolist. This tells us that the optimum allocation is that in which the marginal cost in each plant is the same and equal to the marginal revenue of output as a whole. To see this, assume that MC of member i is greater than that of member j. It is then possible to decrease the cartel's total cost of producing the same total output by marginally decreasing member i's output and marginally increasing

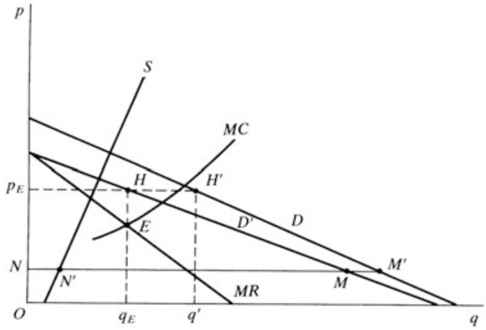


Fig. 10.10 A quasi-monopolistic cartel

(by the same amount) member j's output: in fact, the decrease in total cost (MC_i) is greater than the increase (MC_j). This process continues up to the point where $MC_i = MC_j$. Once the marginal cost has been equalized everywhere for any given output, thus determining the minimum total cost of the cartel, the maximum profit will be as usual determined by equating marginal revenue of output as a whole to the (common) marginal cost of the various producers.

This ideal allocation is not, however, easily realized in practice. In the real world the production is apportioned on the basis of negotiations among the members of the cartel, each of whom has its own interests and different contractual force. The more influential and skilful negotiators will probably get a greater quota than the optimum corresponding to the application of the principle of equalization of marginal costs, even if this will raise the cartel's total cost of production.

We must now consider the more realistic case in which the cartel does not include all but only part of the producers, so that besides the cartel, also independent (i.e., not belonging to the cartel) competitive producers are present in the world market for the commodity. These latter will have to accept the price fixed by the cartel, but the cartel will have to take their supply into account when fixing the price. The market form obtaining here is *quasi-monopoly* (see Henderson and Quandt, 1980). In Fig. 10.10, in addition to the world demand curve D, we have drawn the aggregate supply curve S of the independent producers. If we subtract, for any given price, S from D laterally, we obtain D', which is the demand curve for the cartel's output. For example, at price ON, the supply of the independent producers is NN': if we subtract MM' (equal, by construction, to NN') from world demand NM', we obtain segment NM, that is the quantity that the cartel can sell at price ON.

Once the D' curve has been derived, the cartel can behave along it as a monopolist and will maximize profits by the usual rule, that is, by equating marginal cost to marginal revenue (the latter will, of course, be that concerning curve D'). The cartel, therefore, will fix the price at Op_E and sell a quantity $p_{EH} = Oq_E$, whilst the independent producers will sell a quantity $HH' = q'_{Eq}$.

One can easily check, by drawing the marginal revenue curve concerning curve D (which we leave as an exercise for the reader), that the price is lower and the quantity sold greater than in the case of a monopolistic cartel. It is also possible to check graphically that the greater the elasticity of the supply curve of independent producers S, the smaller the cartel's markup. More precisely, as shown in Sect. 24.2, the (price) elasticity of the D' curve (denoted by η_c) depends on the elasticity of the D curve (η_w), the elasticity of the S curve (η_s), and on the cartel's share in the total consumption of the commodity (k), according to the formula

$$\eta_c = \frac{\eta_w + (1-k)\,\eta_s}{k}. \tag{10.4}$$

Consequently, the cartel's markup is

$$\frac{1}{\eta_c} = \frac{k}{\eta_w + (1-k)\,\eta_s}. \tag{10.5}$$

From Eq. (10.5) we can readily derive the conditions for the success of a cartel, as measured by the capability of imposing a substantial markup and so reaping high monopolistic profits. These are:

(a) A low elasticity of total world demand (a small η_w);
(b) A low elasticity of independent producers' supply (a small η_s);
(c) A high cartel share in the world market for the commodity (a high k: for $k = 1$, Eq. (10.5) reduces to (10.3)).

These are the purely economic conditions, to which a further condition must be added, namely

(d) The members of the cartel must accept and adhere to the official decisions taken by the cartel (by means of majority voting or some other way) as regards price and output.

Condition (d) is essential for the life itself of the cartel. If, in fact, the members begin to decartelize by selling greater amounts (than those allotted to each) at lower prices, the cartel will soon break up. But why should there be any incentive to behave in this manner? The answer is that, though the profits of the cartel as a whole are maximized by respecting the official decisions, the *single* member can obtain vastly greater profits by slightly lowering the price below the official one, *provided that* the other members adhere to the official price. In fact, buyers will be willing to buy all the quantity demanded—previously bought from the cartel—from the single producer who charges a slightly lower price, so that the demand curve facing this single producer is in practice almost perfectly elastic. This producer will therefore realize increasing profits by increasing output, because his selling price is greater than his marginal cost,[10] and he can sell increasing amounts without further reducing

[10]We must remember that in the initial situation the official price fixed by the cartel is higher than the marginal cost (this is true in both the monopolistic and quasi-monopolistic cartel). From the point of view of the cartel as a whole, it is not profitable to reduce the price (this, in fact, would lead to lower profits), whilst the single member can—for the motives explained in the

the price. He will therefore profit from increasing output up to the point where his marginal cost has increased to the level of the selling price charged by him.

Naturally, greater profits for the single producer who does not adhere to the official price mean lower profits for the other cartel-abiding members, but the single producer, especially if relatively small, can always hope that the other members will not become aware of his infringement or will not react. If, for example, his share in the cartel's output is 1 %, he may think that a 50 % increase in his output (this means that his share goes up to 1.5 %) will cause so small a loss (spread out through all the other members) as to be negligible. This is undoubtedly true, but if the same idea occurs to a sufficient number of members and is put into practice by them, the cartel dissolves. Therefore the cartel, to persist, must be able to put pressure (of an economic or political or some other nature) on the single members to make them adhere to the official decisions.

But unfaithful members are not the only cause of the dissolution of a cartel. There are at least three other motives leading to a progressive erosion of the markup (and so of the profits) of the cartel. They can be analysed with reference to formula (10.5) and are:

1. The increase in η_w. Even if world demand is sufficiently rigid when the cartel is set up, the very success of the cartel, paradoxically, helps to make this demand more elastic. As a consequence of the (usually very large) price increase, buyers will put their every effort into the search for substitutes for the cartelized commodity (it suffices to mention the search for energy sources alternative to oil and the research into energy-saving production processes and commodities that were set into motion as a consequence of the Organization of Petroleum Exporting Countries—OPEC—cartel) and so η_w increases.
2. The increase in η_s. Even if independent producers' supply is rigid when the cartel is set up, the success itself of the cartel, again, helps to make this supply more elastic, since these producers will multiply their efforts to increase output. If the cartel concerns an agricultural commodity, such as sugar or coffee, the price increase will induce independent producers to shift increasing amounts of resources (land, labour, capital) to the production of the cartelized commodity. If an exhaustible natural resource is concerned, such as oil or copper, independent producers will multiply their efforts to find new fields. Similar efforts will also come from countries previously not exploiting the resource, these efforts, if successful, will increase not only the output but also the number of independent producers (think of the oil fields found by England under the North-Sea). All this causes an increase in η_s.
3. The decrease in k. In order to increase the price without building up excessive inventories of the commodity, the cartel must restrict output and sales relative to the pre-cartel situation. This, coupled with the efforts of independent producers (point 2), leads to a decrease in k.

---

text—obtain higher profits by slightly lowering his selling price below the official one; this lower price is nevertheless higher than his marginal cost.

These three forces jointly operate to erode the cartel's monopolistic power. Also, note that as the markup is wearing away, the incentive for the single members to decartelize (see above) becomes greater and greater.

Economic theory, therefore, predicts that, in the long run, any cartel is bound to dissolve, even if new cartels are always being set up, so that at any moment a certain number of cartels is in existence. Historical experience seems to confirm this conclusion, even in the most dramatic cases. Among these one must undoubtedly count the cartel which gathers the main oil producing countries into OPEC. Conditions (a), (b) and (c) above certainly held in 1973: very rigid world demand for oil, low elasticity of the supply of independent producers, high share (above 50 %) in world production controlled by the cartel. Furthermore, for various political motives, the degree of cohesion of the cartel was high.

The great initial success of OPEC is, therefore, not surprising. However, forces (1), (2) and (3), slowly but steadily got down to work.

The high price of oil set into motion or intensified the search for alternative energy sources, for productive processes less intensive in energy, for less energy-consuming commodities and ways of life (energy-saving cars, limits to domestic heating, better insulation of new buildings, etc.) began or was intensified. As a consequence, the share of oil in world energy consumption decreased, and energy consumption per unit of real GDP fell in industrial countries as a whole.

Another element that reinforced the drop in demand for oil was the world depression which, by slowing down (and sometimes by causing a decrease in) the level of activity in the various industrialized countries, reduced their energy needs. The supply of independent producers steadily increased (the case of England, which became a net exporter of oil, is sensational). The cartel's share in the world market decreased well below 50 %.

As a consequence of all this, cases of members not adhering to the cartel's official decisions were not lacking, often not because of greed, but out of sheer necessity (many OPEC countries had set up development programs based on estimates of an increasing—or at least not decreasing—flow of oil revenues in real terms, and found themselves in trouble when this flow started to decrease).

Alternative explanation of OPEC's behaviour (based on game theory or on coalition-formation theory) also exist. See, for example, Razawi (1984) and McMillan (1989, chap. 6).

10.6.4 Other Impediments to Free Trade

We give here a (by no means exhaustive) list of other impediments to free trade with a brief description of each. A more in-depth treatment will be given in Chap. 12.

(a) *Export Subsidies.* In general, they may take various covert forms besides the overt one of a direct payment by the government to the exporter (usually in proportion to the volume of exports). Examples of covert subsidies are: more favourable credit conditions (the difference between these and the normal conditions applied to producers for the home market is paid by the government); insurance of certain risks (for example, that the foreign importer defaults) paid by the government; promotional activities (such as trade fairs, advertising, etc.) organized by public agencies. Export subsidies are usually considered legitimate when they are a rebate of the tariff paid by the exporting industry on imported inputs.

(b) *"Voluntary" Export Restraints (VER)* and *Import Expansion (VIE)*. In the case of a VER, the exporting country "voluntarily" curtails exports to the importing country. In the case of a VIE, the importing country "voluntarily" increases its imports from the exporting country. It is, of course, a relative "voluntarity", for it is negotiated between the importing and the exporting country as an alternative to traditional measures such as tariffs or quotas.

(c) *Production Subsidies*. If the government subsidizes the domestic production of a commodity, this subsidy automatically becomes an export subsidy as regards the exported part of the output, or a subsidy to the importables sector if the commodity is an importable.

Box 10.2 Regulatory Protectionism

As conventional trade barriers decline, there is growing concern that countries are resorting to technical regulations to protect domestic producers (Technical Barriers to Trade-TBTs). These barriers, resulting from national regulations and standards on product safety, testing, labelling, packaging, certification, labour and environmental standards, have proliferated in recent years. Since these regulations can be wielded for protectionist ends, their proliferation has led to widespread complaints of regulatory protectionism. TBTs result from norms that control the sale of goods in a particular market. There are two distinct aspects of this control: contents of the norm and testing procedures necessary to demonstrate that a product complies with the norm. Content-of-norm or, generally speaking, regulatory differences between countries, can be broadly classified as *horizontal* or *vertical*. Horizontal norm involve, for example, imposing different technologies as certain plug forms for appliances. With vertical standards a regulator insists that goods achieve at least certain minimum standard of safety or performance (for example, that cars do not exceed certain maximum levels of emissions). Product norms and testing procedures can distort trade when they increase foreign firm's costs relative to those of domestic firms. Of course, the major problem with the economic assessment of TBTs is that they are potentially much more complicated to analyse than tariffs or quotas: the main problem with TBTs is that it is difficult to ascertain whether a certain norm serves the citizens' interests or protectionist interests. The problem that different setting of regulations by EU governments might be hampering trade and competition has been a major reason for the institution of the Single Market program, and mutual recognition agreements have been agreed between the EU and several other countries. A similar rationale underlies the articles on Technical Barriers to Trade and Sanitary and Phytosanitary Standards in the WTO Agreement from the Uruguay Round.

For example, the *Technical Barriers to Trade Agreement* tries to ensure that regulations standards and testing and certification procedures do not create unnecessary obstacles. The agreement recognizes the countries' right to adopt the standards they consider appropriate (for human, animal or plant life or health, for the protection of the environment)

but discourages any methods that would give domestically produced goods an unfair advantage.
A separate agreement on food safety and animal and plant health standard (the *Sanitary and Phitosanitary Measures Agreement*) sets out the basic rules. It allows countries to set their own standards. But it also says that regulations must be based on science and should be applied only to the extent necessary to protect human, animal or plant life or health, but should not arbitrarily or unjustifiably discriminate between countries where identical or similar conditions prevail.
Notwithstanding these agreements there is of course considerable resistance by many countries to conform their policies to trade treaties and to recognize each other's rules and procedures. Liberalization of TBTs often entails preferential arrangements between rich countries, creating a two-tier system of market access with developing countries in the second tier.

(d) *Tied Aid*. Developed countries often grant financial assistance to developing countries with the constraint that the recipient spends the sum received to purchase commodities from the donor. This causes distortions, which are all the greater when the price (and/or other conditions) in the donor country is not the cheapest.
(e) *Advance-Deposit Requirements*. Importers are required to deposit funds (in the central bank, in a commercial bank, etc.) in an amount proportional to the value of the imported commodities, with no interest and for a given period of time (usually prior to the receipt of the commodities). Thus importers are burdened with an additional cost, which depends on the percentage of the value of imports, on the length of the period and on the rate of interest (which measures a direct cost, if the importer has to borrow the funds, or an opportunity cost, if he owns them). The advance deposit is equivalent to a tariff with a rate that can be easily computed: if, for example, the rate of interest is 10 % per annum, the period of time is 3 months and the percentage of the value of imports is 80 %, then the equivalent tariff rate is 2 %. In fact, the rate of interest per quarter is 2.5 % (10 % : 4), and since the importer must deposit 0.8 dollars per dollar of imports, the additional cost is $0.8 \times 2.5\,\% = 0.02$ dollars per dollar of imports, which is equivalent to an ad valorem tariff with a 2 % rate.
(f) *Government Procurement*. Governments buy a large amount of goods and services, and usually prefer to buy domestic rather than equivalent foreign goods of the same price (in some cases they are allowed by domestic legislation to buy domestic goods even if equivalent foreign goods have a lower price, not below a certain percentage); besides, governments may have recourse to a series of techniques aimed at limiting the opportunity for foreign producers to tender for the supply of goods to the public sector. All this amounts to a discrimination in favour of domestic producers, which restricts imports.
(g) *Formalities of Customs Clearance*. These are connected with the imposition of tariffs, such as the classification and evaluation of the commodities in transit at the customs and other bureaucratic formalities. A more rigid application of these formalities hinders trade and involves a cost for importers.

(h) *Technical, Safety, Health and Other Regulations* (so-called *regulatory protectionism*, see box). Countries often have different regulations, and this is in itself an impediment to international trade, for producers have to bear additional costs to make the commodities conform to the different regulations, according to the country of destination. Besides, a country may use these regulations to reduce or even stop the imports of certain commodities from certain countries, for example, by checking with particular meticulousness and slowness their conformity to the regulations, or even by issuing regulations which actually prevent the acceptance of certain foreign commodities (an example is the case of the United States, which in the past sometimes drew up health regulations in such a way that Argentinian beef could not possibly comply with them).

(i) *Border Tax Adjustments*. Governments usually levy an "import equalization tax" on imported goods equal to the indirect tax levied at home on similar goods domestically produced and, vice versa, they give back to exporters the national indirect tax. This may cause distortions if the import equalization tax is higher than the national indirect tax (the difference is a covert import duty) or if the sum returned to exporters is greater than the amount of the national indirect tax (the difference is a covert export subsidy).

(j) *Embargo*. The government of a country decrees that certain commodities must not be exported to certain countries. This is usually done for motives concerning foreign policy, for instance to prevent (actual or potential) enemy countries from having access to advanced technologies or to put political pressure on them by economic means.

(k) *State Trading*. The government of a country takes all of the country's international trade upon itself. This is by itself a non-tariff barrier, for the government (directly or indirectly) has a monopolistic-monopsonistic power as the one and only supplier of domestic goods to foreign markets and the one and only buyer of foreign goods for the domestic market. If, in addition, the country has a planned economy, the determination itself of the commodities to be exported and imported, of the relative amounts, of their prices, etc., is outside the scope of the pure theory of international trade dealt with in this book, but falls within the field of the theory of planning, which is not treated in the present work.

10.7 Intermediate Goods and Effective Rate of Protection

So far, the models used consider solely tariffs on final goods, given the assumption that production takes place by making exclusive use of internationally immobile primary factors (capital and labour). Actually, however, production also requires *intermediate inputs* (raw materials, semi-finished goods, etc.) which can be, and normally are, internationally traded. This has led to the elaboration of the concept of effective rate of protection (or implicit tariff), defined as the percentage increase in the value added per unit in a specific economic activity, made possible by the tariff structure (but other things—including the exchange rate—being equal) with respect to the situation without tariffs.

The basic idea is that, when intermediate goods are brought into the picture, the *nominal* tariff on a certain commodity (which is applied to the price of the commodity) may be quite different from the *implicit* tariff or *effective* rate of protection provided for the economic activity which produces the commodity in question, namely for the value added (which is distributed to the primary factors of production) in the production of the commodity. Nominal tariffs are applied to commodities, but factors move between economic activities; thus, in order to find the effects on resource allocation of a tariff structure, one must calculate the rate of protection given to each activity, that is the effective rate of protection. Besides, if the aim is to protect a certain sector, and since what is relevant for an industry is—*ceteris paribus*—its value added, the true or effective protection is that which gives rise to an increase in value added.

Given the definition, it is intuitive that the effective rate of protection depends not only on the tariff on the commodity under consideration, but also on the inputs of intermediate goods and on the tariffs on these. The usual way of incorporating intermediate goods in the analysis of tariffs is to use an input-output model, in which these goods are input according to fixed (given and constant) technical coefficients. It is also assumed that the world price of imports remains the same.

Here we examine the simplified case of a single intermediate good; the general case will be examined in Sect. 24.3. Suppose, for example, that to produce one unit of cloth 1.5 units of yarn are required, that the unit prices in the absence of tariffs are 100 and 50 for cloth and yarn respectively, that both goods are importables. Given the assumption of constant fixed technical coefficients, we can consider the unit value added, which is 25, that is, the difference between 100 (value of a unit of the final good, cloth) and 75 (value of 1.5 units of the intermediate good, yarn). Let us now introduce a tariff on both cloth and yarn, with rates of 40 and 20 % respectively: the domestic prices go up to 140 and 60 respectively and the cum-tariff value added is $50 = 140 - 90$. This represents a 100 % increase in the pre-tariff value added $[(50 - 25)/25]$. Thus the tariff structure has provided an effective protection with a rate of 100 % to the domestic industry producing cloth.

This numerical example can be transformed into a general formula by using simple algebra. Let us define the following symbols:

$$\begin{aligned}
v_j &= \text{unit value added in activity } j \text{ without tariffs,}\\
v'_j &= \text{unit value added in activity } j \text{ with tariffs,}\\
q_{ij} &= \text{technical coefficient in physical terms (quantity of the}\\
&\quad\ \text{intermediate good } i \text{ input in one unit of the final good } j\text{),}\\
&\quad\ \text{assumed fixed and constant,}\\
a_{ij} &= \text{share of } i \text{ in the value of } j \text{ at free trade prices,}\\
d_j &= \text{nominal tariff rate on goods } j\text{,}\\
d_i &= \text{nominal tariff rate on good } i\text{,}\\
p_j, p_i &= \text{prices.}
\end{aligned}$$

The unit value added without tariffs is

$$v_j = p_j - p_i q_{ij}. \tag{10.6}$$

Since a_{ij} and q_{ij} are, by definition, related by

$$a_{ij} = p_i q_{ij}/p_j \text{ whence } p_i q_{ij} = p_j a_{ij}, \tag{10.7}$$

we can rewrite (10.6) as

$$v_j = p_j - p_j a_{ij} = p_j \left(1 - a_{ij}\right). \tag{10.8}$$

After the imposition of the tariffs, the unit value added becomes

$$\begin{aligned} v'_j &= \left(1 + d_j\right) p_j - (1 + d_i)\, p_i q_{ij} = \left(1 + d_j\right) p_j - (1 + d_i)\, p_j a_{ij} \\ &= p_j \left[\left(1 + d_j\right) - (1 + d_i)\, a_{ij}\right]. \end{aligned} \tag{10.9}$$

The effective rate of protection is defined by

$$g_j = \frac{v'_j - v_j}{v_j}, \tag{10.10}$$

so that, by substituting v_j and v'_j from Eqs. (10.8) and (10.9) we get

$$\begin{aligned} g_j &= \frac{p_j \left[\left(1 + d_j\right) - (1 + d_i)\, a_{ij}\right] - p_j \left(1 - a_{ij}\right)}{p_j \left(1 - a_{ij}\right)} \\ &= \frac{\left[\left(1 + d_j\right) - (1 + d_i)\, a_{ij}\right] - \left(1 - a_{ij}\right)}{\left(1 - a_{ij}\right)} = \frac{d_j - d_i a_{ij}}{1 - a_{ij}}. \end{aligned} \tag{10.11}$$

To transform this expression into a mathematical equivalent which is more illuminating from the economic point of view, we add to and subtract the same quantity $d_j a_{ij}$ from the numerator of the last fraction, whence

$$\begin{aligned} g_j &= \frac{d_j - d_j a_{ij} + d_j a_{ij} - d_i a_{ij}}{1 - a_{ij}} = \frac{d_j \left(1 - a_{ij}\right) + \left(d_j - d_i\right) a_{ij}}{1 - a_{ij}} \\ &= d_j + \left(d_j - d_i\right) \frac{a_{ij}}{1 - a_{ij}}. \end{aligned} \tag{10.12}$$

Since the value of intermediate goods must in general be smaller than the value of output, ay must be smaller than one and so $1 - a_{ij} > 0$. From Eq. (10.12) we can then see that $g_j \gtreqless d_j$ according as $d_j \gtreqless d_i$ that is, *the effective rate of protection is greater than, equal to, or smaller than the nominal tariff rate on the final good according as the latter rate is greater than, equal to, or smaller than the nominal tariff rate on the intermediate good.*

Table 10.1 Nominal and effective rates of protection (per cent)

	United States		Japan		Korea	
SECTOR	NRP	ERP	NRP	ERP	NRP	ERP
Agriculture	1.80	1.91	18.40	21.40	72.3	85.7
Food products	4.70	10.16	25.40	50.31	11.7	−27.6
Wearing apparel	22.70	43.30	13.80	42.20	29.0	93.8
Wood products	1.70	1.72	0.30	−30.59	8.6	6.5
Chemicals	2.40	3.66	4.80	6.39	28.5	50.9
Iron and steel	3.60	6.18	2.80	4.34	12.9	31.5
Electrical machinery	4.40	6.34	4.30	6.73	26.2	44.8
Transport equipment	2.50	1.94	1.50	0.03	31.9	12.4

In the numerical example, we have illustrated the case in which $g_j > d_j$. Let us now assume that, other things being equal, the tariff rate on the intermediate good is 50 % instead of 20 %. The new value added is 27.5 (in terms of the symbols defined above we have $p_j = 100$; $q_{ij} = 1.5$; $p_i = 50$; $a_{ij} = 0.75$; $v_j = 25$; $d_j = 0.40$; $d_i = 0.50$; $v'_j = 27.5$) and so the effective rate of protection is 10 %, $[(27.5 - 25)/25]$, smaller than the nominal tariff rate on the final good.

In the case in which the conditions are fulfilled which make the effective rate lower than the nominal rate, the *effective rate may even be negative*. In the last example, change the nominal rate on yarn to 60 %: the new value added is now $140 - 120 = 20$ and the effective rate of protection is negative. What happened is that the tariff structure caused an increase in the price of the intermediate good (which, it should be remembered, represents a cost for the firm) so much greater than the increase in the price of the final good, that the industry producing this is in a worse situation than before the tariff was imposed.

Several empirical studies (see, for example, Yeats, 1974. More recent studies are reported in Table 10.1) have been carried out to calculate effective rates of protection.[11] In Table 10.1 we give the results obtained by Deardoff and Stern (1984) for the United States and Japan, and by Yoo (1993) for the Republic of Korea. From this table we see that the effective tariff rate (ERP) was greater than the nominal one (NRP) in most sectors; in a few cases, however, it was smaller or even negative. Negative effective rates are not a mere theoretical curiosity.

10.8 Imperfect Competition and Trade Policy

The new explanations for international trade introduce new arguments in the old debate on free trade versus protectionism (see Chap. 11). These new arguments, however, instead of leading the debate towards a conclusion, have complicated it

[11]To perform those calculations, the general formula derived in the appendix has to be used, and adjustments have to be made for the fact that prices of commodities include other taxes besides tariffs.

further. The traditional theory had a set of precise results on the preferability of trade to autarky and, if we exclude second-best situations, on the preferability of free trade to restricted trade. The new theories, conversely, give rise to contradictory results: the reason is due to competing assumptions (Markusen & Venables, 1988). We shall touch on these arguments below; for extensive surveys see Baldwin (1992) and Pomfret (1989). See also Puga and Venables (1997) for a study of preferential trading arrangements in the context of an imperfectly competitive environment.

We now come to *strategic* trade policies. The adjective *strategic* hints to the presence of some form of interaction between the firms involved in international trade, when the action taken by any one firm may have significant effects on other firms. This interaction is certainly absent in perfect competition, and is certainly present in oligopoly, so much so that strategic trade policies and oligopolistic models of international trade go hand in hand. This is why the theory of strategic trade policy has been developed in the context of the new theories of international trade, as by definition no strategic trade policy may arise in the context of the traditional theory.

Sometimes the meaning of *strategic* trade policies is extended to include the case in which the interaction arises between governments pursuing optimal (for each) trade policies rather than between the firms involved in international trade. Under this extension strategic trade policies may also arise in the context of the traditional theory: the optimum tariff (see Sect. 11.1) would be a typical example.

Since the results of the theory of strategic trade policy are contradictory, it is no surprise that this literature is not a useful guide to government policy at this time (Guerrieri & Padoan, 1996; Haberler, 1990; Markusen, Melvin, Kaempfer, & Maskus, 1995, p. 293). It is however important to examine it better to understand why and how these results are heavily model-dependent.

10.8.1 A Tariff Under Vertical Product Differentiation

It is possible to introduce the presence of impediments to trade in Falvey's model (see Sect. 9.1), with interesting results. Suppose that a country, say country 1, introduces a tariff d. This means that the prices of country 2's commodities in country 1 will rise from $p_2(\alpha)$ to

$$p_2'(\alpha) = (1 + d)p_2(\alpha) = (1 + d)W_2 + (1 + d)R_2\alpha.$$

In terms of Fig. 9.1 this implies an upward shift (accompanied by an increase in the slope) of the relevant straight line (see Fig. 10.11).

Country 1 will now only import products of quality lower than α_0'. On the other hand, country 2—for which the relevant comparison is still between $p_1(\alpha)$ and $p_2(\alpha)$—will continue importing solely products of quality higher than α_0. The qualities between α_0' and α_0 will no longer be traded.

Thus we see that the introduction of a tariff gives rise to a range of non-traded qualities, and decreases intra-industry trade. This trade-reduction effect is an

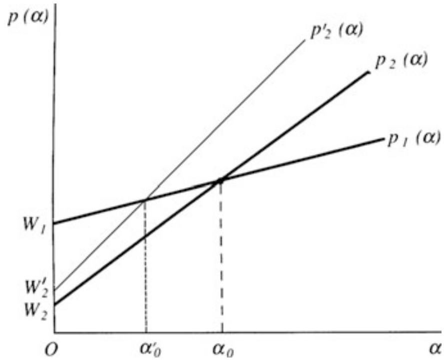


Fig. 10.11 Trade policy with vertical differentiation

increasing function of the tariff rate. Consequently, the elimination or the reduction of the tariff (for example because a customs union is created) will certainly have a trade creation effect, since it will increase international (intra-industrial) trade by causing an increase in the range of exported and imported qualities. This, according to Falvey, is consistent with the empirical evidence, for example that of the European common market.

For further study of the effects of a tariff in this model, see the Appendix, Sect. 24.4.1.

10.8.2 Monopolistic Competition and Welfare-Improving Tariff

It is on the study of the effects of a tariff (or other protective measures) on the number of varieties produced and on the unit cost of production that the free trade versus protectionism debate has concentrated its attention. Protectionist measures could indeed allow domestic firms to increase the scale of production and so bring about a price decrease thanks to internal economies. This can be seen as a new argument in favour of the protection of the infant industry, a problem examined in Sect. 11.2 in the context of the traditional theory. The effect on the number of varieties produced is not so clear. In addition, the use of protectionist measures can give rise to retaliation, in which case the result is probably a reduction in social welfare even when the effect of protectionism would have been positive in the absence of retaliation. For details see Baldwin (1992, sect.İV) and Pomfret (1992, chap. 6), and the Appendix, Sect. 24.4.2.

10.8.3 Strategic Trade Policy Under Oligopoly with Homogeneous Good

In this model (see Sect. 9.3.2), we obtain conventional results as regards the effects of tariffs. The imposition of a tariff on imports causes a decrease in the foreign

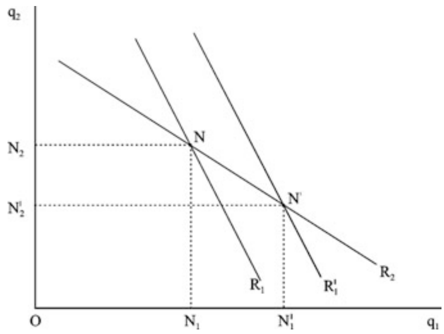


Fig. 10.12 Duopoly and strategic trade policy: subsidies

firm's share of the domestic market (hence a reduction in imports and an increase in the domestic firm's share). In this process the size of the market (i.e., the overall quantity supplied by the foreign and domestic firm) will decrease, causing a price increase. For details see the Appendix, Sect. 24.4.3.

More interesting results are obtained if we consider subsidies. Let us modify the model by assuming (Brander & Spencer, 1985) that the two oligopolistic firms, one in each country, only produce for export into a third market. Interaction between them is always of the Cournot type, and their reaction functions are drawn in Fig. 10.12. The equilibrium point is N, so that ON_1 and ON_2 are the quantities respectively produced by firms 1 and 2; the total quantity $ON_1 + ON_2$ is sold in the third market. Suppose now that country 1 grants a subsidy to the domestic firm. This lowers firm 1's marginal cost and hence shifts its reaction curve to position R_1'. The new equilibrium point is N', and ON_1', ON_2' are the new equilibrium quantities. Firm 1 produces more while firm 2 produces less. It can also be shown (see the Appendix, Sect. 24.4.3.2) that firm 1 enjoys higher profits while firm 2 suffers a profit reduction. This is fairly intuitive; what is less intuitive (see the Appendix) is that country 1 enjoys an increase in welfare, since the domestic surplus (the profit of firm 1 minus the cost of the subsidy if any) increases with respect to the no-subsidy situation.

This shows that in an oligopolistic market, an export subsidy can provide a strategic advantage to the domestic firm and hence shift rents (from the foreign to the domestic firm) and ultimately cause a welfare increase in the country that subsidizes the domestic firm.

10.8.4 Strategic Trade Policy Under Oligopoly with Differentiated Good

In the Eaton-Kierzkowski model (see Sect. 9.3.4) we can show that the imposition of a tariff by country 2 on its imports of commodity B will improve country 2's welfare. We have just seen that the price charged by country 1's monopolistic producer is at the limit of indifference for country 2's consumers. It follows that, because of the tariff, this producer will have to reduce the export price to country 2 in such a way that the final price (export price+tariff) to country 2's consumers does not increase; otherwise there can be no export. The firm under consideration will be willing to accept such a reduction insofar as its profits, though lower than before, are still greater than those that it would obtain giving up any export to country 2 and only producing for its domestic market like in the pre-trade situation.

In such a situation the government of country 2, being aware of the strategic interaction, may even calculate (and impose) a tariff that takes away from the foreign firm all profits in excess of profits this firm earns by selling only to consumers in country 1. In such a case this firm is indifferent between selling only in the domestic market or exporting as well.

Be it as it may, country 2 will be better off because—although there is no welfare increase for the consumers, who pay the same price as before—there is the benefit of the increase in the fiscal revenue (the revenue of the tariff) of country 2's government at no cost. This, again, is contrary to the traditional theory, according to which the imposition of a tariff does in general cause social costs. The difference in results is clearly due to the different market form assumed as well as to the particular nature of demand.

Box 10.3 Strategic Trade Policy: Boeing vs. Airbus

The aircraft sector provides a textbook example of governmental startegic trade policy, namely an industry in which trade policy could affect the strategic interaction between a domestic and an international rival: by subsidizing production, the government can affect the outcome of the competitive game in such a way as to shift rents in favour of the domestic firms as argued by Brander and Spencer (1985). In the commercial aerospace industry the production has been directly and indirectly supported by using the market failure argument. The aerospace industry is surely subject to market failure, notably because of large scale economies in production and the importance of research and development. Given this industry's market structure it is difficult for individual countries to face international competition, so aircraft industry has given rise to significant international cooperation. One of the most famous case of such cooperation is the European Airbus Consortium which was formed in the late 1960s to challenge the dominance of the Boeing Corporation in international world markets. The public support to Airbus has mostly taken the form of a reduction in fixed development cost.

Although not recent, the case of Boeing vs. Airbus contains useful background information on the subsidy issue. Boeing has long been the leader in the world aviation industry and when Airbus was created the commercial aircraft was almost controlled by US firms. Airbus slowly but steadily expanded its market share during the first two decades of its existence and with other competitors out of the picture (Lockheed and McDonnel Douglas) the battle for market share in the 1990s and beyond is being waged directly at Boeing's expense.

Boeing and McDonnell Douglas accused Airbus to be state-supported with virtually unlimited (hence unfair) financial resources in the form of cheap loans, the repayment of which was contingent on Airbus's profits. On the other side the Europeans argued that American aircraft manufactures received indirect government subsidies—from the Department of Defense and NASA—of comparable magnitude. The battle over the appropriateness of subsidies raged for the first 22 years of Airbus' presence. The US government lodged a complaint against Airbus under the GATT and in 1992, an "Airbus Agreement" was signed between the United States and the European Community. This agreement contained three main points:

- Direct government subsidies for aircraft were capped at 33 % of developments costs. Loans made to the consortium were to be repaid according to strict scheduling and interest-rate requirements.
- Indirect subsidies were limited to 3 % of the turnover of civil aircraft manufacturers;
- A bilateral panel would monitor compliance of the previous two points, and increase the "transparency" of the commercial aircraft industry.

Strategic trade policy emphasizes its results in the presence of oligopoly: any external intervention alters the strategic interaction between players on the market. If a domestic firm is a part of an international oligopoly and receives any kind of support from its government, it competes successfully. There seems to be little doubt that the Airbus project would not be in a position of such prominence without government support.

References

Balassa, B. (1965). Tariff protection in industrial countries: An evaluation.
Baldwin, Robert E. (1971). *Non-tariff distortions of international trade*.
Bhagwati, J. N. (1965). On the equivalence of tariffs and quotas.
Bhagwati, J. N. (1971). The generalized theory of distortions and welfare.
Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (2nd ed., Pt. II).
Caves, R. E. (1979). International cartels and monopolies in international trade.
Chacholiades, M. (1978). *International trade theory and policy* (Pt. VII).
Corden, W. M. (1971a). *The theory of protection*.
Corden, W. M. (1974). Trade policy and economic welfare.
Corden, W. M. (1984b) The normative theory of international trade.
Deardoff, A., & Stern, R. M. (1984). The effects of the Tokyo round on the structure of protection.
El-Agraa, A. M. (1984). *Trade theory and policy*.
Greenaway, D. (1983). *International trade policy: From tariffs to the new protectionism*.
Henderson, J. M., & Quandt, R. E. (1980). *Microeconomic theory: A mathematical approach* (Sects. 3.7, 7.1–7.3, 11.7).
Hicks, J. R. (1981). *Wealth and welfare* (pp. 100–132).
Johnson, H. G. (1969). The standard theory of tariffs.
Laird, S. & Yeats, A. (1990). Trends in non-tariff barriers of developed countries.
Lerner, A. P. (1936). The symmetry between import and export taxes.
Lloyd, P. J. (1974). A more general theory of price distortions in open economies.
Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: Theory and evidence* (chaps. 15–16).
McMillan, J. (1989). *Game theory in international economics*.
Meade, J. E. (1952). *A geometry of international trade* (chap. VI).
Meade, J. E. (1955). *The theory of international economic policy* (Vol. 2).
Metzler, L. A. (1949). Tariffs, the terms of trade, and the distribution of national income.
Morey, E. S. (1984). Confuser surplus.
Pearce, I. F. (1970). *International trade* (Book I, chaps. 7 and 9).
Razawi, H. (1984). An economic model of OPEC coalition.
Stern, R. M. (1973). Tariffs and other measures of trade control: A survey of recent developments.
Stolper, W. F. & Samuelson, P. A. (1941). Protection and real wage.
Takacs, W. E. (1978). The nonequivalence of tariffs, import quotas, and voluntary export restraints.

Takayama, A. (1972). *International trade: An approach to the theory* (Pt. V).
Vanek, J. (1962). *International trade: Theory and economic policy* (chaps. 16–18).
Vousden, N. (1990). *The economics of trade protection.*
WTO (1995). Trading into the future: WTO-The World Trade Organization.
Yeats, A. J. (1974). Effective tariff protection in the United States, the European Economic Community, and Japan.
Yoo, J.-H. (1993). The political economy of protection structure in Korea.

Chapter 11
Free Trade vs. Protection, and Preferential Trade Cooperation

In the previous chapter we have implicitly talked of protectionism by talking of tariffs, quotas, etc. This chapter explicitly examines the main arguments in favour of protectionism and the rebuttal of them by the advocates of free trade; here the theory of second best will throw new light on this age-old debate. We shall then go on to examine preferential trade cooperation among countries. This cooperation has the purpose of reducing or eliminating protection among the participating countries, and may take various forms, but in any case the main question is whether these countries are better off. For other treatments of the topics examined in this chapter see, e.g., Bhagwati et al. (1998), Corden (1971a, 1974, 1984b), Deardoff and Stern (1984), El-Agraa (1984), Ethier and Horn (1984), Greenaway (1983), Hicks (1981), Jones (1979), Kemp (1976), Krauss (1972), Lipsey (1960), Lloyd (1974), Markusen et al. (1995), Meade (1952), Pearce (1970), Puga and Venables (1997), Swann (2000), Takayama (1972), Vousden (1990).

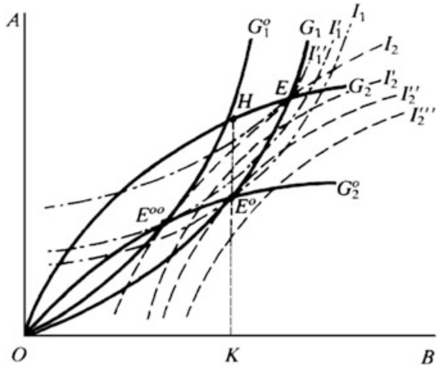
It should be pointed out that by *protectionism* in the broad sense we mean any intervention of the government (which may consist of tariffs and/or any other non-tariff barrier) giving rise to a divergence between domestic relative prices and world relative prices of the same commodities. More precisely, this divergence must be greater than that accounted for by costs of transport (including insurance). But, as usual, we shall ignore these costs to simplify the analysis.

In our treatment we shall refer exclusively to the welfare of the society as a whole and not to the welfare of single rent-seeking groups within the society (for a treatment of these problems see, e.g., Bhagwati, 1982a, 1995; Krueger, 1974).

Finally, we must point out that our treatment will be confined in the context of the traditional theory: strategic trade policy in the context of the new trade theories has been examined in Sect. 10.8.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_11,

Fig. 11.1 The optimum tariff



11.1 The Optimum Tariff

Protectionism is better than free trade because—so the argument runs—it is always possible to find a tariff such that the imposing country's welfare is greater than under free trade.

The examination of this argument requires the study of the welfare effects of a tariff in the general equilibrium context presented in Sect. 10.5.2. This can be done by introducing social indifference curves.

In Fig. 11.1, besides the two countries' offer curves (assumed to be normal), also their social indifference curves are drawn. It should be noted that the latter curves, unlike those drawn in previous diagrams—see for example Fig. 10.1—are *increasing*. This can be explained as follows, considering, for example, country 2. Whilst on the horizontal axis there are the quantities of B obtained (imports), on the vertical axis there are the quantities of A released (exports). It is therefore obvious that a greater amount of the commodity acquired will have to correspond to a greater amount of the commodity given up so as to remain on the same indifference curve. The (ordinal) index of satisfaction increases as we move downwards and to the right, for in I_2' the amount of the commodity acquired is greater than in I_2 with the same amount of the commodity given up (take any horizontal straight line—not shown in the diagram—parallel to the B axis). Finally, these curves are convex to the export axis (concave to the import axis) because in order to maintain a given satisfaction level, ever decreasing successive increments of the commodity given up (exports) will correspond to equal successive increments of the quantity of the commodity acquired (imports). This is the equivalent of the principle of decreasing marginal rate of substitution along the usual curves.

In like manner we can draw the family of country 1's social indifference curves: I_1, I_1', I_1'' etc. Let us now assume that country 2 imposes a tariff, so that OG_2 shifts to OG_2^0: the new international equilibrium point is E^0. Country 2's welfare has increased, for I_2'' represents a higher welfare than I_2 does, and this confirms that in normal circumstances the imposition of a tariff improves the terms of trade and the welfare of the imposing country. From the diagram we also see that country 1's welfare has decreased, for this country is now on I_1 which represents a lower welfare than I_1'' does. Therefore the tariff-imposing country increases its own welfare at the expense of the other country, which confirms the opinion that *the free-trade situation is a Pareto-optimum* (so that it is not possible to improve the situation of a country without worsening the other country's). It goes without saying that, given the ordinal nature of the social indifference curves, it is not possible to ascertain whether the welfare of the world as a whole has increased or decreased as a consequence of the movement from E to E^0.

Let us go back to country 2 and investigate the welfare-maximizing tariff from its point of view, that is, country 2's optimum tariff. Graphically, this amounts to finding country 2's highest social indifference curve compatible with the *given* offer curve of *country* 1. It turns out that this curve is exactly I_2'', tangent to OG_1 at point E^0. It should be stressed that the tangency is to be found between a social indifference curve of country 2 and country 1's offer curve, which is the constraint of the problem. In fact, as we know, each country can, by imposing a tariff, cause a shift in *its own* offer curve, but cannot influence the other country's offer curve. This explains why the constraint for country 2 is country 1's offer curve and vice versa.

Thus in our case the optimum tariff is that which shifts country 2's offer curve downwards so as to make it pass exactly through point E^0, namely from OG_2 to OG_2^0. The corresponding optimum tariff rate can be computed graphically as shown in Sect. 10.5.2, for example as HE^0/E^0K.

We have thus demonstrated the proposition that *for the single country there always exists a cum-tariff (the optimum tariff) situation superior to free trade*. But of course the other country's welfare worsens, as we have shown above, and this may give rise to retaliation.

In fact, we have so far assumed that OG_1 is given, thus implicitly assuming that *country* 1 *does not introduce tariffs*. But, if we exclude non-economic factors, it is not plausible that country 1 should not retaliate: this country, therefore, will also levy a tariff, presumably the optimum one from its own point of view. As the first step has already been made by country 2, country 1 will take the OG_2^0 curve as given, and determine its own optimum tariff as that corresponding to point E^{00}, where an indifference curve (I_1') is tangent to OG_2^0.

We observe that thanks to the retaliation, country 1 recovers part of (though not all) the loss due to the initial imposition of a tariff by country 2: country 1, in fact, passes from I_1 to I_1' which, though better than I_1, is worse than the initial I_1''. We also observe that in E^{00} international trade is further reduced with respect to E^0.

But not even point E^{00} is a stable equilibrium: in fact, once the tariff war begins, there is no reason why country 2 will not counter-retaliate and impose a new optimum tariff in correspondence to OG_1^0, and so forth.

It is not possible to determine a priori a precise outcome of the tariff war, for in general it is possible either that the process continues until trade disappears because tariffs have reached the prohibitive level in both countries[1] or that it stops before

[1]This cannot happen with the curves drawn in Fig. 11.1, but it is conceivable that it may happen with other curves.

for various reasons, for example because a stable equilibrium situation has been reached: this happens when a point is reached where the optimum tariff change is zero for both countries, that is, each country, by taking the other's offer curve as a constraint and maximizing its own welfare, finds that the optimum situation is the current one. This possibility can be readily verified by experimenting with diagrams similar to Fig. 11.1. It is also possible for a "tariff cycle" to occur: see Johnson (1953), who considers all possible cases. Other reasons for the tariff war to stop before the disappearance of trade are that one country yields, or the two countries reach an agreement (in this case it is even possible for the initial free trade situation to be restored or for a bilateral tariff cut to be negotiated). The outcome can also be studied in the context of game theory (McMillan, 1989, chap. 4).

Therefore the statement made above, that for the single country there always exists a cum-tariff situation better than the free trade one, must be taken with caution, as it may be no longer valid in the presence of retaliation.

That statement, however, enables us to show the lack of general validity of the *first* of the two traditional propositions concerning the relationships between international trade and social welfare, which are:

1. Free international trade is better than restricted (tariff-ridden) trade;
2. Some international trade, even if restricted, is better than no trade.

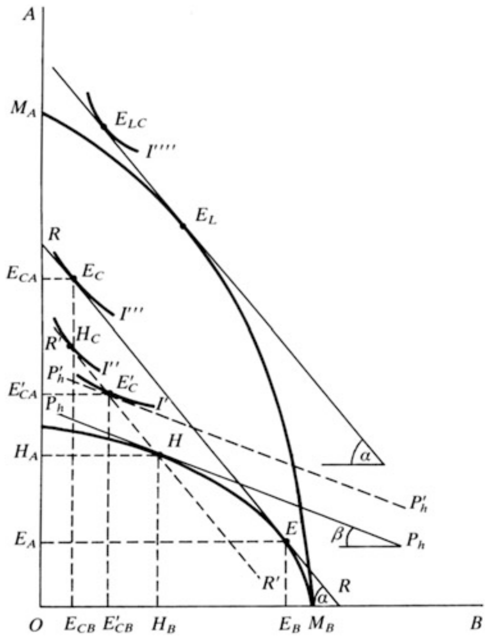
On the contrary, the *second* proposition remains valid even in the optimum-tariff context. With reference to Fig. 11.1 above, we see that it is always possible to find a restricted-trade point (for example, E^{00}), such that the social indifference curves passing through it certainly do represent for both countries a social welfare greater than that represented by the respective social indifference curves passing through the origin (these are not shown, but can be readily drawn by the reader).

11.2 The Infant Industry

This is probably the oldest and best known argument for protectionism: a domestic industry in its infancy cannot compete with well-established foreign firms and therefore it must be protected by a tariff, to give it time to grow up and become competitive with foreign firms; at that point the protection can, and must, cease. It is clear that for the validity of this argument it is necessary for the protected industry to have within it the germs for growing up to the level at which it can compete with foreign firms at world prices and, in addition, that the benefits accruing to society from the operation of this industry when protection is discontinued, will more than compensate for the losses deriving from the protection itself.

But, even if these conditions are satisfied, it can be seen that the advantages of the infant industry becoming adult can be obtained with lower costs by way of non-tariff protection, for example by giving the infant industry a subsidy which enables it to charge domestic consumers a price for the commodity equal to the world price.

Fig. 11.2 The infant industry



This can be shown by way of the analysis made in Sect. 10.5.1 and, in particular, of Fig. 10.3, reproduced in Fig. 11.2. A tariff levied on commodity A (the importable) shifts the production point from E to H, thus favouring the domestic output of A. But, as we saw, the consumption point shifts from E_C to E'_C, so that social welfare decreases, for the indifference curve I' is lower than I'''. If instead of imposing a tariff the government subsidized the domestic output of A so as to reach the same production point H, the situation would improve. In fact, as there is no difference between the domestic and the world relative price, the consumption point would be H_C, which lies on I'', higher than I'.

The same diagram can be used to see the long-run advantages deriving from the protection of the infant industry, provided that it succeeds in becoming competitive with foreign firms. Thanks to the protection, there is a continuous improvement in production techniques, labour skills, etc., in the sector producing commodity A, so that the country's transformation curve shifts gradually upwards and to the right,[2] up to the long-run position $M_A M_B$. At this point protection can cease and (for simplicity's sake we assume that the terms of trade remain the same) the country will produce at E_L and consume at E_{LC}. In the diagram we have illustrated the case in which E_{LC} is to the left of E_L so that the country remains an importer of commodity A, but it may equally well (with different shifts of the transformation

[2]This amounts to saying that protection has enabled sector A to benefit from technical progress in a broad sense. On technical progress and (free) international trade see Sect. 13.5.

curve and/or different shapes of the social indifference curves) become an exporter of this commodity (point E_{LC} is to the right of E_L along the terms-of-trade line). In any case the long-run consumption point E_{LC} will lie on a higher indifference curve than E_C does.

This is a comparative-static result; dynamically, the economic system can follow various paths to pass from E_C to E_{LC}, but in any case there is an initial fall in social welfare, from I''' to I' (if a tariff is used) or to I'' (in the case of a subsidy). As the transformation curve shifts, welfare increases, but will remain below I''' for a longer or shorter time before overtaking it and increasing towards I''''.

It is therefore clear that the protection of an infant industry, even if it may give benefits in the long run, will cause welfare losses in the short and medium run. Isn't it possible, then, to balance benefits and costs and check whether there is a net benefit or a net cost of protecting the infant industry? In theory the answer is yes, provided that one has a sufficient amount of very precise information. One must, in fact, not only know the precise *dynamic path* followed by the economic system but also assume that social welfare can be *measured* (or proxied) by a cardinal function and, finally, determine a *social discount rate* to bring to the same point in time the various quantities of future welfare and thus be able to compare the various alternatives. Now, even if it is granted that the required information can be obtained, it would nevertheless remain true that the aforementioned elements would be different from case to case, so that it is *not* possible to state *in a general* way that protection of the infant industry is definitely beneficial or definitely harmful. It is however possible to state that, with the same benefits, costs are lower if a subsidy is used instead of a tariff, as shown above.

11.3 Distortions in Domestic Goods Markets

We consider here all those situations in which the domestic relative price of commodities does not reflect, as it should, the marginal rate of transformation. These distortions may be due to monopolistic elements (which make the selling price higher than the marginal cost) or to external economies or diseconomies (which make the producer's marginal cost different from the social marginal cost, that is, cause a divergence between the private and the social marginal cost).

When the domestic relative price and the marginal rate of transformation are unequal, free international trade may even cause a *decrease* in welfare with respect to the autarkic situation. This (possible but not necessary) case is shown in Fig. 11.3. In the pre-trade equilibrium, the country is producing and consuming at point Q, where—because of distortions—the domestic relative price (p_B/p_A, represented by the absolute value of the slope of $P_h P_h$) is different from the marginal rate of transformation (slope of $P_e P_e$). More precisely, since the slope $P_h P_h$ is greater than that of $P_e P_e$, the price of commodity A is too low relative to that of B. The world price ratio is represented by the slope of $P_i P_i$, smaller than that of $P_h P_h$: this signals the fact that the country has a comparative advantage (at distorted prices) in

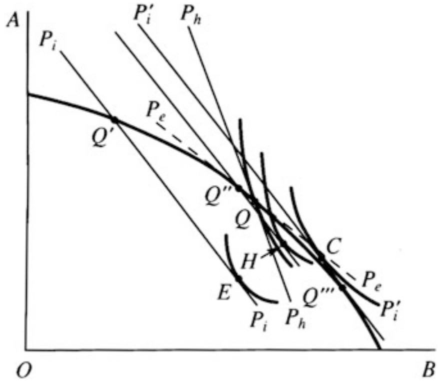


Fig. 11.3 Distortions in domestic goods markets: specialization in the wrong direction

commodity A, since p_B/p_A is higher, and so p_A/p_B is lower, on the international than on the domestic market. Summing up, the situation at point Q is given by the double inequality

$$(p_B/p_A)_e < (p_B/p_A)_i < (p_B/p_A)_h \,, \tag{11.1}$$

that is,

$$(p_A/p_B)_h < (p_A/p_B)_i < (p_A/p_B)_e \,. \tag{11.2}$$

Since, as we said, the signal to which domestic producers respond is given by the comparison between $(p_B/p_A)_h$ and $(p_B/p_A)_i$, or, what amounts to the same thing, between $(p_A/p_B)_h$ and $(p_A/p_B)_i$, when trade is opened up, the country increases the output of A and moves to a production point on the left of Q, for example Q'. As the country can exchange goods at the given terms of trade—we are making the small country assumption—the consumption point will be E. Point E is clearly inferior to point Q, as it lies on a lower indifference curve. But, as we said at the beginning, the welfare loss with respect to the initial autarkic situation is not a necessary outcome. If, for example, the production point is Q'' instead of Q', the country will be able to consume at H, which lies on a higher indifference curve than Q.

We observe that, in all the cases examined, the distortion has induced the country to specialize in the wrong direction. In fact, in the absence of distortions, the domestic price ratio at Q would have been equal to the slope of P_eP_e, showing the *true* comparative advantage to be in commodity B. Now—so the protectionist argument runs—the introduction of a tariff on commodity B, by increasing its domestic relative price, stimulates the production of this commodity in which—as we saw a moment ago—the true comparative advantage lies, thus increasing the country's welfare. But the imposition of a tariff, which in this case involves a production gain (deriving from a better allocation of resources), causes a consumption loss, so that the net result can be, in general, either a loss or a gain.

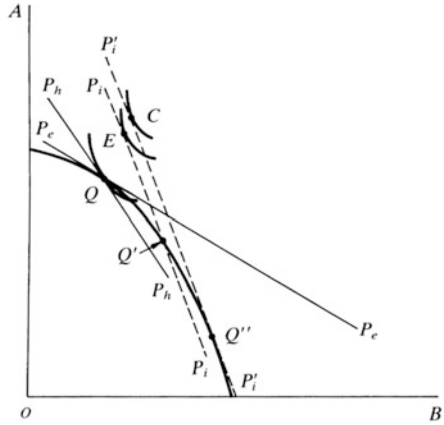


Fig. 11.4 Distortions in domestic goods markets: specialization in the right direction

We must further observe that the imposition of a tariff can never *reverse* direction of international trade: it can, at most, make imports of the commodity cease (prohibitive tariff), but will never make this commodity become a exportable. Now, the optimal situation for the country (determined by comparing the slope $P_e P_e$ with the terms of trade) is to be an exporter, rather than an importer, of commodity B, as can be seen from the fact that the (hypothetical) $P_i' P_i'$ parallel to $P_i P_i$,would give rise to the production point Q''' and the consumption point C (both hypothetical). It is therefore obvious that tariff will never be optimal, even if it were to improve social welfare with respect to the free trade situation.

Thus the imposition of a tariff is not the best policy, even in this case. The optimal policy—better than both free trade and a tariff—is to subsidize the production of B and/or tax that of A, so as to reduce the domestic price ratio p_B/p_A to the level of the marginal rate of transformation (that is, $P_h P_h$ comes coincide with $P_e P_e$). The country can then engage in free trade and obtain maximum welfare by producing at Q''' (which from being hypothetical now becomes actual) and consuming at C.

We have so far examined the case in which the country, as a consequence of the distortions, specializes in the wrong direction; the conclusions, however, do not change even if it specializes in the right direction. A specialization in the right direction occurs, for example, when the terms of trade, instead of being include between the marginal rate of transformation and the distorted domestic relative price, are greater than both. In symbols,

$$(p_B/p_A)_i > (p_B/p_A)_h > (p_B/p_A)_e \,, \tag{11.3}$$

that is,

$$(p_A/p_B)_e > (p_A/p_B)_h > (p_A/p_B)_i \,. \tag{11.4}$$

In this case the signal coming from the comparison between the terms of trade and the (distorted) domestic price ratio points in the right direction—that is, the same direction in which the comparison between the terms of trade and the marginal rate of transformation would point—as can be seen from inequalities (11.3) and (11.4), even if not with sufficient intensity.[3] In terms of Fig. 11.4, the country moves from Q to the right, but does not reach the optimum position Q'' as the too feeble signal induces it to stop beforehand, for example at Q', and to consume at E. Here social welfare is certainly better than that at the autarkic point Q, though lower than at the optimal consumption point C (corresponding to the production point Q'').

In such a situation the advocates of protectionism suggest, to offset the distortion, a commercial policy such as, for example, a subsidy to exports. This subsidy brings about a production gain but a consumption loss (due to the fact that domestic consumers pay a higher price for commodity B than the one that foreign consumers are charged), with an ambiguous net outcome.

Also in this case, the optimal policy is a less protectionist one, that is, a subsidy to the domestic production of B (and/or a tax on the domestic production of A), so as to offset the initial price distortion without introducing consumption losses. In fact, since the subsidy is to *all* domestic production (not only to the part of it being exported), the production of this commodity will be enhanced with no consumption loss, because—as all production is subsidized—domestic consumers will pay the same lower price as the foreign ones. The "less protectionist" qualification is due to the fact that, according to the classification in Sect. 10.6.4, point (c), a subsidy to production is in general listed among the obstacles to free trade. On production subsidies see Sect. 12.4.

11.4 Distortions in Domestic Factor Markets

These distortions imply that the equality between the price of a factor and the value of its marginal productivity and/or the equalization of the price of a factor in all sectors do not hold. For example, the industrial wage rate may exceed the value of the marginal product of labour in industry, or the wage rate in agriculture may be different from that in industry, though both are equal to the value of the respective marginal productivities. This will lead to an inefficient allocation of resources and, consequently, the country will not be on its true transformation curve, say $T_A Q T_B$ in Fig. 11.5, but on a lower curve, say $T_A Q' T_B$.

In other words, the distortions under consideration prevent the country from reaching the efficiency locus in the box diagram (see Sect. 3.1), since the conditions of efficiency require that the marginal rate of technical substitution (given by the ratio between the marginal productivities of the two factors) should be equal in both sectors and equal to the (common) factor-price ratio. One

[3] It is also possible for the signal to be too strong, so that the country overspecializes in the right direction and overshoots the optimal point. See, for example, Chacholiades (1978, pp. 509–510).

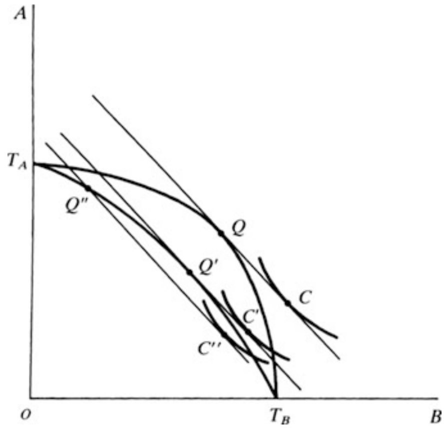


Fig. 11.5 Distortions in domestic factor markets

might think that no problem arises so long as this ratio is the same though the absolute factor prices are different (of course by a common factor, say, $p_{LA} = qp_{LB}$, $p_{KA} = qp_{KB}$, $q > 0$), and each factor is paid the value of its marginal productivity. This, however, is not true, for the optimum conditions also require (see Sect. 19.1) that $p_A/p_B = MPL_B/MPL_A = MPK_B/MPK_A$. Now, if $p_AMPL_A = p_{LA}$, $p_BMPL_B = p_{LB} = qp_{LA}$ etc., we get $p_AMPL_A/p_BMPL_B = q$ etc., whence $p_A/p_B = qMPL_B/MPL_A$ etc., which is not consistent with the optimum conditions unless $q = 1$.

Since the transformation curve T_AQT_B is derived—as shown in Sect. 3.1—from the efficiency locus, if the country is not on this locus its production possibilities will also be lower than the maximum ones (represented by the T_AQT_B curve), whence the curve $T_AQ'T_B$. Let us note that the intercepts with the axes are the same, because, when all productive factors are employed in the production of one good, no problem of resource allocation arises and the distortions will be irrelevant.

In Fig. 11.5 we have drawn a family of parallel straight lines with a slope equal to the given terms of trade. If we assume for the time being that the distortions in factor markets have no effect on goods markets, what happens is that the country will produce at point Q' (which is optimum with reference to the distorted transformation curve) instead of producing at Q, and will consume at C' instead of C, thus achieving a lower welfare level.

It is however to be presumed that the distortions in factor markets will cause distortions in goods markets so that—as shown in Sect. 11.3—the country's production point will be to the left of Q', for example, at Q'', and the consumption point will be at C''. We must now distinguish between two aspects of the problem: the achievement of the optimum point on the distorted transformation curve and of the optimum point on the true transformation curve. As regards the former, the prescription is the same as that given in Sect. 11.3: the optimal policy is not the imposition of a tariff, but a subsidy to the production of A and/or a tax on that of B, so as to cause the country to reach the optimum point Q' on the distorted transformation curve $T_A Q' T_B$. As regards the latter aspect, the optimal policy will consist in taxes (and/or subsidies) on the use of factors, so as to eliminate the divergences which cause the distortions: in this way the efficiency conditions are restored and the country can move to the true transformation curve $T_A Q T_B$, then producing at Q and consuming at C which denotes a higher welfare level than C'.

11.5 Non-economic Motives for Protection

The most frequently cited non-economic motives for protection are three in number. The first and perhaps oldest motive is *national defence*. The seventeenth century British mercantilists already used this argument to advocate protection for the domestic shipbuilding industry, which in their opinion had to be kept strong and flourishing so that, in case of war, warships could be rapidly built. More generally, seeing that, if war breaks out, international trade will be reduced or even discontinued, the country must maintain domestic production of certain strategic commodities (even if in period of peace it is more expensive to produce them domestically than to import them) so as not to find itself at the mercy of the enemy should war come.

The second motive is *national pride*. To produce a certain commodity at home may become a motive of national pride, much as winning Olympic medals or the America Cup. In such cases, the industry producing that commodity will be protected in any event, even if this involves a very high cost.

The third motive refers to *foreign policy*. Countries often use economic means (aid, tariffs, embargoes, etc.) to obtain political benefits.

It is obvious that in all cases in which non-economic motives are present, protectionism will be brought into being even if it were shown not to be advantageous from the strictly economic point of view. This, it should be stressed, is not at all irrational, for it simply means that in the social welfare function the arguments “national defence”, “national pride”, etc., are also present and predominate over economic arguments.

11.6 The Theory of Second Best

It is time to inquire whether it is possible to reach general results on the outcome of the free-trade-versus-protection debate. Many authors share the opinion that free trade is better than restricted trade (excluding the case of the optimum tariff without retaliation) and that, if the country wants to help infant industries or correct the effects of distortions, it had better use subsidies and/or taxes rather than tariffs, as was shown in Sects. 11.2–11.5. This opinion, however, must be qualified, as its validity has been demonstrated in a situation of free competition in all national and international markets (of both commodities and factors).

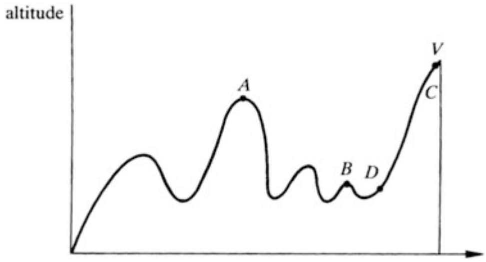


Fig. 11.6 Intuitive graphic representation of the theory of second best

When this situation does not occur (and as a rule it doesn't), the problem is quite different and we must have recourse to the theory of second best. This theory purports to find the (second) best situation when (because of distortions or whatever) it is not possible to fulfil all the conditions for a Pareto optimum (first best). The fundamental principle of this theory is that *once one or more of the Pareto-optimum conditions is violated, it is not necessarily true that the (second) best situation is that in which all the remaining conditions are fulfilled.*

A corollary of this principle is that it is not possible to ascertain on purely a priori grounds whether the replacement of a violation of the Paretian conditions with another violation improves or worsens the situation. Another corollary is that the elimination of a violation (except when it is the only one) does not necessarily improve the situation, and that the introduction of a further violation does not necessarily worsen it. In other words, this means that, in a world in which are present non competitive situations, distortions, and various restrictions to free trade, *the elimination of one or more of these restrictions does not necessarily mean the achievement of a better situation, and the introduction of one or more further restrictions does not necessarily mean a deterioration of the situation but, paradoxically, might even lead to a better situation, though still suboptimal.*

A rigorous proof of the fundamental principle of the theory of second best will be given in Sect. 25.2; here we give an intuitive idea of it, by elaborating on an analogy due to Meade (1955, p. 7). Imagine a person who wishes to reach the highest point on a range of hills. In walking towards this point, the person will have to climb lower hills and then go downhill: it is therefore not true that to reach the goal this person will always have to walk uphill. Furthermore, as the highest hill is surrounded by lower ones of different heights, after having climbed one hill the person will probably have to climb yet another one but of lower height: it is therefore not true that any movement towards the target brings the climber to an ever higher point.

Elaborating further on this effective analogy, if for example a gorge or another insuperable obstacle prevents the climber from reaching the summit and if this person's objective is despite everything to climb to the highest possible point, our climber may have to go back quite a long way if the second highest hill is a great distance from the very highest. In terms of Fig. 11.6, the climber arrives at B and sees that the way to V is blocked by an insuperable gorge at D. Then, instead of staying at B or, worse, walking towards V as far as D, the climber will have to backtrack to A to reach the second highest point.

Now, if we apply the theory of second best to the free-trade-versus-protection debate, it immediately follows that, *in the real world*, it is not possible to ascertain a priori whether a protectionist policy improves or worsens the situation nor is it possible to state that any movement towards freer trade automatically gives rise to an improvement.

Similarly, it is not possible to state, as the traditional theory goes, that there exist other policies decidedly better than the imposition of a tariff. This statement, in fact, is certainly true only if *all* the violations of the Pareto-optimum conditions are eliminated; a particular case occurs when there is *only one* violation (for example a distortion in the factor market or in the goods market), in which case the elimination of the violation without the introduction of others restores the optimum situation for certain (in terms of Fig. 11.6, if our climber is at C, the last step uphill will certainly bring this person to V). This, as the reader can check, has implicitly been the line of reasoning followed—in accordance with traditional theory—in Sects. 11.2–11.5.

But if, as is true in the real world, there are numerous violations of the Paretian conditions, it follows from the theory of second best that it is not possible to state for certain that a policy which eliminates one of these without introducing another violation is better than a policy which eliminates the same violation by introducing another.

It is clear that, things being so, it becomes impossible to make statements valid in general and deduce a priori policy prescriptions from a limited number of guiding principles. In reality any outcome is possible and one must ascertain which is the best policy (free trade or protection in its various forms) in each actual case without being blinded by theoretical preconceptions.

11.7 Preferential Trading Cooperation

11.7.1 The Various Degrees of Cooperation

After dealing with tariffs and protectionism it is natural to proceed to the theory of preferential trading cooperation, whose main forms (in order of increasing degree of integration) are:

1. A *preferential trading club* (or *agreement*), which is an agreement between two or more countries to reduce tariffs and other restrictions on imports from one another; each member, however, retains complete freedom to impose different tariffs and other restrictions on imports from non-member countries.
2. A *free-trade area* (or *association*), in which the partner countries abolish tariffs and other restrictions on imports from one another, while retaining complete freedom over their commercial policies towards the rest of the world.
3. A *customs union*, which, in addition to the provisions of the free-trade area, establishes a common external tariff schedule on all imports from non-member countries.

4. A *common market*, in which the countries, in addition to the provisions of the customs union, allow free movement of all factors of production among themselves.

It should be pointed out that cooperation can exceed agreements on free movement of goods and factors. The partner countries may decide to unify their economic policies. This unification can have various degrees, going from the harmonization of a limited range of policies up to the complete unification of all economic policies (including monetary policy, possibly with a common currency). In these cases we are in the field of international economic integration, possibly leading to an economic and monetary union. In the older literature preferential trading cooperation was often called (a form of) international economic integration, but to avoid terminological confusion we do not use this definition.

According to these classifications the EEC (European Economic Community), even before its transformation into the European Union, although article 9 of the founding treaty (Treaty of Rome, 1957) stated that the Community was founded upon a customs union, more properly belonged, at least in theory, to the category of economic unions, since it involved unification of some economic policies (agricultural policy, for example).

We finally note that, as in the world there are several preferential trade cooperation arrangements, the issue of their interrelations arises. This gives rise to complex problems, for example of the "hub-and-spoke" type. This term refers to arrangements that give one region (the hub) better access to other regions (the spokes) than these have to one another (Baldwin, 1994; Kowalczyk & Wonnacott, 1992; Krugman, 1993b). For example, as a consequence of association agreements between the European Union and several CEECs (Central and East European Countries), bilateral trade liberalisation between the EU and each of these CEECs has taken place, but trade barriers between the CEECs have remained. For these problems we refer the reader to the cited authors.

11.7.2 The Effects of a Customs Union

In this section we shall deal mainly with the theory of customs unions but most of the analysis can be applied to other forms of trade cooperation. In general, it might seem that, since a customs union represents a step towards the ideal situation of free trade, it will improve social welfare. But this is not the case: as we know from the theory of second best (Sect. 11.6), when the Pareto-optimum conditions are violated, the elimination of part of these violations does not necessarily bring about an improvement. We must therefore examine the effects of the formation of a customs union more closely. Viner (1950), in examining these effects on the production side, introduced the distinction between *trade creation*, which represents an improvement in resource allocation, and *trade diversion*, which, on the contrary, represents a worsening in this allocation.

Table 11.1 Effects of a customs union

Commodity	Cost	Country 2 (exp to 1)		Country 1		Country 3 (exp to 1)	Effects
A	Cost	12		14		10	
	Cost + tariff before the union	15.6		14	←	13	
	Cost + tariff after the union	12	→	14		13	Trade diversion
B	Cost	14		11		15	
	Cost + tariff before the union	18.2		11		19.5	Neither diversion
	Cost + tariff after the union	14		11		19.5	Nor creation
C	Cost	12		15		13	
	Cost + tariff before the union	15.6		15		16.9	
	Cost + tariff after the union	12	→	15		16.9	Trade creation

Trade creation refers to the fact that, as a consequence of the elimination of tariffs (in this section, for brevity, "tariffs" indicates "tariffs and other barriers to trade") within the union, a commodity—which before the union was produced domestically by each partner country and not traded because of tariffs—is now traded and so is produced by that partner country which is most efficient in its production. This brings about a better allocation of resources.

Trade diversion occurs when the elimination of tariffs within the union induces a partner country to import a commodity from another partner country instead of from a country outside the union as it did before, because, though the latter is the most efficient in producing the commodity, it is no longer competitive on account of the tariff, which has been maintained against it. This leads to a worse allocation of resources.

Better to explain these effects, we must consider at least three countries: two which form a union and a third representing the rest of the world. The following numerical example may be helpful. Consider two countries, 1 and 2, forming a customs union, whilst country 3 remains outside, and three commodities A, B, C. The arrows in Table 11.1 represent the direction of trade flows; no arrow means no trade. The productive efficiency is measured in terms of the unit cost of production (a common unit is used) in the absence of tariffs; for simplicity this cost is assumed constant. Before the customs union, country 1 applied a 30 % tariff on all imports, whilst after the union it keeps the tariff on imports coming from country 3 and eliminates it on imports from country 2. Let us now consider the effects of the union with reference to country 1.

As regards commodity A, the most efficient country is country 3, where the unit cost is lowest. Before the union, country 1 imports commodity A from country 3, as its price, even with the tariff, is lower than the domestic cost of production

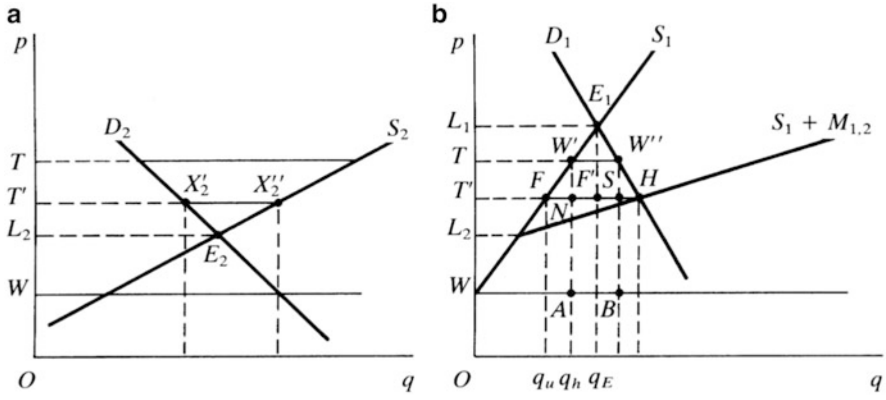


Fig. 11.7 Effects of a customs union

(13 instead of 14). After the union, country 1 imports the same commodity from country 2, because its cost is 12, lower than 13: the lower efficiency of country 2, with respect to country 3, in producing *A* is more than offset by the tariff schedule. Therefore, the union causes a less efficient allocation of resources (trade diversion). As regards commodity *B*, the most efficient country is country 1: the formation of the customs union, therefore, does not change the fact that, for this country, it is better to produce *B* domestically rather than to import it. The situation for country 1 is the same both before and after the union and the union has no effect on its trade.

Finally, the presence of a prohibitive tariff prevented country 1 from importing commodity *C*; the formation of a union with country 2, which is the most efficient in producing *C*, brings about a better allocation of resources, as country 1 now imports this commodity from country 2 (trade creation).

This analysis considers only the *production* effects of the union, but Johnson (1960) and others have rightly observed that, to evaluate the consequences of the formation of a customs union, one must also consider the *consumption* effects and, more precisely, the effects on consumers' surplus. Thus we also have *trade creation* and *diversion* from the point of view of *consumption*. The former derives from the fact that consumers substitute cheaper foreign goods (imported from a member country) for more expensive domestic goods, and so benefit from an increase in consumers' surplus. The latter derives from the fact that consumers' surplus decreases as a consequence of consumers having to substitute more expensive foreign goods (imported from a member country) for formerly cheaper goods (previously imported from a country remaining outside the union) which are now non-competitive because the union has decided to raise tariff rates with respect to non-member countries.

If we add the effects on production and consumption together, we have trade creation and diversion in the *broad sense*. These broader concepts of trade creation and diversion can be illustrated in a partial equilibrium framework by using a diagram. In Fig. 11.7a we have drawn country 2's domestic demand and supply curves for a certain commodity whilst Fig. 11.7b depicts country 1's domestic demand and supply curves for the same commodity. We then calculate, for any given price, the excess supply $S_2 - D_2$, that is, country 2's supply of exports which, in the case where countries 1 and 2 contemplate the formation of a customs union, has to be added to country 1's domestic supply, giving rise to the curve $S_1 + M_{1,2}$ in Fig. 11.7b. This curve originates from S_1 at the point corresponding to country 2's domestic equilibrium price, OL_2.

For simplicity's sake the supply price of the commodity by the rest of the world is assumed constant and is—in the absence of tariffs—equal to *OW*.

Let us now consider various cases, following Robson (1998).

(a) If before the union both country 1 and country 2 levied a prohibitive tariff, the domestic prices were OL_1 and OL_2 respectively. After the union both countries levy a tariff at the same rate (for example an average of the pre-union rates) against the rest of the world, so that the domestic price in both countries is established at a common level, intermediate between OL_1 and OL_2 for example OT. This, however, is not an equilibrium price, as country 2's excess supply is greater than country 1's excess demand; the price, therefore, decreases to OT', where $X_2'X_2'' = FH$, i.e. the combined supply $S_1 + M_{1,2}$ is equal to the demand D_1.

To examine the effects of the customs union we can use the concepts of producers' and consumers' surplus employed in Sect. 10.3. In country 1, as a consequence of the decrease in the domestic price from OL_1 to OT', consumers' surplus increases by the area $T'L_1E_1H$ whilst producers' surplus decreases by the area $T'L_1E_1F$ (domestic producers have had to reduce output from Oq_E to Oq_u as a consequence of the decrease in the domestic price). The net benefit is given by the area FE_1H, which can be divided in two parts. Area FE_1F' represents the production effect, that is the decrease in costs due to the fact that the quantity q_uq_E is imported at a cost (that of country 2) lower than that of producing it at home; this is the *production part of the trade creation effect*. Area $F'E_1H$ represents the *consumption part of the trade creation effect*. The sum of the two constitutes the trade creation effect of the union.

Country 2's domestic price increases from OL_2 to OT', so that there is a decrease in consumer's surplus equal to area $L_2T'X_2'E_2$. But the increase in producers' surplus (area $L_2T'X_2''E_2$) is greater, so that the union's net effect is favourable to country 2 as well.

As regards the rest of the world, the situation is unaltered, as its trade with countries 1 and 2 was nil both before and after the union between these two countries.

We can therefore conclude that in the case examined the formation of the union is unequivocally beneficial.

(b) A second case occurs when before the union only country 2, and not country 1, levied a prohibitive tariff. Let us then assume that in country 1 the pre-union tariff rate was such as to give an absolute unit amount equal to WT, so that domestic output was Oq_h and imports (coming from the rest of the world) were $W'W''$. Tariff revenue was $W'T \times W'W''$, that is, equal to area $W'ABW''$.

A customs union is now formed between countries 1 and 2, and the common tariff rate against the rest of the world will be intermediate between the pre-union tariff rates of the two countries, for example such as to give an absolute unit amount equal to WT', so that the domestic price in both countries changes to OT'. Country 1 now imports FH of the commodity under consideration from country 2 and produces Oq_u of it domestically. Consumers' surplus increases by area $T'TW''H$, producers' surplus decreases by area $T'TW'F$, the government's tariff revenue disappears, that is, decreases by the whole area $W'ABW''$. This last area can be divided into two parts, as $W'ABW'' = W'NSW'' + NABS$. The balance between benefits and costs can then be reduced graphically to the comparison between areas $FW'N$ and $SW''H$ on the one hand, and area $NABS$ on the other.

Area $FW'N$ represents the production part of the trade creation effect, due to the saving on production cost that derives from the fact that the quantity $q_uq_h = FN$, instead of being produced at home, is imported at a lower cost (that of country 2). Area $SW''H$ represents the consumption part of the trade creation effect. Area $NABS$, on the contrary, represents a trade diversion on the side of production, due to the fact that the quantity of imports $W'W'' = AB$, which prior to the union came from the rest of the world, now comes from country 2, with an additional cost in terms of resources equal to the difference between OT' (the supply price, i.e. the marginal cost, of the commodity in country 2) and OW (the supply price, i.e. the marginal cost, in the rest of the world); this difference is $T'W = NA$.

The diagram shows that in the case under examination the balance between country 1's benefits and costs is unfavourable, but of course, in general, the opposite outcome is possible.

As regards country 2, the effects are the same as in case (a), so that this country gains from the union. Therefore, if we were willing to accept the inter-country comparability of the monetary measures of the various effects (expressed in a common unit), we could calculate the algebraic sum $FW'N + SW''H - NABS + X_2'X_2''E_2$ and ascertain whether the union is on the whole beneficial or harmful. Note that we have not included in this calculation the effects on the rest of the world, which sees its net exports drop by $W'W''$. However, as we have assumed that the supply curve of the rest of the world is perfectly elastic, for a first approximation these effects can be ignored.

Those illustrated are only two out of practically unlimited possibilities: the reader can construct other examples *ad lib* and analyse them by way of the same graphical technique. The fact that it is not possible to demonstrate general propositions (except the purely negative one that it is impossible to state any precise result, as anything can happen) is by now obvious if one refers to the theory of second best.

Since it would not be possible to reach definite general conclusions even if one examined the effects of a customs union in a general equilibrium setting, for the same motives related to the theory of second best, we omit the general equilibrium analysis of customs unions (for which see, e.g., Kemp, 1969a; Lipsey, 1970; Lloyd, 1982). It is however possible to give some indications of a probabilistic type (thus likely to be sometimes wrong, sometimes correct). These indications are that a customs union will be more likely to produce beneficial effects:

(i) The greater is the degree of competitiveness among member countries, i.e. the greater the number of similar goods they produce. In such a case, in fact, due to the differences in productive efficiency, each country will expand its comparatively more efficient industries and contract the comparatively less efficient ones; thus there will be more scope for trade creation without much trade diversion from other countries;
(ii) The higher are the initial tariffs between the countries forming the customs union: in fact, the gain deriving from the elimination of these tariffs will be larger;
(iii) The lower are the tariffs with the outside world: trade diversion, in fact, will be less likely;
(iv) The wider is the union, as this increases the probability that trade creation effects will override trade diversion effects (in the extreme case, if the union includes all the world, we have free trade and no trade diversion can occur).

So far the analysis has been of a (comparative) *static* type; in addition to this, the theory of customs unions also examines the *dynamic* benefits of a union; the main benefits are:

1. The increase in the size of the market made possible by the union allows the industries producing traded goods to enjoy the fruits of economies of scale;
2. The elimination of protection with respect to member countries brings about an increase in competition;
3. The fact that the member countries together negotiate the tariffs with the rest of the world, gives them greater bargaining power.

In addition to the possible economic gains so far examined, there are gains of a political nature, which are outside the scope of this treatment, but which, like the non economic motives for protection, may warrant the formation of a union (or the entry into an existing one) even if the strictly economic benefits are not positive.

11.7.3 Empirical Problems

In concluding this treatment it is as well briefly to mention the methods used for the empirical estimation of the effects of economic integration.

A first distinction is between *ex ante* and *ex post* estimates. *Ex ante* estimates aim to evaluate the future effects of a prospective economic union (in what follows we use the term economic union to indicate any one of the five categories of economic integration listed at the beginning) or of the entry of new members into an already existing economic union. In this case the data concerning the existing pre-union situation is known and one has to estimate the hypothetical result of the prospective integration, on which, naturally, no data is available. Ex post estimates aim to evaluate the effects of an already existing economic union. Although in this case the problem might seem simpler, as the post-integration data is known, it should be pointed out that the problem is to ascertain to what extent the events *observed* are due to the union and to what extent they would have come about (even) in its absence. One must, in other words, compare a known situation (the events observed) with an unknown and hypothetical one (what would have happened if the union had not been formed). This is the usual problem that derives from the impossibility, in economics, of carrying out experiments under controlled conditions.

A second distinction is based on the methods used for estimating the hypothetical alternative, which are principally three. The *direct* method consists in using a precise analytical model, the parameters of which are estimated econometrically; simulation procedures are then used to produce the alternatives. The *survey* (or *delphic*) method consists in assessing the views of the experts, for example by asking the managements of the firms how they expect the sales in the domestic market and in the markets of the partner countries to change as a consequence of the modification in the trade barriers. The *indirect* method consists in projecting the pre-integration trade flows into the post-integration period, then calculating the effects of the economic union as the difference between actual and projected flows (so-called residual imputation).

Many empirical studies were carried out as regards the EEC (European Economic Community, now transformed into the European Union); the reader interested in these can consult, for example, Robson (1998, chap. 12), Grinols (1984), and Winters (1989). The results of different studies are often themselves different: for instance, various studies carried out around 1970 on the effects of UK entrance in the EEC yielded *ex ante* estimates all indicating a net cost, but varying from 453 to 1,144 million pounds (1969 prices). *Ex post* estimates (Grinols, 1984), for the UK and for the period 1972–1980, indicated, again, a net cost amounting to about 1.5 % of GDP.

For estimates concerning the United States and NAFTA (North American Free Trade Agreement) see Bhagwati and Panagariya (1996) and Krueger (2000).

11.8 The Main Cases of Preferential Trading Cooperation

11.8.1 The European Common Market (Now European Union)

The European Economic Community (EEC) was founded with the Treaty of Rome signed in 1957. The founding countries were West Germany, France, Italy, Holland, Belgium, Luxembourg. At the beginning it contemplated a common external tariff; the complete liberalization of trade in industrial goods among the members took place only in 1968. The still existing non-tariff barriers were eliminated in 1986, thus realizing a true customs union. Free factor mobility among the members was realized subsequently, first that of capital and then (1993) that of labour, thus giving rise to a true common market.

Over the years other countries joined the initial 6, reaching the number of 27 (Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Holland, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Poland, Portugal, Romania, Slovenia, Slovakia, Spain, Sweden, United Kingdom), and negotiations are under way with other countries (Croatia, Iceland, Macedonia, Turkey) for their admission. In the meantime the name was changed, from EEC to EU (European Union).

The European Union is something more than a common market, because it contemplates various measures of coordination of the members' policies (the best known is the common agricultural policy) and other interventions to homogenize the economies of the member countries. Details can be found in the EU's site, http://europa.eu.int

11.8.2 NAFTA

The North Atlantic Free Trade Association is a free trade area formed in 1993 among the United States, Mexico, and Canada, aimed at the elimination not only of tariffs, but also of non-tariff barriers to the circulation of commodities and services. The possibility is also contemplated for each member country to invest capital in any other member, hence NAFTA is something more than a mere free trade association. The official site is http://www.nafta-sec-alena.org

Box 11.1 European Economic Integration

The European Internal Market became a reality in 1993. Since then the EU countries have experienced convergence in terms of consumer prices though significant differentials remain in some areas. There was a period of widening prices in the mid-1990s, but overall the tendency is clear, and with the price transparency, and with the elimination (thanks to the euro) of currency conversion costs and exchange rate risk, the trend can be expected to continue. However, other costs of trading (such as transport) remain, so significant variations in prices can be expected to remain within the euro area, especially in sectors which are less exposed to trade. The wide wage discrepancies prevailing among EU countries at the start of the integration have considerably decreased. As regards the prices of capital (interest rates), convergence has been observed thanks to the creation of EMU.

To evaluate the economic effect of a regional trade agreement such as the EU on the partners and third countries, theory—as we know—focuses on the concepts of trade creation (switching of imports from a high-cost origin to a low-cost origin) and trade diversion (switching of imports from a low-cost source to a high-cost source) that can be measured by the share of intra-union versus extra-union trade. As a general rule, the greater the absolute growth of extra-union trade, the less the danger of trade diversion. In the EU's case, the share of intra-EU trade in total trade has risen from 42 % in 1961 to 64 % in 2010. The increase in the intra-EU trade share was accompanied by a rapid absolute growth of extra-EU trade. This indicated that trade creation dominated trade diversion. Tsoukalis (1997) argued that overall trade creation dominated in manufactured goods and overall trade diversion in agricultural goods. The latter is the result of the Common Agricultural Policy, which has protected EU agriculture from foreign competition.

11.8.3 MERCOSUR

MERCOSUR (acronym from the Spanish Mercado Común del Sur, Common Market of the South; or MERCOSUL, acronym from the portuguese Mercado Comum do Sul) is an agreement signed in 1991 by some Latin-American countries, originally Argentina, Brazil, Paraguay, Uruguay, which in 1996 were joined by Bolivia and Chile. This agreement aims at the formation of a common market among these countries. The official site is http://www.mercosur.org.uy

11.8.4 ASEAN

The Association of Southeast Asian Nations was formed in 1967 for political reasons, to defend the member countries against the then communist Indochina; the original members were Thailand, Malaysia, Singapore, Indonesia, the Philippines. ASEAN was subsequently transformed into a preferential trading association with the intention of moving on to a customs union and then to a common market. The founding countries have been joined over the years by Brunei Darussalam, Cambodia, Laos, Myanmar, Vietnam. The official site is http://www.aseansec.org

11.8.5 FTAA

The Free Trade Association of the Americas (or ALCA, from the Spanish Área de Libre Comercio de las Américas) was proposed in the Miami conference held in 1994 among 34 countries of the Americas. It aims at giving rise to a free trade area that will eliminate all barriers to trade and investment flows. This area will not be in competition with other existing agreements (such as NAFTA). The official site is http://www.ftaa-alca.org

References

Baldwin, Richard E. (1994). Towards an integrated Europe.
Bhagwati, J. N. (1982a). Directly unproductive, profit-seeking (DUP) activities.
Bhagwati, J. N. (1995). U.S. trade policy: The infatuation with free trade agreements.
Bhagwati, J. N., & Panagariya, A. (Eds.), (1996). *The economics of preferential trade agreements.*
Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (Pt. III).
Chacholiades, M. (1978). *International trade theory and policy* (Pt. VII).
Corden, W. M. (1971a). *The theory of protection*.
Corden, W. M. (1974). Trade policy and economic welfare.
Corden, W. M. (1984b). The normative theory of international trade.

Deardoff, A., & Stern, R. M. (1984). The effects of the Tokyo round on the structure of protection.
El-Agraa, A. M. (1984). *Trade theory and policy*.
Ethier, W. J., & Horn, H. (1984). A new look at economic integration.
Greenaway, D. (1983). *International trade policy: From tariffs to the new protectionism*.
Grinols, E. L. (1984). A thorn in the Lion's Paw: Has britain paid too much for common market membership?
Hicks, J. R. (1981). *Wealth and welfare* (pp. 100–132).
Johnson, H. G. (1953). Optimum tariffs and retaliation.
Johnson, H. G. (1960). The economic theory of customs union.
Jones, A. J. (1979). The theory of economic integration.
Kemp, M. C. (1969a). *A contribution to the general equilibrium theory of preferential trading*.
Kemp, M. C. (1976). *Three topics in the theory of international trade: Distribution, welfare and uncertainty* (Pts. I and III).
Kowalczyk, C., & Wonnacott, R. J. (1992). Hubs and spokes, and free trade in the Americas.
Krauss, M. B. (1972). Recent developments in customs union theory: An interpretive survey.
Krueger, A. O. (1974). The political economy of the rent-seeking society.
Krueger, A. O. (2000). NAFTA's effects: A preliminary assessment.
Krugman, P. R. (1993b). The hub effect: or, threeness in interregional trade.
Lipsey, R. G. (1960). The theory of customs unions: A general survey.
Lipsey, R. G. (1970). *The theory of customs unions: A general equilibrium analysis*.
Lloyd, P. J. (1974). A more general theory of price distortions in open economies.
Lloyd, P. J. (1982). 3×3 theory of customs unions.
Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: Theory and evidence.*
McMillan, J. (1989). *Game theory in international economics*.
Meade, J. E. (1952). *A geometry of international trade* (chap. VI).
Meade, J. E. (1955). *The theory of international economic policy* (Vol. 2).
Pearce, I. F. (1970). *International trade* (Book I, chaps. 7 and 9).
Puga, D., & Venables, A. J. (1997). Preferential trading arrangements and industrial location.
Robson, P. (1998). *The economics of international integration*.
Swann, D. (2000). *The economics of Europe: From common market to European union*.
Takayama, A. (1972). *International trade: An approach to the theory* (Pt. V).
Tsoulakis, L. (1997). *The new European economy revisited*.
Viner, J. (1950). *The customs union issue*.
Vousden, N. (1990). *The economics of trade protection.*
Winters, L. A. (1989). Britain in Europe: A survey of quantitative trade studies.

Chapter 12
The "New" Protectionism

12.1 Introductory Remarks

The typical instruments of the "old" protectionism are tariffs and (non-discriminatory) import quotas. The last few decades have seen a progressive reduction of these traditional trade barriers: GATT and WTO have provided a negotiating framework for such a reduction and outlawed the use in general of import quotas, as well as established the extension to all members of the MFN (Most Favoured Nation) treatment (see Sect. 10.2). Up-to-date information is contained in the annual publication World Tariff Profiles, that can be downloaded from the WTO site (www.wto.org).

However, notwithstanding the dramatic decrease in average world tariffs, protectionism is still around under new forms.

12.2 Why the New Protectionism?

In parallel with the decline of the old protectionism (see above), the last decades have witnessed the emergence of a "new" protectionism or neoprotectionism, based on the type of non-tariff barriers (NTB) exemplified in items (a) through (i) listed in Sect. 10.6.4. (see, e.g., Laird and Yeats, 1990: Schucknecht, 1992; Vousden, 1990)

The common feature of these barriers is that they are less overt and more subject to discretion than the instruments of the old protectionism. Several reasons have been set forth in the literature to explain this trend:

1. The countries members of GATT did agree not to use discriminatory tariffs and quantitative import restrictions, except in special circumstances contemplated by the GATT Articles (e.g. to relieve temporary balance-of-payments pressures, and for the emergency protection of domestic industry). By using the instruments of the new protectionism, GATT members avoided a clash with the letter of the

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_12,

GATT rules (although, of course, these instruments clashed with the spirit of GATT).
2. The barriers under consideration are politically much easier to implement. In fact, the traditional measures (tariffs and quotas) must be implemented through either legislative acts or highly transparent administrative channels. The measures of the new protectionism, on the contrary, can often be negotiated in secret: a typical example is that of voluntary export restraints.
3. For the reason given under (2), pressure groups lobbying the government for protection find it more convenient to ask for measures belonging to the "new" rather than to the "old" protectionism.

This brings us to the question of how protectionist measures are *actually* introduced, a question that, largely neglected in the old theory, is given a lot of attention in the new theory. Now, in reality, protection is usually sought for by interested domestic industries through the lobbying of politicians or the use of administered protection procedures. The difference is that in the former case the possible introduction of a protective measure is a matter of political discretion, while in the latter it is the result of a codified administrative procedure aimed at remedying an alleged injury. These topics will also be dealt with in the present chapter.

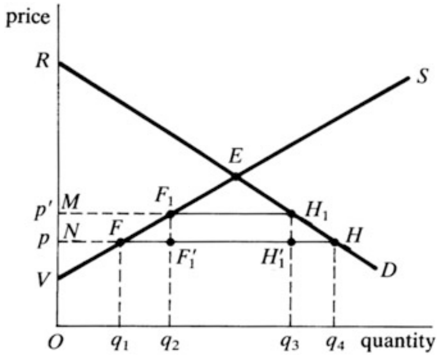
12.3 Voluntary Export Restraints and Import Expansion

A voluntary export restraint (VER) is an agreement negotiated between the exporting and the importing country, whereby the exporting country "voluntarily" curtails exports to the importing country. Another name under which such agreements are also presented in order to avoid conflict with the letter of the GATT articles is "orderly market agreements". A VER is really an alternative, for the exporting country, to the imposition of a tariff or a quota by the importing country.

Since the outcome of a VER is a quantitative reduction in the amount of goods that the importing country receives from the exporting country, the effects on the former country can be analysed—in the small-country, partial-equilibrium context—by the same diagram developed for the analysis of quotas (see Fig. 10.7, that we reproduce here as Fig. 12.1 for the reader's convenience). Suppose that as a consequence of a VER the imports of the home country fall from q_1q_4 to q_2q_3. The effects on the home country's price and output would then be the same as under a quota (Greenaway, 1983, chap. 7).

What is completely different is the destination of the sum represented by the area $F_1F_1'H_1'H_1$. Under a quota this is a gain accruing to importers (unless the government auctions off the licenses). But since a VER is by definition administered by the foreign country, the sum under consideration accrues to this country. Thus this part of the reduction in domestic consumers' surplus is not offset by a redistribution to domestic importers or authorities, but is redistributed abroad. The fact that this "rent" from VER protection accrues to the exporting country is clearly an important

Fig. 12.1 Effects of a VER



motive for this country to prefer a VER to the imposition of a tariff or quota by the importing country. This is an economic reason (other non-economic reasons have been given in Sect. 12.1) for the widespread acceptance of VERs in the place of the measures of the "old" protectionism.

But there is another substantial difference between a VER and a quota, that occurs when the export goods subject to a VER become monopolized. Cartelisation, in fact, may be an outcome of a VER (Murray, Schmidt, & Walter, 1978). This happens, for example, when the government of the exporting country leaves to the industry affected by the VER to decide on the license allocation. Let us assume, for simplicity's sake, that the cartel controls the whole supply of the good. Then we can use the same diagram developed for the study of the monopolistic cartel (see Fig. 10.9, that we reproduce here for the reader's convenience).

Let us then consider the pre-VER and pre-cartel situation of free competition and no barriers to trade. Let D be the demand curve of the importing country for the commodity in question and MC the partial equilibrium supply curve of exports by the exporting country (under perfect competition, the marginal cost curve represents the industry's supply curve). Then, as we know, the equilibrium quantity and price will be determined by the intersection of D and MC (point L). Suppose now that a quota is introduced limiting the quantity to somewhat less than the perfect competition quantity but somewhat more than the monopoly quantity (point q_E in Fig. 12.2). The price can be read off as the ordinate of D corresponding to the quantity fixed by the quota.

We now consider the introduction of a VER having the same quantitative limitation as the quota. As a consequence of the VER a monopolistic cartel is formed, and so the cartel's equilibrium point will be H in Fig. 12.2: we see that *it is perfectly rational for the profit-maximising cartel to export less than the quantity stated in the VER agreement*! It follows that a VER, as compared with a quota involving the same quantitative restraint, may result in the actual volume of imports being lower and price higher. This, in turn, gives a greater protection (because of the higher domestic price) to the import substitute sector.

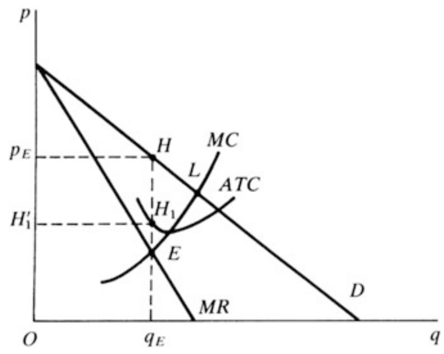


Fig. 12.2 VER and monopolistic cartel

For the analysis of all the differences between a quota and a VER under different market conditions and different degrees of participation in the VER, see Takacs (1978) and Hillman and Ursprung (1988). The political preferability of VERs is examined by Jones (1984) and Dinopoulos and Kreinin (1989). Pomfret (1989) gives a survey of the economic consequences of VERs. For a general equilibrium analysis of VERs (which substantially confirms the results of the partial equilibrium analysis used here) see Herberg (1990).

A trade policy tool that can be used as an alternative to a VER is a *voluntary import expansion* (VIE). Rather than voluntarily restricting exports from country 2 to country 1, trade agreements between the two countries can take the form of country 2 voluntarily increasing imports from country 1. A VIE sets a minimum market share for imports, hence symmetrically sets a maximum share for the domestic producer. To reduce its market share to the level required by the VIE, the domestic firm increases its price, which induces an increase in the foreign firm's equilibrium price.

According to some authors, VIEs are to be preferred to VERs because, while the latter are intended to restrict trade, the former are, on the contrary, designed to increase trade by increasing foreign sales in countries where structural impediments and policies restrict access to foreign suppliers. Actually, the US-Japan trade negotiations seem to have shifted from trying to limit the access of Japanese firms to the US market to trying to increase the access of American firms to the Japanese market, namely from agreements based on VERs to agreements based on VIEs.

For a theoretical analysis of VIEs see Bhagwati (1987) and Greaney (1996).

12.4 Subsidies

Subsidies can be present in both the export and the import sector. As regards the export sector, the subsidy can be either an export subsidy (i.e., given to domestic producers only on the exported part of their output) or a production subsidy (i.e., given to domestic producers on their whole output). Let us begin by

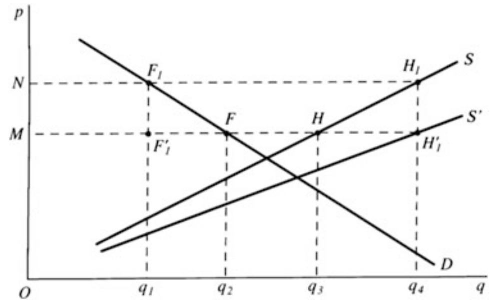


Fig. 12.3 Effects of production and export subsidies

considering an *export subsidy*. In Fig. 12.3, D and S represent as usual the domestic partial equilibrium demand and supply curves. With free trade, given the ruling international price OM, domestic price is the same, and exports are $FH = q_2q_3$. Suppose now that an export subsidy, say MN per unit of output exported, is given to domestic firms. The situation is perfectly symmetrical to that of an export duty (Sect. 10.6.2). The *domestic* price increases from OM to ON: domestic producers, in fact, receive ON per unit of the commodity exported, and will not be willing to serve the domestic market unless they receive the same price. If we exclude the possibility of re-importing the commodity to the domestic market at the world price OM, the domestic price will be driven to ON. Thus domestic producers will sell NF_1 in the domestic market at the price ON and export F_1H_1 at the prevailing world price OM, but actually getting ON given the subsidy MN. The total amount of the subsidy that they receive is thus the area $F_1F_1'H_1'H_1$.

Benefits and costs can be calculated using the concepts explained in Sect. 10.3 as regards an import duty. Producers' surplus increases by the area MNH_1H. Consumers' surplus decreases by the area MNF_1F. The government has to pay an amount $F_1F_1'H_1'H_1$ (note that the area $F_1F_1'F$ appears twice among the costs). Hence the net welfare cost of the export subsidy is the sum of the two triangles $F_1F_1'F$ (the consumption cost) and $H_1H_1'H$ (the production cost). These have the same interpretation as in the case of a tariff (Sect. 10.4).

The case of a *production subsidy* can also be examined using Fig. 12.3. Since this subsidy is given to domestic producers on their whole output, the result is an equiproportional shift downwards of the supply curve (from S to S') by a percentage equal to the (ad valorem) subsidy. In Fig. 12.3 we have assumed a production subsidy of the same percentage as the export subsidy. This is shown by the fact that at output level Oq_4 the vertical distance between S and S' is $H_1H_1' = MN$, denoting that the subsidy to producers is the same per unit as in the case of the export subsidy. The cost to the government is now higher: since the subsidy is given on all domestic output, the total amount is MNH_1H_1'. But now there is no decrease in consumers' surplus, since the price remains at OM. Producers' surplus increases by MNH_1H, as before; hence the net cost is now only the triangle $H_1H_1'H$ (the production cost). Thus it appears that a production subsidy (which creates no wedge between the domestic and the international price) is preferable to a direct trade intervention like an export subsidy (which creates such a wedge).

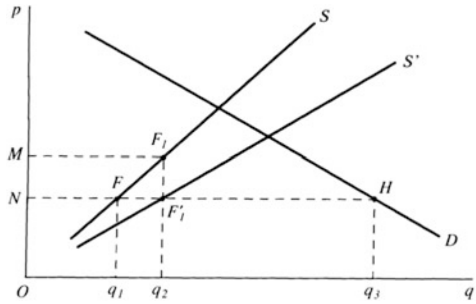


Fig. 12.4 Effects of subsidies to the import-competing sector

Let us finally consider a subsidy to the domestic sector producing import-competing goods (Fig. 12.4). This is actually a production subsidy, hence the supply curve of domestic producers (S) shifts downwards equiproportionally by a percentage equal to the (ad valorem) subsidy (S').

The effect of the subsidy is that, for any output, the price received by producers is greater than the price paid by consumers by the amount of the subsidy. Let us assume a world price *ON* and a subsidy such that the amount received by domestic producers on every unit of output is *MN*. Hence domestic producers will be able to supply $NF_1' = Oq_2$ instead of $NF = Oq_1$. Consumers continue to pay *ON* per unit, but producers receive *OM*. The outcome of this protective measure is that imports fall from q_1q_3 to q_2q_3.

If we now make the usual cost-benefit analysis, we see that there is no decrease in consumers' surplus, since they continue to pay the same price as before. The only cost is the government's outlay for the subsidy, namely area $MNF_1'F_1$. On the side of benefits we have the increase in producers' surplus, which is $MNFF_1$. The balance is a net cost of the subsidy equal to the triangle $FF_1'F_1$ (the production cost). It follows that, if we compare a subsidy to the import-competing sector with a tariff (which entails both a production and a consumption loss), we conclude that the subsidy is preferable to the tariff. *Ceteris paribus*, a tariff creates two distortions (on both the production and the consumption side) while a subsidy only creates one distortion (on the production side).

Let us now put this result together with the previous result, that a production subsidy on the side of exports is preferable to an export subsidy (which can be taken as a negative tariff). The conclusion is that production subsidies are to be ranked above tariffs. This conclusion lies behind the suggestions given in Sects. 11.3 and 11.4, that a subsidy to domestic production is a better policy than a tariff.

12.5 The Political Economy of Protectionism

The previous treatment explains why countries may prefer the instruments of the new protectionism rather than the traditional ones. But it does not explain why protectionism in general is still around. We know quite well from the theory of

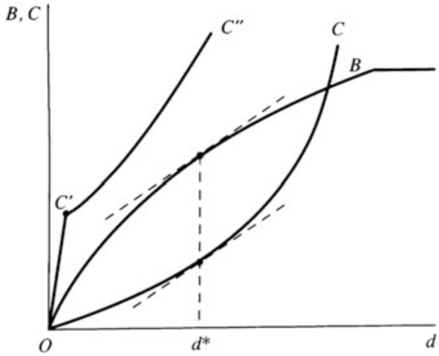


Fig. 12.5 The optimal amount of lobbying

second best (Sect. 11.6) that, when we are in the presence of many violations to the Pareto-optimum conditions, it is not possible to say in general whether the introduction or elimination of a violation (such as a protectionist measure) decreases or increases social welfare. Hence the continuing presence of protectionism would seem to imply that the above uncertainty has been solved in the sense that social welfare increases in the presence of protectionism. This is certainly not the case: no theoretical study exists in this sense.

Thus we must look elsewhere, and the political economy of protectionism offers an interesting answer. This school of thought starts from the observation that protectionist measures are not introduced (or eliminated) by a benign, omniscient government aiming at the maximization of social welfare. Rather, they are the result of pressure groups lobbying the government for particular policy changes (see, e.g., Jones and Krueger, 1990; Lopez and Pagoulatos, 1994; Markusen et al., 1995). Hence, protectionism can be interpreted as a *rational* policy for decision makers in a democracy.

12.5.1 The Demand for and Supply of Protection

The observation of actual decision making in a democracy suggests that there exists a *political market* for protection, where there is a *demand* for, and a *supply* of protection (Baldwin, 1982; Brock & Magee, 1978; Frey, 1984). The demand for protection comes from particular groups of voters, firms, and associated interest groups. The supply comes from politicians and government officials.

Let us begin with the demand side. The economic agents who will gain from protectionist measures invest resources in order to influence political decisions in their favour. Hence the situation can be examined in the context of cost-benefit analysis, as shown in Fig. 12.5 (Baldwin, 1982; Frey, 1984). The amount of protection is measured by the variable d, that for simplicity we take as the tariff rate but can be any other protectionist measure. Benefits (B) and costs (C) of lobbying are measured in money terms. The cost-of-lobbying curve OC is drawn on the assumption of increasing marginal costs, since it is reasonable to expect that it becomes more and more difficult to obtain higher and higher tariff rates from the government. The benefits-from-lobbying curve *OB* increases up to a maximum, which corresponds to the prohibitive tariff. It is not inconceivable that increasing tariff rates might yield increasing marginal benefits over a certain range, but for simplicity's sake *OB* has been drawn assuming decreasing marginal benefits everywhere. As in any case of cost-benefit analysis with well-behaved curves, the net benefit is maximised where the marginal benefit equals the marginal cost. This gives the associated tariff rate d^* (endogenously determined), where the slopes of the *OB* and *OC* curves are equal, and the vertical distance between these two curves (i.e., the net benefit) is highest.

The figure also shows that a lobbying activity is not always worthwhile. The $OC'C''$ cost curve shows that the initial costs of lobbying may be so high that the curve lies above the benefits curve everywhere. This cost situation may occur when the interest groups benefiting from protection are difficult to organise, and no organisation for other purposes (e.g., for social gatherings) already exists that could be used for setting up the lobby, thus avoiding the initial cost OC'. This explains why protection is not "demanded" by everybody and why the interest groups that are already organised tend to get additional advantages, while newcomers are in a difficult position in the political market for protection.

Before turning to the supply side, it should be observed that, in addition to the interest groups gaining from protection, there are groups losing from protection. Pro-tariff groups mainly consist of firms (including the workers) producing import-competing goods. These have strong lobbies, because such groups are usually well organised. On the contrary, anti-tariff groups (typically the consumers and the exporters) have weak lobbies as they find it difficult to organise (see Olson, 1964).

We now consider the *supply of protection*. The protectionist measures of a country are determined by politicians (typically the government) and by government officials (even if they are not entitled to decide the introduction of a protectionist measure, bureaucrats prepare, formulate and implement trade bills). A government has certain ideological goals (amongst which there may be a specific position with regard to free trade or protection), but also has a number of other goals, amongst which the need or desire of being re-elected. Since the interest groups demanding protection are much better organised than the anti-protection ones and have greater lobbying power (which includes financial help for the election campaign), a government will pay more attention to them. Furthermore, a government also has constraints, such as the budget and the balance of payments. A high balance-of-payments deficit may induce protectionist measures, and a budget deficit may be an additional element for a tariff (which gives a revenue to the government).

As regards the bureaucrats, it has been argued (Messerlin, 1981) that they favour greater protection than politicians. Amongst the various reasons there is the fact that they must reach their goals (which are not the common interest or collective welfare, but the maximization of their utility function) by using the instruments available to them (which are more limited than those available to politicians) more intensely.

Actual tariff rates are the outcome of the interaction between the demand and the supply in the political market for protection. Various models exist for the analysis of this interaction (Brock & Magee, 1978; Brock, Magee, & Young, 1989; Findlay & Wellisz, 1983; Mayer, 1993).

Brock and Magee, for example, consider the case of two lobbies, one pro-tariff and the other anti-tariff, and two political parties. The pro-tariff lobby is better organised and has more money but less votes than the anti-tariff lobby. Hence there is a trade-off between the number of votes that the lobby can offer to politicians and the amount of money that the lobby can give the politicians to finance the electoral campaign. The parties want to maximise the probability of re-election by choosing an appropriate position on the free trade-protectionism issue. Each party knows that it can obtain less votes but more financial resources (which in turn can be used to obtain more votes through the electoral campaign) by taking a position in favour of protectionism, and vice versa. The evaluation of the effects of this trade-off on the probability of re-election is specific to each party.

It is intuitive that each party reaches the optimal position when the marginal benefit (on the probability of re-election) of more financial resources equals the negative marginal benefit of the votes lost. It can be shown (Brock & Magee, 1978) that the equilibrium solution of the model (a game-theoretic Cournot-Nash equilibrium) endogenously determines the tariff level, the amount and distribution of the financial resources employed in the financing of the parties, and the distribution of votes between the two parties. In this context, tariffs can be seen as a "price" that clears the political market for protection (Frey, 1984).

12.6 Administered and Contingent Protection, and Fair Trade

In Sect. 12.5 we have seen how domestic industries can seek protection by lobbying politicians. Another avenue for seeking protection is to petition for import relief through 'administered protection' (AP) procedures. These procedures are rules-oriented and are codified in both national legislation and international agreements. Their characteristic is that they are based on objective criteria rather than on political discretion, and offer protection when an alleged injury occurs. Antidumping, countervailing duty, and safeguard actions are the foremost examples of such procedures.

However, AP is a somewhat broader concept than protection granted to offset the domestic injury deriving from allegedly 'unfair' foreign trade practices (antidumping, countervailing duty: see below) or from an occasional import surge (safeguard actions: see below). It also includes all kinds of protection deriving from domestic regulations whose primary aim is *not* that of (directly or indirectly) influencing international trade as such: for example, regulations aimed at environmental protection. A country worried that a domestic industry is generating an excessive amount of pollution might subsidize imports of the commodity produced by that industry, so as to reduce its domestic production and hence pollution. Or export restrictions might be imposed to curb exports of a commodity (say, timber) so as to prevent excessive exploitation of natural resources (deforestation).

Box 12.1 Free Trade or Fair Trade?

The pursuit of *free trade* involves activities as the harmonization of trading rules and the reduction of barriers to trade. In its simplest sense the issue of free trade should be conducted on a *level-playing field*: more free trade would result from the application of the same policies, rules, mechanisms and institutions to each participant in the trade regime, regardless of origin and capacity. This last point brings us to the conceptual notion of *fair trade*. The term *fair trade* is used to indicate a position that calls for protectionist measures by developed countries against products that have been produced in developing countries at prices developed countries cannot compete with because of their different economic circumstances. As an example we can consider demands by the rich countries for imposing higher environmental and labour standards on the poor countries as preconditions for trade liberalization to prevent social dumping and a so called "race to the bottom" in wages and benefits. Trade sanctions or eco-dumping duties (sometimes referred to as a "social clause") are often imposed in response to violations of labour and environmental standards. Developing countries consider such sanctions as disguised protectionism. Let us shortly analyse the issue of labour standards.

The core labour standards—freedom of association and the right to organize and bargain collectively, freedom from forced labour, the abolition of child labour and freedom from discrimination—are recognized as fundamental rights to which all workers are entitled regardless of the level of development of the country or sector they work. Such list of labour standards is the OECD set of the core standards which corresponds with the International Labour Organization's (ILO) core standards.

The literature on international labour standards can be broadly divided in two categories. The first focuses on the evaluation of the appropriateness of linking labour standards with trade. For surveys see Brown, Deardoff and Stern (1996) and Stern (2003). The second includes recent writings by development economists such as Basu (2001), that link the issue of international labour standards to broader perspectives on development.

The central question is whether implementation and enforcement of global labour standards should be explicitly linked to trade agreement.

The reason why the issue of trade and labour standards is so much debated in trade negotiations is that labour interests in high-standards countries argue that low labour standards are an unfair source of comparative advantage and that increasing imports from low-standards countries will have an adverse impact on wages and working conditions: low wages and labour standards in developing countries threaten the living standards of workers in developed countries. For low-standards countries there is the fear that the imposition of high labour standards upon then is just a form of protectionism and is equally unfair as regards their competitiveness.

The broader concept of administered protection brings us to the problem of 'fair' trade and harmonization. Trade between countries with different environmental and labour standards, as well as with different competition rules, raises a number of new issues. Demands for harmonization to reduce the diversities of domestic policies and institutions, so as to foster free (or at least "fair") trade are now at the centre of the new debate on protectionism versus free trade.

This problem involves not only economic, but also legal aspects, both of which are fully addressed in the two-volume set of essays edited by Bhagwati and Hudec (1996, see also Krugman, 1997). While referring the reader to these works, we shall however briefly show how these two aspects (the economic and the legal ones) are intimately intertwined even in the subset consisting of *contingent protection.*

Under the broader concept of administered protection, in fact, the subset consisting of protection against unfair foreign practices or to offset an import surge is called contingent protection, to which we shall limit our analysis.

Before going on, however, it is interesting to point out that lobbying and contingent protection can be viewed as alternative means for seeking protection in the presence of an alleged injury, so that the interested industry can choose between them through an optimization process that maximizes the expected net benefit (Moore & Suranovic, 1992). In addition, also antidumping law (see below) can be considered as a strategic business tool (alternative to lobbying), since it can be used by domestic firms as an *offensive* tool, even though the law is meant to be a defensive tool (Hartquist, 1987).

12.6.1 Dumping and Antidumping

Dumping is an international *price discrimination* which takes place when a producer sells a commodity abroad at a price lower than that charged in the domestic market (for a case in which dumping occurs in the intermediate goods market see Bernhofen, 1995). The export price considered is f.o.b. (free on board), and so transport cost and insurance are excluded. Also excluded are export duties (if any) and the (possible) markup of the foreign wholesale importer. Dumping is not necessarily a synonym of a bargain-sale below cost, as is often thought, for, on the contrary, it may be a way of maximizing profits. In general three types of dumping can be distinguished: *sporadic*, *predatory*, and *persistent*.

Sporadic dumping, as the name suggests, occasionally occurs when a producer, who happens to have unsold stocks (e.g., because of bad production planning or unforeseen changes in demand) and wants to get rid of them without spoiling the domestic market, sells them abroad at reduced prices. This is the type nearest to the concept of a sale below cost.

Predatory dumping takes place when a producer undersells competitors on international markets in an effort to eliminate them. Of course this producer also suffers losses but can subsequently (in case of success) raise the price to the monopoly level, once competitors leave the market. This type of dumping is, therefore, only temporary.

Persistent dumping is that started off by a producer who enjoys a certain amount of monopolistic power and exploits the possibility of price discrimination between domestic and foreign markets in order to maximize profits. This case can therefore be analysed by using the theory of the discriminating monopolist. It should be recalled that this theory is based on the assumption that the markets are completely

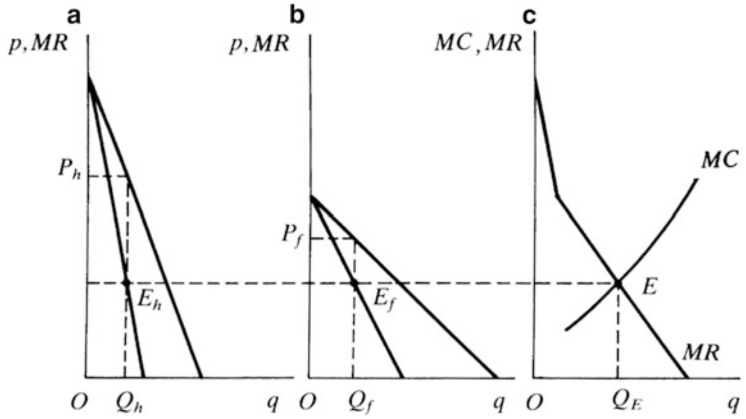


Fig. 12.6 Persistent dumping

separated, so that it is not possible for consumers to purchase the commodity in the market where the price is lower (and even less to carry out arbitrage operations, that is, to buy the commodity in the low-price market and resell it in the high-price one). This separation usually occurs in international trade: transport costs, customs duties and other barriers, imperfect (and costly) information, administrative regulations etc., effectively separate the domestic and the foreign market.

Now, the theory of the discriminating monopolist tells us that, to maximize profits, the monopolist must equalize the marginal revenues in the various markets with one another and with the marginal cost of output as a whole (for simplicity, we assume that the monopolist produces the commodity at one plant, situated at home). If, in fact, the marginal revenue in market i were greater than that in market j, the monopolist could—with the same output and so with the same total cost—increase total revenue (and so profits) by selling one unit less in market j (total revenue decreases by MR_j) and one unit more in market i (total revenue increases by $MR_i > MR_j$). The process would continue up to the point where $MR_i = MR_j$. Once the marginal revenues are equalized (this equalization gives the maximum total revenue corresponding to any given output), profits will be maximized by equating the (common level of the) marginal revenue to marginal cost.

This procedure is represented in Fig. 12.6a–c, where we have drawn, from left to right, the demand and *MR* curves in the home market, the demand and *MR* curves in the foreign market and the overall *MR* and the *MC* curves.

The *MR* curve in panel (c) is obtained by horizontal summation of the *MR* curves in (a) and (b): in such a way, for any given output, for example OQ_E, the common level of the *MR* in the two markets is immediately found ($Q_E E$), which is then carried into panels (b) and (a) ($Q_E E = Q_f E_f = Q_h E_h$), showing the optimal allocation between the two markets, OQ_f and OQ_h (note that $OQ_f + OQ_h = OQ_E$ by construction). The intersection between *MC* and *MR* in panel (c) determines the equilibrium point E; from panels (b) and (a) one finds the price to be charged in the foreign (OP_f) and domestic (OP_h) market respectively, and the corresponding quantities sold.

We see from the figure that $OP_f < OP_h$, but this is *not* due to a sale below cost: on the contrary, it is the condition required by profit maximization (this explains why persistent dumping is also called *equilibrium dumping*). The fact that it is profitable to sell on the foreign market at a lower price than on the home market depends on the fact that *the elasticity of demand is higher on the foreign market*, so that the monopolist's optimum markup—which equals the reciprocal of the elasticity of demand—is smaller in the foreign than in the domestic market. And since the markup is applied to marginal cost, which is one and the same, it follows that the price charged to foreign buyers is higher than that charged domestically.

Whilst sporadic and predatory dumping are undoubtedly harmful to the foreign importing country, it might seem that persistent dumping is beneficial, as the consumers of the importing country will pay a *systematically* lower price for the commodity. But this opinion ignores the loss of the foreign producers of the commodity (or of close substitutes), who will ask for antidumping protection. This (subject to a legal procedure) is granted through an *antidumping tariff*, namely a duty on imports equal to the *dumping margin*.

The dumping margin may be calculated as the difference between OP_f and OP_h (so as to equalize the price to that in the domestic market of the exporting country); alternatively it may be calculated as the difference between OP_f and the so called "*fair value*" of the commodity, which is usually taken to be average cost of production by the exporting firm.

Subject to country-specific institutional differences, the process leading to antidumping action may be broadly described as follows:

(a) A domestic firm (or group of firms representing an industry) files a petition against a foreign firm or industry. This petition is filed with the domestic institution legally entitled to examine it. In the United States, this petition has to be filed with both the International Trade Commission and the Department of Commerce; in the European Union (where trade policy vis-à-vis the rest of the world is centralized) with the European Commission. This action is costly, for it entails data collection costs and legal expenses. Let us call C_0 this initial (sunk) cost.

(b) Within a time t_0 the institution issues a preliminary determination, which may be interlocutory or negative. In the latter case, the procedure ends, in the former case it continues with the next stage.

(c) On the basis of the preliminary findings of the institution, the domestic industry may decide to withdraw the petition or to pursue it. In the latter case further ongoing legal expenses are incurred, say C_1, and the institution continues its investigation, issuing the final decision within a time t_1.

(d) The decision may be positive or negative. In the former case, an antidumping duty is levied (in the United States, the basis is usually the fair value, see above).

Box 12.2 Antidumping Measures in the World

Antidumping, contrary to other neoprotectionist measures, has a long history. The first antidumping laws in the United States go back to 1916 and 1921, while the first multilateral regulation is contained in article VI of GATT 1947. However, up to the 1980s AD actions concentrated within a restricted number of countries: US, European Community, Australia, Canada. It was only in the late 1980s that AD actions began to come from developing countries, mainly Argentina, Brazil, India, and Mexico. In the period 1995–2008 more than 60 % of the world AD actions (3,305 initiated and 2,106 enforced) concerned developing countries. About two thirds of these are directed at other developing countries.

China is the main target of AD actions: in the period 1995–2008 there were 446 petitions against China (about 15 % of the total), of which 441 gave rise to an antidumping measure The next target is the Republic of Korea with 247 petitions (of which 147 accepted) and the United States with 183 petitions (of which 112 accepted).

All the data are drawn from the WTO web site, http://wto.org

Let us now examine the domestic welfare effects of a successful antidumping petition. The traditional view is that the antidumping duty, as any duty, increases producers' surplus at the expense of consumers' surplus. This view has however been challenged on the basis of possible collusive behaviour of the domestic and foreign industry. Prusa (1992) started from the observation that in the United States each of the three possible outcomes of antidumping cases initiated in the period 1980–1985 (petition accepted, rejected, withdrawn) accounted for approximately a third of the total. Now, since most of the costs of a petition are sunk (C_0 is much greater than C_1), one would expect few cases to be withdrawn. However, as Prusa (1992, p. 2) remarks, frequently a petition is withdrawn only after the domestic industry has achieved some type of out-of-court settlement with its foreign rival. The settlement may involve either a price undertaking (i.e., a voluntary price increase by the foreign firm) or a quantity restriction. Furthermore, settlements can be made with or without government approval. Subsequent periods have confirmed this view: in the United States in the period 1980–2005, 1,132 AD procedures were initiated, and about 20 % of these ended with the withdrawal of the petition (Morkre and Kelly, 1994; WTO, 1995 and following years).

Similar results hold for the European Union, where in the same period of the 631 petitions more than 35 % have been withdrawn (Bown, 2010).

Hence most if not all of the withdrawals are really out-of-court settlements. This is interpreted by Prusa in terms of a game-theoretic bargaining model which gives rise to a unique Nash solution. The result is that within this bargaining process the domestic and foreign firms cooperate on pricing decisions so as to achieve a collusive level of profits.

Thus antidumping cases may actually be used as a stratagem that paves the way for collusion among (domestic and foreign) oligopolistic firms. In these cases, as Prusa observes, the welfare conclusion is exactly the opposite of conventional wisdom: the imposition of an antidumping duty, instead of decreasing consumers' surplus, might actually increase it, because the alternative is not free trade, but a collusive oligopolistic situation.

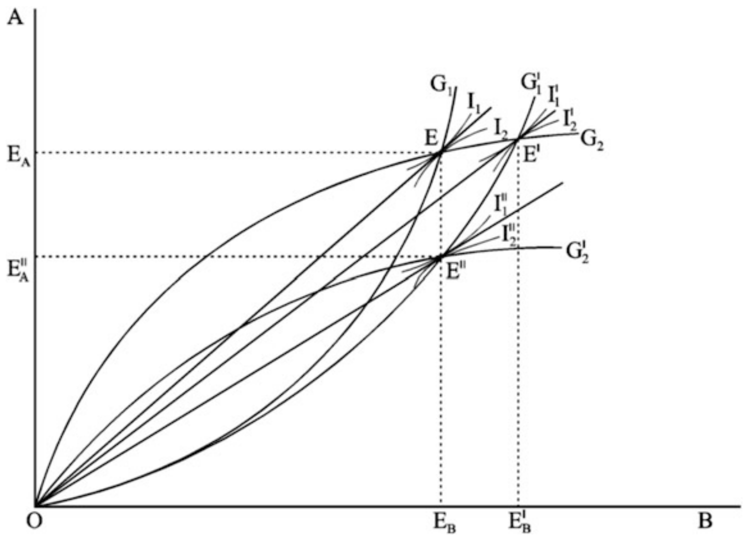


Fig. 12.7 Export subsidies and countervailing duties

For further analysis of the effects of antidumping duties see Anderson et al. (1995).

12.6.2 Countervailing Duty

A countervailing duty (CVD) is a duty levied in retaliation to an export or production subsidy by a foreign country. It is interesting to observe that export subsidies constitute a sort of official dumping, since they are paid out by the government to domestic producers-exporters, enabling them to sell abroad at a lower price than at home. This explains why export subsidies are prohibited, except when they are rebates of indirect taxes (see above, point (i) in Sect. 10.6.4).

Since an export subsidy increases consumers' welfare in the importing country, why should there be a retaliation? The answer is the same as in the case of dumping: the producers of the importing country are harmed, hence they will ask for protection by filing a petition (the procedure is similar to that described above in the case of an antidumping petition).

Subsidies have been treated above (Sect. 12.4) in a partial equilibrium context, but in order better to examine the effects of a CVD levied against a foreign subsidy we need a general equilibrium setting. Let us start from the free-trade situation depicted in Fig. 12.7, and assume that country 1 introduces an export subsidy.

The export subsidy causes country 1's offer curve to shift downwards and to the right, from OG_1 to $OG_1^{\prime}$. To show this, take for example point E on country 1's offer curve. At the price ratio given by the slope of OE, consumers and producers are willing to exchange OE_B of exportables for OE_A of importables. However, due to the presence of an export subsidy (measured in terms of commodity A taken as *numéraire*)[1] equal to EE'', the foreigners actually give $OE''_A = E''E_B$ of importables. Hence we have a domestic price ratio

$$p_d = \text{slope of } OE = EE_B/OE_B$$

and an international price ratio (terms of trade)

$$p = \text{slope of } OE'' = E''E_B/OE_B.$$

The subsidy rate (s) is then measured by

$$s = EE''/EE_B.$$

Another way of showing this is to note that $(1-s)p_d = p$, hence $s = (p_d - p)/p_d = EE''/EE_B$.

Let us now consider what happens to international equilibrium. Given country 2's offer curve OG_2, the equilibrium point shifts from E to E', and country 2's imports increase from OE_B to OE'_B.

Note that the terms of trade move in favour of country 2, whose welfare improves while that of country 1 decreases: this can be seen by checking that the social indifference curve I'_2 is better than I_2, while I'_1 is worse than I_1 (these curves are drawn according to the same technique described in Fig. 11.1).

Suppose now that country 2 retaliates by a CVD, whose amount is calculated so as to bring imports back to the pre-subsidy situation, namely to OE_B. Country 2's offer curve shifts downwards (see Sect. 10.5.2) to position OG'_2, and the new equilibrium point is E''. The terms of trade further move in favour of country 2, whose welfare again increases at the expense of country 1's welfare (I''_2 is better than I'_2, and I''_1 is worse than I'_1).

Note that such a CVD also restores country 2's domestic price ratio to its initial (free trade) value. In fact, at the beginning, the domestic price ratio coincided with the terms of trade (slope of OE). In the final situation, the terms of trade are given by the slope of ray OE'', but country 2's domestic price ratio is given by the slope of ray OE, since EE'' is the amount of the tariff (see Sect. 10.5.2).

Hence both the quantity of imports and the domestic price ratio are back to the initial situation: the impact of the export subsidy has been completely offset by the CVD. One might then ask "why can't countervailing duties deter export subsidization?". This is the title of a paper by Qiu (1995), who shows that there are three factors that explain the coexistence of export subsidies and CVDs. He

[1]This neglects the problem of how the government raises the funds required to pay the subsidy. For an in-depth treatment of this problem see Meade (1952, chap. VI).

works in the context of a duopolistic model, but his considerations can be applied to the standard competitive model as well.

The first reason is a delay in retaliation. A petition against an alleged foreign export subsidy requires time to be examined by the domestic institution, and hence, even assuming a 100 % probability of success, during this time the export subsidy exerts all its effects. International agreements, in fact, allow retaliation but not vengeance, which means that no CVD can be levied if the foreign country withdraws the export subsidy at the end of the procedure.

The second reason is the upper limit to a CVD. According to international agreements, in fact, the CVD rate cannot exceed the subsidy rate. Now, we see from Fig. 12.7 that *the fully offsetting CVD rate is greater than the export subsidy rate*. In fact, the CVD rate is measured by the proportional downward shift of the OG_2 curve, for example by $EE^{''}/E^{''}E_B$ (see Sect. 10.5.2), which is clearly greater than $EE^{''}/EE_B$. The application of a CVD with the same rate as the subsidy would entail a smaller downward shift of the OG_2 curve, that would bring this curve to an intermediate position (not shown in the diagram) between OG_2 and $OG_2^{'}$. It is then easy to check that the CVD would not completely offset the subsidy, since the final quantity of country 2's imports would be somewhere between E_B and $E_B^{'}$.

The third reason is the phenomenon of out-of-court settlements that give rise to VERs. This is the same phenomenon already examined above in the case of AD petitions. The data are also similar: between 1980 and 2005, about 30 % of CVD petitions were withdrawn in the US, most of them resulting in VERs. Of the remaining ones, 64 % have been rejected and only 36 % accepted.

Box 12.3 THE US-EU Dispute on Steel

Among the 12 active WTO disputes between the European Union, as a complaining party, and the United States—mostly associated to misuse of trade defence instruments, four of them relate to the steel sector. Under case number WT/DS248, in particular, the EU is complaining on the US definitive safeguard measures on imports of certain steel products adopted on 5 March 2002, with the belief that such measures are in breach of both the US obligations under the provisions of GATT 1994 and of the Agreement on Safeguards (SA).

Following the recommendations of the International Trade Commission (ITC), which, on 22 June 2001, initiated a safeguard investigation on imports of four broad groups of steel products, the US President announced, on 5 March 2002, definitive safeguard measures in the form of an increase in duties ranging from 8 to 30 % on imports of certain steel products, effective as of 20 March 2002.

Although three rounds of consultations took place over March–April 2002, the last jointly with Korea, Japan, China, Switzerland and Norway, they did not succeed in solving the dispute, and a panel was established, under request by the EC, at the special meeting of the Dispute Settlement Body (DSB) of 3 June 2002. More precisely, a single Panel was established against the US steel safeguards under Article 9.1 of the Dispute Settlement Understanding (DSU), following requests presented by Japan, Korea, China, Switzerland, Norway, New Zealand and Brazil.

The claims put forward relate to violations of both the Article XIX of the GATT agreement on "unforeseen developments" and a number of SA provisions, including, among others, the lack of increased imports, the incorrect definition of the domestic industries that produce like products, the lack of serious injury or threat thereof serious injury and the absence of causal link between imports and serious injury.

The Panel report, which was circulated to all WTO Members on 11 July 2003, found that all safeguard measures lacked a legal basis. However, on 11 August 2003 the United States decided to appeal the panel report. The Appellate Body rejected the appeal on 10 November 2003, and authorized an appropriate relaliation by the EU against the United States in case the United States maintained the tariffs on steel imports. On 4 December 2003 the United States withdrew these tariffs.

12.6.3 Safeguard Actions

International agreements also allow a country to protect domestic producers against fair imports (that is, imports that are not dumped or subsidized by the foreign country) under certain circumstances. The characteristic of this form of administered protection (called a *safeguard* action) is that it must be temporary and nondiscriminatory. For example, a country experiencing a sudden surge of imports that threatens severe injury to domestic producers, may impose a temporary nondiscriminatory tariff.

Although SA, AD, and CVD are called *emergency measures* by WTO, there are three important differences among them.

The first is that, while for the application of AD and CVD measures it is necessary that unfair competition has taken place, for the application of SA it is enough that the industry of the country has been damaged by the increase in imports, in the absence of export subsidies or dumping by the exporting country.

The second concerns the application. AD and CVD measures are only applied against the country guilty of unfair competition. On the contrary, SA measures are applied against all countries (with some exceptions concerning developing countries) whose exports have harmed the country which presents the petition.

The third concerns the compensation. AD and CVD measures are a penalty imposed on countries guilty of an unlawful behaviour, hence obviously they do not imply compensation to these countries. On the contrary, a country that obtains an SA is required to compensate the exporting countries for possible losses that they undergo because of the SA measure.

Safeguard actions are a small minority with respect to AD and CVD actions: for example, at the world level, in the period 1995–2008, 3,305 AD petitions were undertaken (of which 2,106 were accepted), while there were only 209 requests for CVD (of which 121 were enforced) and 164 requests for SA (of which 83 were applied).

References

Anderson, S. P., Schmitt, N., & Thisse, J.-F. (1995). Who benefits from antidumping legislation?

Baldwin, Robert E. (1982). (1988). The political economy of protectionism.

Basu, K. (2001). On the goals of development.

Bernhofen, D. M. (1995). Price dumping in intermediate good markets.

Bhagwati, J. N. (1987). VER, quid pro quo DFI and VIEs: Political-economy-theoretic analyses.

Bhagwati, J. N., & Hudec, R. E. (Eds.). (1996). *Fair trade and harmonization: Prerequisites for free trade?*

Bown, C. P. (2010). Temporary trade barriers database (TTBD).

Brock, W. A., & Magee, S. P. (1978). The economics of special interest politics: The case of the tariff.

Brock, W. A., Magee, S. P., & Young, L. (1989). *Black hole tariffs and endogenous policy theory: Political economy in general equilibrium*.

Brown, D.K., Deardoff, A.V., & Stern, R.M. (1996). International labor standards and trade: A theoretical analysis.

Dinopoulos, E., & Kreinin, M. (1989). Import quotas and VERs: A comparative analysis of a three-country framework.

Findlay, R., & Wellisz, S. (1983). Some aspects of the political economy of trade restrictions.

Frey, B. S. (1984) *International political economics*.

Greaney, T. M. (1996). Import now! An analysis of market-share voluntary import expansions (VIEs).

Greenaway, D. (1983). *International trade policy: From tariffs to the new protectionism*.

Hartquist, D. A. (1987). Trade wars – arming for battle.

Herberg, H. (1990). Welfare effects of non-tariff barriers: A general equilibrium analysis.

Hillman, A. L., & Ursprung, H.-W. (1988). Domestic politics, foreign interests, and international trade policy.

Jones, K. (1984). The political economy of voluntary export restraint agreements.

Jones, R. W., & Krueger A. O. (Eds.). (1990). *The political economy of international trade*.

Krugman, P. R. (1997). What should trade negotiators negotiate about?

Laird, S., & Yeats, A. (1990). Trends in non-tariff barriers of developed countries.

Lopez, R. A., & Pagoulatos, E. (1994). Rent seeking and the welfare costs of trade barriers.

Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: Theory and evidence* (chaps. 19–20).

Mayer, W. (1993). Lobbying for tariff policies.

Meade, J. E. (1952). *A geometry of international trade.*

Messerlin, P. A. (1981). The political economy of protectionism: The bureaucratic case.

Moore, M. O., & Suranovic, S. M. (1992). Lobbying vs. administered protection: Endogenous industry choice and national welfare.

Morkre, M. E., & Kelly, K. H. (1994). Effects of unfair imports on domestic industries: US antidumping and countervailing duties cases, 1980 to 1988.

Murray, T., Schmidt, W., & Walter, I. (1978). Alternative forms of protection against market disruption.

Olson, M. (1964). *The logic of collective action.*

Pomfret, R. (1989). The economics of voluntary export restraint agreements.

Prusa, T. J. (1992). Why are so many antidumping petitions withdrawn?

Qiu, L. D. (1995). Why can't countervailing duties deter export subsidization?

Schucknecht, L. (1992). *Trade protection in the European community.*

Stern, R.M. (2003). Labor standards and trade agreements.

Takacs, W. E. (1978). The nonequivalence of tariffs, import quotas, and voluntary export restraints.

Vousden, N. (1990). *The economics of trade protection.*

WTO. (1995). Trading into the future: WTO-the World trade organization.

Part V
Growth and Trade

Chapter 13
International Trade and Growth: Comparative Statics

13.1 Introduction

At the cost of some simplification, the causes of growth are traditionally classified in two categories: increase in factor endowments and technical progress. Many believe that such a classification is artificial, for in the real world these two causes are not separable: for example, the increase in the stock of capital often consists in the purchase of new machines rather than of machines identical to those already owned. Since the new machines contain the latest technological improvements ("embodied" technical progress), it becomes impossible to distinguish the increase in the capital stock from technical progress.

It should also be noted that in traditional theory technical progress is *exogenous*, in the sense that technological improvements fall on the economy like manna from heaven. Technical progress, however, usually derives from activities directed at procuring it (for example R&D: Research and Development), hence it is normally *endogenous*.

We refer the reader to the textbooks on economic growth for a detailed examination of these problems and we maintain the traditional distinction for simplicity's sake, also assuming that technical progress is exogenous and of the "disembodied" type. More sophisticated forms of technological change will be examined in Chap. 15.

The theoretical analysis of the relations between growth and international trade was initially directed to the examination of the effects of the various forms of growth on international trade, in particular on the volume and pattern of trade, on the terms of trade, and on welfare. In this analysis—which is essentially of a comparative-static nature and usually adopts the assumptions of first-degree homogeneous production functions and of no factor intensity reversal—growth and its causes are considered as given and their impact on international trade is explored.

This is an inherently incomplete or partial analysis, as it examines solely one aspect of the problem: the increase in the stock of capital, for example, is not a windfall but depends on investment; besides, international trade can influence

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_13,

growth. Therefore, in a more general setting, one must consider the interrelationship between trade and growth, as these influence each other. The analysis of these problems requires the use of dynamic models, which will be briefly examined in the next chapter. Besides, as stated above, the sources of growth, and in particular technological progress, cannot be taken as exogenous. The relations between the theory of endogenous growth and international trade will be examined in Chap. 15.

It is as well to inform the reader that we shall not deal with the relations between international trade and economic development (as distinct from growth), that is, with the specific problems arising when one considers the role of international trade in the development of less-developed countries. This is a topic of great importance but cannot be adequately dealt with here, as it pertains more to (and in any case requires the knowledge of) development economics. Part of the material examined in this and in other chapters (as, for example, the infant industry argument illustrated in Sect. 11.2, and some of the "new" theories of international trade examined in Part III) could of course be relevant to issues of economic development, as shown, for example, by Findlay (1984), but for the reasons stated we do not examine these aspects.

13.2 The Effects of Growth on the Volume of Trade

The study of these effects requires a preliminary examination of the consumption and production effects of growth, which we shall perform following the classification of Johnson (1955, 1959); see also Chacholiades (1978), Corden (1971b), Ghosh (1984), Takayama (1972), Woodland (1982).

13.2.1 Consumption Effects

As regards the *consumption effects*, the question of interest is whether growth, *at unchanged relative price* of the commodities (terms of trade), will increase the demand for the importable[1] more than proportionally to, in the same proportion as, or less than proportionally to, the increase in national income (measured in terms of either commodity, i.e. in real terms), that is, whether growth will make the country relatively less self-sufficient (more dependent on trade), neither more nor less dependent on trade, or relatively more self-sufficient. It is in fact clear that, if the demand for the importable increases more than proportionally to the increase

[1]The importable is, as usual, the commodity for which there exists a domestic excess demand in the relevant price range; as we assume incomplete specialization, this commodity is also produced domestically.

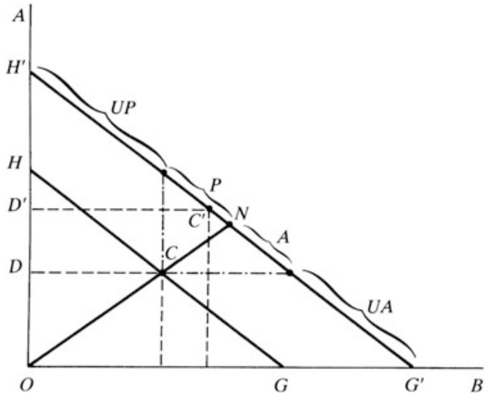


Fig. 13.1 Consumption effects of growth

in national income, this will tend— *ceteris paribus*, i.e. ignoring for the moment the production effects[2]—to increase international trade (in the sense of the share of imports in national income[3]) and vice versa in the opposite case. In the former case growth is defined as *pro-trade biased*, whilst it is called *anti-trade biased* in the opposite case and *neutral* in the case of equiproportional increases.

There are, in addition, the case of *ultra-pro-trade-biased* growth, in which the demand for the importable increases not only more than proportionally to the increase in national income, but also by a greater absolute amount, and the case of *ultra-anti-trade-biased* growth, in which the demand for the importable decreases when national income increases (this implies that we are in the presence of an inferior good).

The five cases listed above can be conveniently illustrated in terms of Fig. 13.1. The budget line or equal national income line (which we call *isoincome*) *HG* represents the value of the initial national income, and *OH* is the initial real national income in terms of commodity *A* (for isoincome lines see Sect. 13.2.3). The initial consumption point is C, so that the consumption of the importable is *OD*. Growth shifts the isoincome line (at constant prices) in parallel upwards to $H'G'$, and the proportional increase in income is $H'H/OH$. If, for example, the new consumption point is C', the proportional increase in the demand for the importable is $D'D/OD$. As $D'D/OD > H'H/OH$, growth is pro-trade biased (P). It is possible to check graphically that, if the new consumption point falls:

[2]It should be pointed out that, in the case of *complete specialization,* the imported commodity is not produced at home, so that there are no production effects and the effects of growth coincide with the consumption effects.

[3]Since in the context of the pure theory of international trade the value of imports equals the value of exports, it makes no difference whether this share is measured in terms of imports or exports. Given the relative price, once the effects on the excess demand for *A* (demand for imports) have been determined, Walras' law (see Sect. 3.3) immediately allows us to determine the effects on the excess supply of *B* (supply of exports).

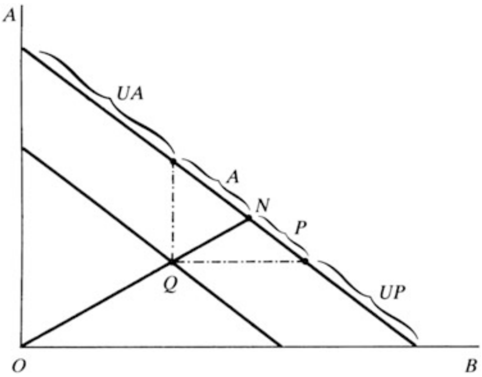


Fig. 13.2 Production effects of growth

1. In segment *UP* (ultra-pro-trade), the absolute increase in the demand for the importable is greater than the absolute increase in national income (this, of course, implies that the other good is inferior);
2. In segment *P* (pro-trade), the proportional increase in the demand for the importable is greater than the proportional increase in national income;
3. At point *N* (neutral), the proportional increases in the two variables are equal;
4. In segment *A* (anti-trade), the proportional increase in the demand for the importable is smaller than the proportional increase in national income;
5. In segment *UA* (ultra-anti-trade), the demand for the importable decreases in absolute terms (this, of course, implies that it is an inferior good).

13.2.2 Production Effects

If we now consider the *production effects*, we can give an analogous classification on the basis of the relations between the variation in the domestic output of the importable and the increase in national income, at unchanged relative price of the commodities. In this context, the degree of self-sufficiency is positively related to the increase in the domestic output of the importable. If, for example, this output increases more than proportionally to the increases in national income, the country will—*ceteris paribus*—become more self-sufficient (less dependent on trade): growth is anti-trade biased as regards its production effects. In Fig. 13.2, point Q represents the initial output bundle; with a reasoning similar to that used in relation to Fig. 13.1 we can see that if the point representing the new output bundle falls in segment *UA*, the absolute increase in the domestic output of the importable is greater than the absolute increase in national income, so that growth is ultra-anti-trade-biased, and so on for segments A, P, *UP* and for point N.

13.2.3 A Reformulation in Terms of Elasticities: The Total Effect

What has been stated in terms of relations between proportional increases can be reformulated in terms of domestic demand and supply income-elasticities. The income elasticity of the domestic demand for the importable is defined as the ratio between the proportional increase in this demand and the proportional increase in national income, that is

$$\eta_{dY} = \frac{\Delta A^D/A^D}{\Delta Y/Y}, \tag{13.1}$$

that is, in terms of Fig. 13.1

$$\eta_{dY} = \frac{D'D/OD}{H'H/OH} = \frac{D'D/OD}{G'G/OG}.$$

It can be easily checked that this elasticity can also be written as the ratio between the marginal propensity μ_{dY} and the average propensity α_{dY} to consume commodity A; in fact

$$\eta_{dY} = \frac{\Delta A^D/\Delta Y}{A^D/Y} = \frac{\mu_{dY}}{\alpha_{dY}}. \tag{13.2}$$

Given these definitions, the consumption effects of growth will be pro-trade-biased , neutral, anti-trade-biased according to whether $\eta_{dY} \gtreqless 1$; the ultra-pro-trade and ultra-anti-trade cases occur when $\mu_{dY} > 1$ and $\mu_{dY} < 0$ respectively.

As regards the production effects, we can define an elasticity of domestic supply (production) of the importable as

$$\eta_{sY} = \frac{\Delta A^S/A^S}{\Delta Y/Y} = \frac{\Delta A^S/\Delta Y}{A^S/Y} = \frac{\mu_{sY}}{\alpha_{sY}}, \tag{13.3}$$

where μ_{sY} and α_{sY} are the marginal and average propensity to produce commodity A. The production effects of growth will be pro-trade-biased, neutral, anti-trade-biased according to whether $\eta_{sY} \lesseqgtr 1$; the ultra-pro-trade and ultra-anti-trade cases occur when $\mu_{sY} < 0$ and $\mu_{sY} > 1$ respectively.

The effects of growth on the demand for imports (this demand must not be confused with the demand for importables: the two coincide only in the case of complete specialization, whilst in the normal case of incomplete specialization the demand for imports equals the demand for the importable less the domestic production of this commodity) depend on the combination of the consumption and production effects. The result will be pro-trade-biased, neutral, or anti-trade-biased according to whether the demand for imports increases more than proportionally

Table 13.1 Classification of the effects of growth on trade by combining the consumption and production effects

	Consumption effect				
Production effect	Neutral	Pro-trade	Ultra-pro-trade	Anti-trade	Ultra-anti-trade
Neutral	N	P	P or UP	A or UA	UA
Pro-trade	P	P	P or UP	Not UP	UA
Ultra-pro-trade	P or UP	P or UP	UP	Not UA	All types possible
Anti-trade	A or UA	Not UP	Not UA	A or UA	UA
Ultra-anti-trade	UA	UA	All types possible	UA	UA

to, in the same proportion as, or less than proportionally to the increase in national income; it will be ultra-pro-trade-biased or ultra-anti-trade-biased when the increase in the demand for imports is greater than the absolute increase in income, or when this demand decreases as income increases.

The result can easily be determined when the consumption and production effects have the same bias. If, for example, they are both pro-trade-biased, the demand for imports will certainly increase: in fact, this means that, for the same (proportional) increase in income, the demand for the importable increases more than proportionally to the increase in its domestic production, so that the demand for imports must increase to make up the difference. Besides, this increase is proportionally greater than the increase in income. In fact, if we denote by g_d, g_s, g_m, g_Y the (proportional) growth rates of the demand for the importable, the domestic production of this, the demand for imports, and national income, respectively, then, in general (see Sect. 27.1),

$$g_m = g_Y + \frac{A^D}{A^D - A^S}\left(g_d - g_Y\right) - \frac{A^S}{A^D - A^S}\left(g_s - g_Y\right), \tag{13.4}$$

so that, in our case, $g_m > g_Y$ as $g_d > g_Y$ and $g_s < g_Y$. Unfortunately the results are less obvious when the consumption and production effects have an opposite bias. The results of all possible combinations are given in Table 13.1: for example, the result of a growth which has a pro-trade-biased production effect and a neutral consumption effect can be read off the intersection of the row labelled pro-trade and the column labelled neutral.

Most results are intuitively clear and are those occurring when both consumption and production effects have the same kind of bias or when one of the two effects has a certain bias whilst the other is neutral; these results can easily be checked by means of (13.4). It is similarly intuitive that the table is symmetric with respect to the diagonal: for example, the result of a pro-trade-biased production effect combined with an ultra-anti-trade-biased consumption effect is qualitatively the same as the result of a pro-trade-biased consumption effect combined with an ultra-anti-trade-biased production effect.

Less intuitive is the fact that whilst an ultra-anti-trade-biased effect prevails on a pro-trade-biased effect (the result is an any case *UA*: see Table 13.1), on the contrary an ultra-pro-trade-biased effect does not prevail on an anti-trade-biased effect (the result is in any case not *UA*, so that a result A is also possible). To understand this asymmetry we must remind that imports are the excess demand for the importable, $A^D - A^S$. Now, in the case of a *UA* consumption effect combined with a P production effect, A^D decreases whilst A^S increases (though less than proportionally to the increases in income), so that $A^D - A^S$ certainly decreases (a *UA* result). Similarly, in the case of a *UA* production effect combined with a P consumption effect, A^S increases by more than the absolute increase in income and A^D also increases, but by less than the absolute increase in income; therefore A^D increases by less than A^S and the demand for imports $A^D - A^S$ decreases (a *UA* result).

On the contrary, in the case of a *UP* consumption effect combined with an A production effect, A^D increases by more than the absolute increase in income and A^S also increases, but by less than the absolute increase in income. It follows that the demand for imports $A^D - A^S$ certainly increases (so that the result cannot be *UA*), but we do not know whether it increases more or less than proportionally to the increase in income, so that the result might be A. Similarly in the case of a *UP* production effect combined with an A consumption effect, A^S decreases whilst A^D increases, but less than proportionally to the increase in income: the demand for imports $A^D - A^S$ certainly increases (and so the result cannot be *UA*), but we do not know whether it increases more or less than proportionally to the increase in income, so that the result might be A.

13.3 Growth and Terms of Trade: Immiserizing Growth

13.3.1 The Large Country and the Terms of Trade

We have so far assumed that the relative price of commodities (terms of trade) is given. This assumption is acceptable in the context of a small country model, where the changes in the country's demand for imports and supply of exports have negligible effects on the world market. But in the opposite case one must investigate the effects of the various types of growth on the terms of trade. For this purpose, it is necessary to determine the shifts of the offer curve of the growing country (country 1, say) due to the various types of growth. In Fig. 13.3 we have the initial offer curve (OG_1) and terms of trade (slope of OR); for the time being, we ignore curve OG_2. The initial equilibrium point is E. Since in all types of growth—except for the ultra-anti-trade-biased one—there is an increase in the demand for imports at unchanged terms of trade (and so in the supply of exports, given Walras' law: see

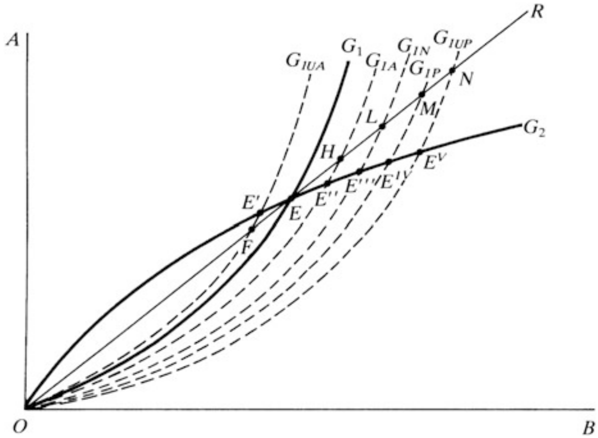


Fig. 13.3 Growth shifts the offer curve and changes the terms of trade

Chap. 3), the points on the terms-of-trade ray *OR* corresponding to the new offer curve will all be to the right of E, except for the *UA* case, in which the point will be to the left of E (lower imports and so lower exports) . The order of the points will be that indicated in Fig. 13.3, since the absolute increase in the demand for imports for a given increase in income is greater as growth is more favourable (or less unfavourable) to trade.

Since the reasoning can be repeated for any given terms of trade, if we imagine rotating ray *OR* we obtain the broken-line offer curves OG_{1UA}, OG_{1A}, and so on. In the case of a small country, *OR* would be given, and we would only have to consider points F, H, L, M, N, which illustrate the effects of growth on the volume of trade described in Sect. 13.2. In the case in which the growing country is not small, the shifts in its offer curve will influence the terms of trade. This can be verified by introducing the (given) offer curve of country 2, OG_2 and finding the intersection between this and country 1's new offer curve, so as to determine the new international equilibrium point. This will be E', or E'', etc., depending, as the case may be, on the type of growth actually occurring.

It can be seen from the figure that the terms of trade become worse and worse in all cases of growth (except in the *UA* case, in which they improve), the more favourable (or less unfavourable) to trade growth is. One only has to draw straight line segments from the origin to the various points E'', E''', etc. and verify that their slope (equal to the terms of trade p_B/p_A) gets smaller and smaller than the slope of *OR* (except for the slope of OE', which is greater).

This result can be explained in the following way: if we exclude *UA* growth, in all other cases country 1's demand for imports increases at the given terms of trade, so that an excess demand for A will arise in the world market (and, given Walras' law, there will be a correlative excess supply of B): this will cause a decrease in p_B and an increase in p_A, thus a decrease in p_B/p_A. As these forces grow more intense the greater the excess demand for A and the excess supply of B in the world market, there will be a greater decrease in p_B/p_A the greater the excess demand and supply become. However, these price changes will put a brake on country 1's demand for A and supply of B, while at the same time stimulating country 2's supply of A and demand for B. This explains why in the new situation of equilibrium country 1's demand for imports of A and supply of exports of B will ultimately increase by less than the initial effect of growth: it is, in fact, sufficient to compare the coordinates of any one of the points $E'', \ldots, E^V$ with those of the corresponding equilibrium points $H, \ldots, N$.

In the case of *UA* growth, on the contrary, the results are quite the opposite: the initial decrease in country 1's demand for imports, etc., gives ultimately rise to an increase in p_B/p_A and so to a boost to that demand, etc.; thus at the new equilibrium point E', there will be an improvement in the terms of trade and a decrease in the volume of trade, which is, however, less intense than the initial decrease (point F).

It should be stressed that the analysis so far is valid in the case of normal offer curves: in the case of anomalous offer curves the results might be different, but we do not wish to burden our treatment with the examination of these, which the reader can in any case easily perform by way of the same technique. Instead it is important to mention the possible negative effects of growth on social welfare: this is the so-called immiserizing growth case.

13.3.2 Immiserizing Growth

Growth is called immiserizing when it reduces the welfare of the growing country. This possibility was first pointed out by Edgeworth (1894, pp. 40–42) and taken up again by Bhagwati (1958, 1973), who gave it its name, and other authors: see Bhagwati et al. (1998, chap. 29); Hatta (1984). On the relations between immiserizing growth and donor-enriching "recipient immiserizing" transfers see Bhagwati, Brecher, and Hatta, 1984.

This phenomenon involves the relations between growth, changes in the terms of trade, and changes in welfare. In general, as we have seen, growth can bring about either an improvement or a deterioration in the terms of trade. The deterioration in the terms of trade can, in turn, improve, leave unchanged, or cause a deterioration in social welfare. It follows that the deterioration in the terms of trade is a necessary, but not a sufficient condition for the decrease in social welfare. Let us now examine the case we are concerned with.

If we assume that the terms of trade deteriorate as a consequence of growth, the possibility of a decrease in social welfare is illustrated in Fig. 13.4, where *TT* is the initial transformation curve. Given the initial terms of trade represented by the slope of *RR*, the country produces at Q and consumes at E by trading $Q_B E_B$ of B (exports) for $Q_A E_A$ of A (imports), thus reaching the social indifference curve I. As a consequence of growth the transformation curve shifts to $T'T'$ and the terms of trade deteriorate (the slope of $R'R'$ is lower in absolute value than the slope

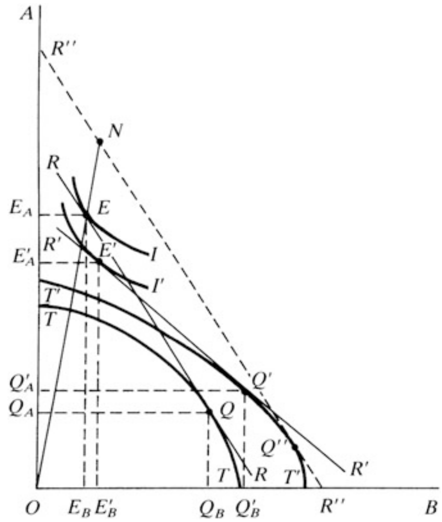


Fig. 13.4 Immiserizing growth

of *RR*). The country produces at Q' and consumes at E' by trading $Q'_B E'_B$ of B for $Q'_A E'_A$ (it is easy to see that $Q'_A E'_A < Q_A E_A$ and $Q'_B E'_B > Q_B E_B$, that is, the country obtains less imports in exchange for more exports). As E' is on the social indifference curve I', lower than I, social welfare has decreased.

It is interesting to ascertain which is the type of growth (according to the classification examined in the previous section) represented in Fig. 13.4. For this purpose we have to determine the consumption point and the production point along the $T'T'$ curve at unchanged terms of trade. By drawing a tangent to the $T'T'$ curve parallel to the *RR* straight line, i.e. $R''R''$, we find the production point Q'', which implies a lower output of A and a higher output of B with respect to Q'. Therefore, growth has ultra-pro-trade-biased production effects. As regards the consumption effect, the point of tangency between $R''R''$ and a social indifference curve may in general occur either to the left or to the right of N or even at N, so that in the absence of further information we cannot classify the consumption effect. However, if we exclude the possibility that A is an inferior good (thus excluding a *UA* consumption effect), then, on the basis of Table 13.1, we can conclude that the type of growth is certainly not *UA*. The same result is obtained if we observe, on the basis of Fig. 13.3, that the condition for the terms of trade to move against the growing country is that growth must not be *UA*.

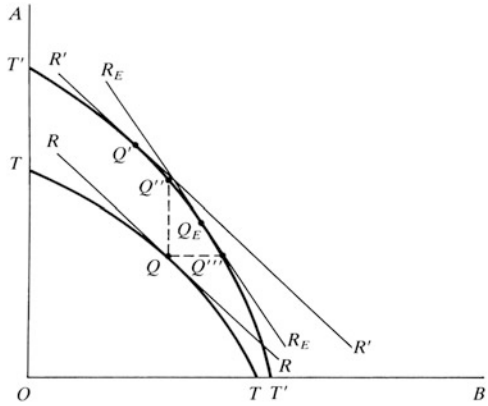


Fig. 13.5 Rybczynski's theorem and relative price of goods

13.4 Increase in Factor Endowments and International Trade: Rybczynski's Theorem

In this section we examine the effects of an increase in factor endowments; the effects of technical progress will be examined in the next section. More sophisticated forms of growth will be analysed in Chap. 15.

The point of departure for examining the effects of an increase in factor endowments is Rybczynski's theorem, (Rybczynski, 1955; see also McDermott, 1985) according to which the increase in the quantity of a factor (given the other) will cause an increase in the output of the commodity which is intensive in that factor and a decrease in the output of the other commodity, at unchanged commodity and factor prices.

The proof of this theorem has been given in Sect. 5.4. We now consider a trading open economy, where we must distinguish the small country case from the case in which the country is sufficiently large for its demands and supplies on the world market to influence the terms of trade.

Let us consider Fig. 13.5 (which reproduces Fig. 5.2) and assume, by way of example, that commodity A is the importable. We further assume, for simplicity, that no commodity is inferior, so that, when income increases, the demand for both A and B increases (each, of course, increases by less than income). In the passage from Q to Q', the output of commodity A has increased more than income whilst the output of B has decreased. It follows that, within the country: (a) the excess demand for A (demand for imports) decreases, as output has increased more than demand; (b) the excess supply of B (supply of exports) decreases, as output has decreased whilst demand has increased. Therefore, on the world market, at the given world relative price, there will be a decrease in both the demand for A and the supply of B.

It is at this point that the distinction between the small and the large country case becomes relevant. In the former case the terms of trade do not change, the country will go on producing at Q' and we shall be in the presence of a case of *UA* growth, as the country's demand for imports (and supply of exports) have decreased.

In the latter case, the excess supply of A on the world market (due to the decrease in the country's demand for imports), and the correlative excess demand for B (due to the decrease in the country's supply of exports) will cause changes in world prices, since the excess supply of A will put a downward pressure on p_A and the excess demand for B an upward pressure on p_B; therefore the terms of trade p_B/p_A increase. This confirms the closed-economy result. Note that, since we have assumed A to be the importable, the terms of trade have improved.

An alternative way to arrive at the same results is to employ the analysis carried out in the previous sections. Since A is, assumedly, the importable, with reference to Fig. 13.2 we find that Q' lies in the *UA* stretch of the isoincome line, so that the increase in the amount of labour has given rise to a growth with *UA* production effects. It is therefore unnecessary to know the consumption effects: in fact, from Table 13.1 we know that a growth with *UA* production effects is globally *UA*, except for the case of *UP* consumption effects, which is, however, ruled out by the assumption of no inferior goods (*UP* consumption effects on A, in fact, imply a decrease in the consumption of B). As regards the change in the terms of trade of a large country, we know from Fig. 13.3 that a *UA* growth causes an increase in the relative price p_B/p_A, that is, an improvement in the terms of trade as A is the importable.

Let us now consider the case in which the importable is commodity B, maintaining the assumption that there are no inferior goods. When the production point shifts from Q to Q', the consequences for the country will be: (a) the excess supply of A (supply of exports) increases, since its output (which increases by more than income) increases by more than demand (which increases by less than income); (b) the excess demand for B (demand for imports) increases, because output decreases whilst demand increases. Therefore—leaving aside the small country case—on the world market at unchanged prices there will be an increase in both the supply of A and the demand for B and so—since the initial situation was of equilibrium—an excess supply of A and an excess demand for B. This will cause a decrease in p_A and an increase in p_B, so that p_B/p_A will increase, confirming the closed-economy results. As B is the importable, the terms of trade have moved against the country.

13.5 Technical Progress and International Trade

13.5.1 Types of Technical Progress

Before coming to grips with the analysis of the effects of technical progress on international trade, it is necessary to introduce the notions of neutrality and bias of technical progress. It should be remembered that we are considering solely

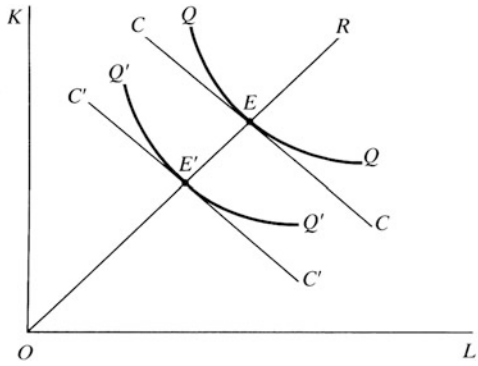


Fig. 13.6 Neutral technical progress

disembodied exogenous technical progress. For a general treatment of technical progress see, for example, Allen (1967, chap. 13) and Burmeister and Dobell (1970, chap. 3).

Among the various possible definitions, we shall adopt the traditional Hicksian definition (Hicks, 1932, 2nd ed.: 1963), according to which technical progress is neutral if, at unchanged capital/labour ratio, it causes an equiproportional increase in the marginal productivities of both factors, whilst it has a factor-saving bias if it increases the marginal productivity of the other factor more than proportionally to the increase in the marginal productivity of the saved factor. Instead of the factor-saving bias one can define a factor-using bias: technical progress has a factor-using bias if it increases the marginal productivity of a factor (the used factor) more than proportionally to the increase in the marginal productivity of the other factor.

It is then clear that capital-using is synonymous with labour-saving, and labour-using with capital-saving.

An equivalent definition is that—at unchanged factor-price ratio—neutral technical progress leaves the optimum factor ratio unaltered, whilst a factor-saving progress reduces the optimum ratio between this factor and the other. In other words a labour-saving technical progress reduces the optimum labour/capital ratio (that is, relatively less labour is used), and a capital-saving progress reduces the optimum capital/labour ratio, always at unchanged factor-price ratio.

In Fig. 13.6, *QQ* is the typical isoquant before technical progress; given the factor-price ratio represented by the slope of the isocost *CC*, the optimum input combination is found at E, where the factor-price ratio equals the marginal rate of technical substitution, which in turn is equal to the ratio between the marginal productivities. After technical progress, the isoquant shifts to $Q'Q'$ and, given the isocost $C'C'$ parallel to *CC* (the same factor-price ratio), the new equilibrium point is found at E', which lies on the same ray *OR* as E. Therefore K/L is the same at E' as at E.

Let us now consider the case of labour-saving technical progress. In Fig. 13.7, *QQ* and $Q'Q'$ are the isoquants before and after technical progress. Since the marginal productivity of capital has increased by a greater proportion than the

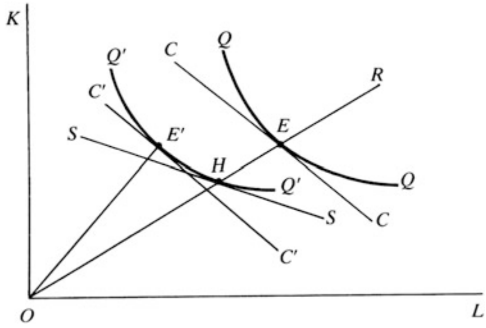


Fig. 13.7 Labour-saving technical progress

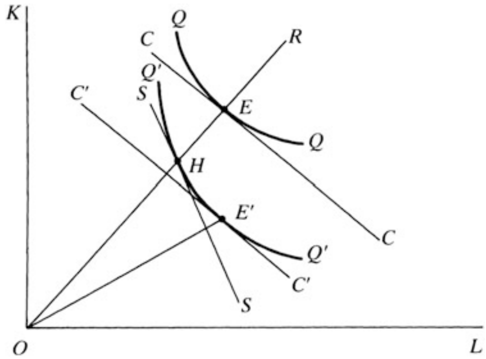


Fig. 13.8 Capital-saving technical progress

marginal productivity of labour, at the point of isoquant $Q'Q'$ where the K/L ratio is the same (point H), the MRTS (equal to *MPL*/*MPK*) is lower, as can be seen from the fact that *SS* is less sloped than *CC*. The new optimum input combination at unchanged factor-price ratio will be found to the left of H, for example at E', where the optimum K/L is higher and so L/K is lower.

Similarly, it can be checked (see Fig. 13.8) that in the case in which *MPL* has increased more than proportionally to the increases in *MPK* (capital-saving technical progress), the optimum K/L is lower at E' than at E.

13.5.2 Effects of Neutral Technical Progress on Production Levels and the Terms of Trade

The first result to be demonstrated is that neutral technical progress in a sector brings about—at an unchanged relative price of goods—an increase in the output of that sector and a decrease in the output of the other sector. For this purpose, as suggested by Findlay and Grubert (1959), we can use the Lerner-Pearce diagram (see Fig. 4.3). In Fig. 13.9 the isoquants of A and B are denoted by *AA* and *BB*, and the productive levels they represent correspond to the given commodity-price ratio (for details see

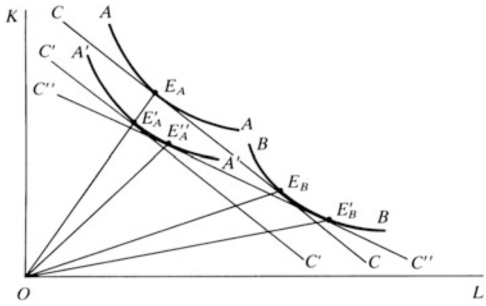


Fig. 13.9 Effects of neutral technical progress on factor intensities and price ratio

Fig. 4.3); the respective optimum input combinations are E_A and E_B. Let us assume that sector A enjoys a neutral technical progress: the AA isoquant shifts to $A'A'$ and, at unchanged factor-price ratio, the new optimum point is E'_A. However this is not a situation compatible with an unchanged commodity price ratio: in fact, at unchanged factor-price ratio, the same quantity of A (isoquant $A'A'$ represents the same output as isoquant AA, thanks to technical progress) now has a lower production cost (isocost $C'C'$ is below isocost CC), while the cost of producing the same quantity of B is unchanged; therefore, the exchange ratio (relative price) of the two commodities cannot remain unchanged.

For the commodity-price ratio to remain unchanged the factor-price ratio must change so that the cost of producing the quantity of B represented by isoquant BB and the cost of producing the quantity of A represented by isoquant $A'A'$ (which is the same as that represented by the old isoquant AA) are equalized. Graphically this amounts to finding a new isocost ($C''C''$) simultaneously tangent to $A'A'$ (at E''_A) and BB (at E'_B). The reader will note that capital intensity has decreased in both sectors and that the p_L/p_K ratio is lower (p_K/p_L is higher).

To sum up: given neutral technical progress in a sector, at unchanged relative price of commodities, the intensity of the factor used relatively intensively in that sector decreases in both sectors, and the relative price of this factor increases.

Let us examine the effects of neutral technical progress on the productive levels. For this purpose the box diagram is useful. In Fig. 13.10, let P be the initial point on the efficiency locus. The capital/labour ratio in sector A is clearly higher than in sector B and corresponds to the slope of OE_A and of OE_B, respectively, in Fig. 13.9.

Let us now inquire whether point P lies on the new efficiency locus which comes about as a consequence of technical progress. The answer is affirmative, since with first-degree homogeneous production functions, a neutral technical progress is equivalent to a mere renumbering of the isoquants: in other words, in Fig. 13.6, isoquant $Q'Q'$ occupies exactly the same place occupied by a lower-index isoquant before technical progress.

The system, however, cannot remain at P after technical progress has taken place. We have in fact seen above that, at an unchanged relative price of commodities, a neutral technical progress in sector A (the capital intensive commodity)

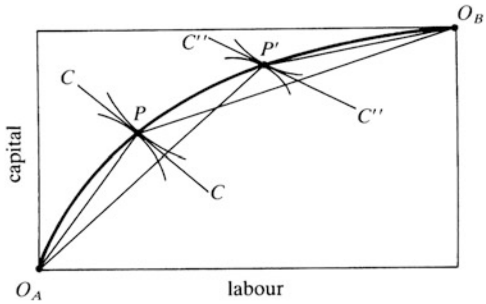


Fig. 13.10 Effects of neutral technical progress on outputs

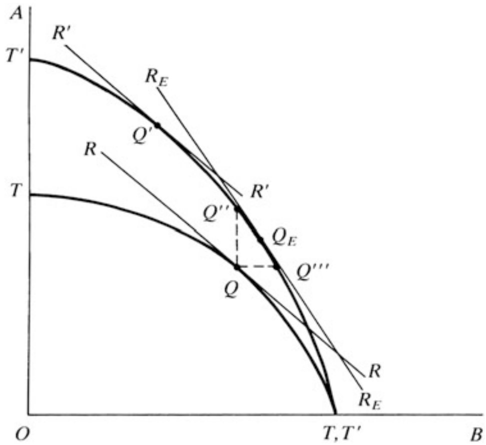


Fig. 13.11 Neutral technical progress and relative price of goods

causes a decrease in the capital intensity in both sectors. Therefore, the new equilibrium point will have to be somewhere to the right of P: in fact, the capital intensity will be lower in both sectors only at points on the PO_B stretch of the efficiency locus. Let P' be the new equilibrium point, where both sectors have a lower K/L ratio, corresponding to the slope of OE''_A and OE'_B respectively (Fig. 13.9). We observe that, at an unchanged commodity-price ratio, the output of commodity B is lower whilst that of A is higher.

We have thus proved the result stated at the beginning, that neutral technical progress leads to an increase in the output of the sector enjoying this progress and a decrease in the other sector's output, at unchanged relative price of commodities.

Point P', however, cannot be a point of general equilibrium if we bring demand into the picture. Technical progress brings about an increase in national income at constant prices: see Fig. 13.11, where the isoincome $R'R'$ represents a higher national income at constant prices than *RR*. Hence, if we exclude inferior goods, the demand for both commodities will increase. Now, since at an unchanged commodity-price ratio the output of B has decreased, there will be an excess demand for this commodity which will cause an increase in its relative price (and so a decrease in the relative price of A).

To sum up, neutral technical progress in a sector brings about a decrease in the relative price of the commodity produced by this sector.

This result can also be illustrated graphically by using transformation curves (as we did in the case of an increase in the endowment of a factor). In Fig. 13.11, *TT* is the initial transformation curve and $T'T'$ that which occurs as a consequence of neutral technical progress in sector A. Note that, as no technical progress has occurred in sector B, the intercept with the B axis of the new transformation curve is the same as that of the old one, because when all factors are employed in the B sector (where technology is the same) the maximum output of B remains the same.

At an unchanged commodity-price ratio (the line $R'R'$ is parallel to *RR*) the economy shifts from the equilibrium (production) point Q to Q', where the output of A is higher and that of B lower. As, assumedly, no commodity is inferior, point Q' (which corresponds to P' in Fig. 13.10) cannot be a general equilibrium point. The final equilibrium point will lie somewhere in the portion $Q''Q'''$ of the $T'T'$ transformation curve, where the outputs of both A and B are higher. At any such point—for example Q_E—the relative price p_B/p_A is higher (and so p_A/p_B is lower) than at Q'.

All this concerns the closed economy. As regards the open economy, we can follow exactly the same line of reasoning as in Sect. 13.4 with reference to Rybczynski's theorem. In fact, once we know that—at unchanged relative price of commodities—neutral technical progress in sector A brings about an increase in the output of A and a decrease in the output of B, and having assumed away inferior goods, we can proceed exactly in the same way as in Sect. 13.4 and show that the terms of trade p_B/p_A increase in any case, so that the situation will be better or worse according to whether A is the exportable or the importable.

This parallelism will be intuitive if one thinks that, as regards the effects on international trade, what matters is the increase in the domestic output of B and the decrease in the domestic output of A at unchanged relative price of commodities, as the causes of these changes in output (increase in the quantity of the factor used intensively in sector A or neutral technical progress in this sector) are irrelevant.

13.5.3 Effects of Biased Technical Progress

The effects of biased technical progress are more complicated, and we must distinguish the factor-saving technical progress occurring in the sector which is more intensive in the saved factor from that occurring in the sector which is more intensive in the other factor (the used factor).

13.5.3.1 Capital-Saving Progress in the Capital-Intensive Sector

As regards the former, we examine the case of capital-saving progress in the sector intensive in capital (the case of labour-saving progress in the labour-intensive

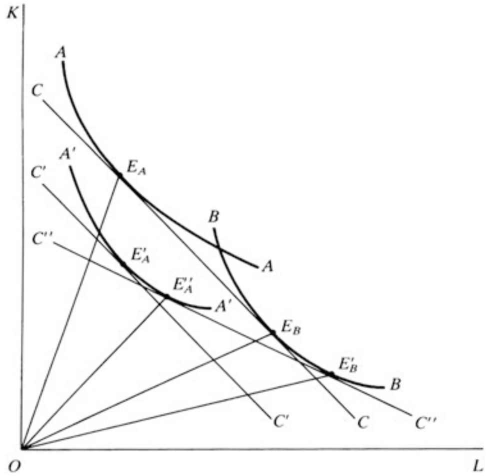


Fig. 13.12 Effects on factor intensities and price ratio of capital-saving technical progress in the capital-intensive sector

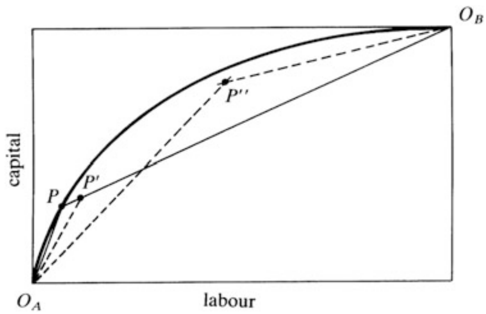


Fig. 13.13 Effects on productive levels of capital-saving progress in the capital-intensive sector

sector is perfectly symmetrical). In Fig. 13.12, which has the same structure as Fig. 13.9, capital-saving technical progress occurs in sector A (the capital-intensive commodity): the AA isoquant shifts to $A'A'$, as described in Fig. 13.8 and, at unchanged factor-price ratio, the new optimum point is E_A'. But, as the commodity-price ratio is assumedly unchanged, this point is not acceptable, as the same quantity of A now has a lower production cost whilst the production cost of B is unchanged; thus the exchange ratio (relative price) of the two commodities could not remain unchanged. It is then necessary for the factor-price ratio to change, so as to determine a new isocost ($C''C''$), tangent to both $A'A'$ (at E_A'') and BB (at E_B'): only in this way, in fact, will the production cost of the quantity of B represented by isoquant BB be the same as that of the quantity of A represented by isoquant $A'A'$ (which has the same index as isoquant AA).

It can be readily seen from the diagram that the capital intensity has decreased in both sectors, and that the p_L/p_K ratio is lower (p_K/p_L is higher). These effects are qualitatively similar to those found in the case of neutral technical progress in sector A (Fig. 13.9), and the effects on productive levels are also similar. In Fig. 13.13 P is the initial equilibrium point on the efficiency locus, with K/L ratios corresponding to the slopes of OE_A and OE_B in Fig. 13.12.

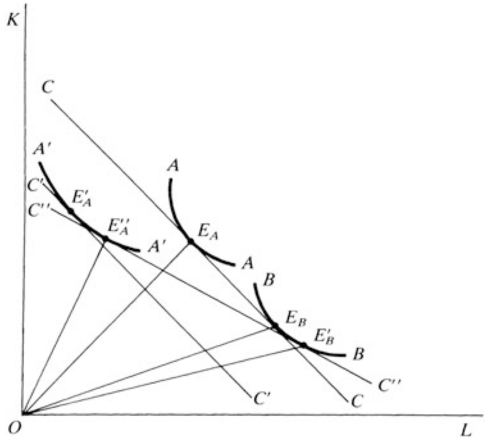


Fig. 13.14 Effects on factor intensities and price ratio of labour-saving technical progress in the capital-intensive sector: first case

As technical progress is biased, the new efficiency locus will not coincide with the old one, but it is possible to arrive at the results we are interested in without drawing it all. Let us begin by observing that point P'—at the intersection of the old ray O_{BP} (this has the same slope as OE_B in Fig. 13.12) and the ray O'_{AP} (this has the same slope as OE'_A in Fig. 13.12)—belongs to the new efficiency locus. In fact, point E'_A in Fig. 13.12 has been determined at unchanged relative factor prices, so that E'_A the isoquant $A'A'$ has the same slope as isoquant AA has at E_A. Now, given the property of radiality of first-degree homogeneous production functions (see Sect. 19.1), along ray O_{BP} the isoquants of B maintain the same slope, so that at P' the slopes of the isoquants of A and B (not shown in the diagram) are the same as the respective slopes at P and thus are equal (the A and B isoquants are tangent at P'): it follows that P' is an efficient point belonging to the new locus. It goes without saying that, as P' is nearer than P to the origin O_B, it represents a lower amount of B and, of course, a higher amount of A.

But, as we have shown above, point P' cannot be accepted if the relative price of commodities has to remain unchanged: from Fig. 13.12 we see that the capital/labour ratio further decreases in sector A, and decreases in sector B as well. Thus we shall get to a point P'' on the new efficiency locus (this is not drawn in the diagram for simplicity) such that: slope of O''_{AP} = slope of OE''_A in Fig. 13.12, and slope of O''_{BP} = slope of OE'_B in Fig. 13.12. As point P'' is still nearer to the origin O_B, we have proved that the output of B decreases whilst that of A increases. From this point onwards the analysis of the effects on international trade and on the terms of trade is identical with that explained with regard to neutral technical progress.

13.5.3.2 Labour-Saving Progress in the Capital-Intensive Sector

We must now examine the effects of labour-saving technical progress in the capital-intensive sector (the case of capital-saving progress in the labour-intensive sector is

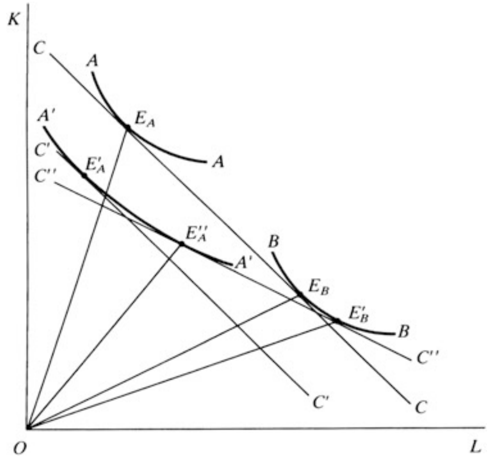


Fig. 13.15 Effects on factor intensities and price ratio of labour-saving technical progress in the capital-intensive sector: second case

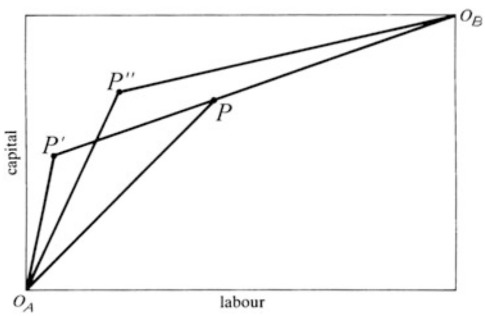


Fig. 13.16 Possible effect on productive levels of labour-saving technical progress in the capital-intensive sector

perfectly symmetrical). In Fig. 13.14, labour-saving technical progress occurs in the capital-intensive sector A. With the usual procedure, by now familiar to the reader, it can be seen that for the relative price of commodities to remain unchanged the isocost must shift to $C''C''$, whence a decrease in p_L/p_K. The capital intensity certainly decreases in sector B whilst the outcome in sector A is ambiguous. In Fig. 13.14 we have shown the case in which K/L increases in sector A; the opposite case in shown in Fig. 13.15 (the borderline case in which K/L remains unchanged in sector A is also possible, but unlikely). In any case K/L decreases in sector B. Thus the effect on productive levels is ambiguous. In fact, in the case in which the K/L ratio decreases in both sectors, the result will be the same as in the cases analysed above (the output of B decreases and that of A increases), whilst in the case in which K/L decreases in sector B but increases in sector A, it is possible (though not necessary) for the output of B to increase and that of A to decrease.

This possibility is represented in Fig. 13.16, where for simplicity's sake we have drawn only the equilibrium points: the initial one (P); the one corresponding to E'_A in Fig. 13.14 (i.e. P'), which is found at the intersection of ray O_BP with the ray having the same slope as OE'_A and is a point of the new efficiency locus; the one corresponding to the slopes of OE''_A and OE'_B in Fig. 13.14, i.e. P''. The equilibrium point P'' is farther than P from origin O_B (so that the output of B is higher) and nearer to origin O_A This is not sufficient for the output of A to be lower as we have to account for technical progress; it is however possible that the initial A isoquant through P shifts downwards by an amount insufficient to bring it below P'', so that we shall find that the isoquant through P'' has a lower index than that of the initial isoquant through P.

We must then conclude that in the case of a labour-saving progress in the capital-intensive sector, the outputs can move in any direction. As a consequence, the direction in which the terms of trade will move is indeterminate.

13.5.4 Conclusion

It may be useful to sum up the results concerning the effects of technical progress on the terms of trade.

1. Neutral technical progress in a sector causes a decrease in the relative price of that sector's product. The movement of the terms of trade will therefore be favourable (unfavourable) to the country if the sector concerned produces an importable (exportable).
2. Capital-saving technical progress in the capital-intensive sector and labour-saving technical progress in the labour-intensive sector have unambiguous effects, qualitatively similar to those of case (1): the relative price of the commodity produced in the innovating sector decreases. The terms of trade will therefore shift in favour of (against) the country if the innovating sector produces an importable (exportable).
3. Capital-saving technical progress in the labour-intensive sector and labour-saving technical progress in the capital-intensive sector have indeterminate effects, as the relative price of the commodity produced in the innovating sector may either increase or decrease. Note, finally, that once we know the effects of technical progress on the terms of trade we can determine the effects on the country's welfare: if the terms of trade improve, social welfare will certainly improve, whilst if they move against the country, there is the possibility of immiserizing growth (see Sect. 13.3.2).

References

Allen, R. G. D. (1967). *Macro-economy theory: A mathematical treatment* (chap. 13).

Bhagwati, J. N. (1958). Immiserizing growth: A geometrical note.

Bhagwati, J. N. (1973). The theory of immiserizing growth: Further application.

Bhagwati, J. N., Brecher, R. A., & Hatta, T. (1984). The paradoxes of immiserizing growth and donor-enriching "recipient-immiserizing" transfers: A tale of two literatures.

Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (chaps. 29).
Burmeister, E., & Dobell, A. R. (1970). *Mathematical theories of economic growth.*
Chacholiades, M. (1978). *International trade theory and policy* (Pt. V).
Corden, W. M. (1971b). The Effects of trade on the rate of growth.
Edgeworth, F. Y. (1894). *The theory of international values, I, II, III* (pp. 40–42).
Findlay, R. (1984). *Growth and development in trade models*.
Findlay, R., & Grubert, H. (1959). Factor intensities, technological progress and the terms of trade.
Ghosh, D. K. (1984). *Trade, distortions and growth.*
Hatta, T. (1984). Immiserizing growth in a many-commodity setting.
Hicks, J. R. (1932). *The theory of wages* (2nd ed., 1963).
Johnson, H. G. (1955). Economic expansion and international trade.
Johnson, H. G. (1959). *Economic development and international trade*.
McDermott, J. (1985). The rybczynski theorem: Short-run adjustment.
Rybczynski, T. M. (1955). Factor endowment and relative commodity prices.
Takayama, A. (1972). *International trade: An approach to the theory* (Pt. IV, chaps. 9, 10).
Woodland, A. D. (1982). *International trade and resource allocation* (chaps. 13, 14).

Chapter 14
International Trade and Growth: Dynamics

14.1 Dynamic Models

The models examined in the previous chapter analyse the relations between trade and growth (the latter taken as exogenous) in a comparative-static context, but for a more in-depth analysis of these relations explicitly dynamic models are necessary, as stated in Sect. 13.1. Several attempts have been made in this direction, the early ones consisting in an extension of Harrod's growth model to an open economy, the later ones in an extension of the two- and multi-sector growth models (including optimal growth models), whether neoclassical or non-neoclassical, to an open economy.

Anyone even vaguely acquainted with the immense literature on growth that flourished in the late 1950s and in the 1960s (for representative textbooks, see, e.g., Burmeister & Dobell, 1970; Wan, 1971), and who has followed the present textbook, will realize that there can be many more models of growth and trade than there are of growth in a closed economy, as each of the latter can be "opened" with several alternative assumptions on the international side (different theories on the causes of international trade, small or large country, free trade or tariff-ridden trade, etc.).

Even a survey, not to speak of an exhaustive treatment, of the topic under consideration would therefore require a volume to itself. What we can do is to give the reader a taste of this topic by means of a brief treatment of one of the possible ways of tackling it, consisting in a combination of the neoclassical model of international trade (or the Heckscher-Ohlin model, which is a particular case of it) with the traditional two-sector neoclassical growth model in a context of positive economics (see, e.g., Oniki and Uzawa, 1965; Johnson, 1971a, 1971b; Petith, 1974; Brems, 1980; Findlay, 1984). Models that emphasize the role of endogenous growth in international trade theory will be examined in Chap. 15.

The first point to emphasize is that in a context of growth we cannot ignore the fact that capital is not a primary factor of production but is a produced means of production, the increase in which does not come out of the blue but is determined by investment. The simplest way of accounting for this fact in our two-sector

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_14,

model is to assume that of the two commodities A and B, one (say, the former) is a fixed capital good, while the other (say, the latter) is a final consumption good, instead of both being final consumption goods. To simplify to the utmost, we assume away depreciation and technical progress, so that in a closed economy the increase in the capital stock coincides with the output of A, whilst in an open economy we have to add the imports (or subtract the exports) of this commodity. This concerns the production side; as regards the demand side, we can no longer assume that all income is consumed, but we must introduce a saving function, which, for simplicity's sake, we take as proportional to national income (thus ignoring the possibility of different propensities to save between different classes of income-earners, for example capitalists and workers). In the neoclassical growth model, saving is automatically invested in the purchase of the capital good, so that the domestic demand for A coincides with saving.

Finally, as regards the labour force, we assume that it grows exogenously, for example at a constant exponential rate depending on exogenous factors.

We now investigate, according to a well-established methodology, the existence and properties of the steady-state growth path and whether the system converges towards it. In Sect. 14.2 we examine the closed-economy situation; international trade will be introduced in Sect. 14.3.

14.2 A Simple Closed-Economy Two-Sector Growth Model

Following Johnson (1971a, 1971b) we first ascertain the steady-state growth path; for this purpose, we use Johnson's diagram, reproduced in Fig. 14.1. The axes in the right-hand part measure the amounts of the two commodities per worker, A/L (which is, assumedly, investment per worker) and B/L (consumption per head). The curve T_2T_2 is the transformation curve corresponding to a given stock of capital per worker in the economic system; a different curve corresponds to each different stock of capital per worker (for graphic convenience we have drawn only another such curve, T_4T_4, corresponding to a higher stock of capital per worker).

Given the relative commodity price, p_B/p_A, we can determine the equilibrium point on the transformation curve in the usual way; with reference to the curve T_2T_2 and assuming that the commodity price ratio corresponding to the steady-state equilibrium is equal to the slope of $R_{I2}R_{B2}$, the equilibrium point is E. Since, as we know, $R_{I2}R_{B2}$ can be interpreted as an isoincome, its intercept with the vertical axis, OR_{I2}, represents real national income (product) in terms of the capital good.

If we keep the commodity price-ratio constant and consider the different transformation curves corresponding to different stocks of capital per worker, we can determine the locus of equilibrium points, i.e. of the points of tangencybetween a transformation curve and an isoincome (parallel to $R_{I2}R_{B2}$). This locus is given by the line *DD*, which Johnson (1971a, 1971b) called the *Rybczynski line*;

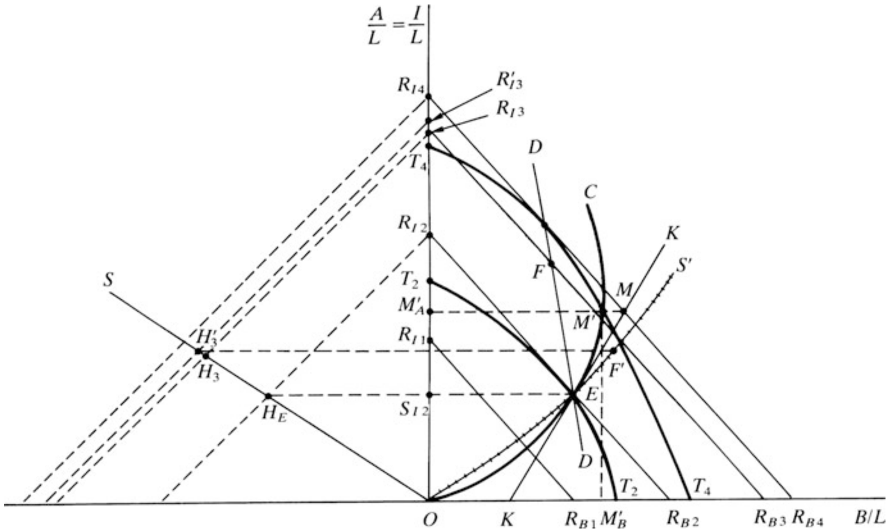


Fig. 14.1 Steady-state growth in a closed economy

its negative slope reflects the assumption that commodity A is capital-intensive.[1] In fact, the increase in a factor brings about—at given relative prices of factors and of commodities—an increase in the output of the commodity intensive in that factor and a decrease in the output of the other commodity (Rybczynski's theorem: see Sect. 5.4). Reformulated in per-capita terms, this means that the increase in the economy's stock of capital per head brings about an increase in the output of the commodity with a higher capital/labour input ratio (assumedly A) and a decrease in the output of the other commodity (the consumption good).

The line KK represents the investment requirements locus, where by investment requirements we mean the investment per worker required to maintain a constant stock of capital per worker, that is to provide each new worker (recall that the labour force grows continuously) with the same amount of capital as the existing workers. More precisely, since an isoincome goes through each point of the KK line, the ordinate of the intersection represents the investment per head required to provide the new workers with the same stock of capital per head as that which equips the existing workers, who produce the income represented by that isoincome. For example, in relation to point M, an investment per head equal to OM'_A is required to maintain the stock of capital per worker at the same level as that which equips

[1] We recall from growth theory that a sufficient stability condition in the two-sector neoclassical growth model is that the sector producing the capital good is less capital-intensive. When—as in the present case—this condition is not fulfilled, other sufficient stability conditions come into play (concerning the elasticity of substitution: see, for example, Gandolfo, 1971, pp. 454–455). We assume that these are satisfied.

the existing labour force which produces the national income (in real terms) OR_{I4}. Since at each point above (below) the KK line, investment per head is higher (lower) than at the corresponding point on KK, it follows that at all points above (below) this line the economy's stock of capital per head is increasing (decreasing).

Since the transformation curve tangent to the isoincome line $R_{I4}R_{B4}$ is T_4T_4, the output per head of the consumption good (corresponding to OM'_A of output per head of the capital good) is OM'_B, that is, the abscissa of point M'. The locus of all point like M' is the consumption-per-head possibilities curve OC, which represents the per-capita consumption possibilities of the economy as the stock of capital per head varies. This curve, initially upwards sloping, after a certain point bends back on itself, to denote that there exists a maximum attainable level of consumption per head,[2] after which further increases in capital per head have negative effects. Let us now consider the left-hand part of Fig. 14.1. The straight line represents the per-capita saving function, constructed so that its intersection with the 45° line drawn from a certain per-capita income gives the corresponding per-capita saving. For example, given the per-capita income OR_{I2}, the intersection under consideration occurs at point H_E which has an ordinate equal to OS_{I2}. This is the per-capita saving, which, as we know, coincides with per-capita investment, that is, with the output per head of commodity A (i.e. with the ordinate of the equilibrium point E on the transformation curve T_2T_2: this point, in fact, gives rise precisely to the per-capita income OR_{I2}).

As the stock of capital per head increases, the transformation curve shifts upwards and, as we said above, the succession of the production points (at unchanged commodity-price ratio) lies on the Rybczynski line DD. However, as the consumption good is not an inferior good,[3] the per-capita demand for this commodity increases as per-capita income increases, so that DD (along which B/L decreases) cannot be a locus of equilibrium points when we bring demand into the picture. The pressure of excess demand for B will cause an increase in the relative price p_B/p_A, so that the isoincome lines will take on a higher slope (in absolute value) with respect to the B/L axis and, consequently, their intercept with the A/L axis will increase in correspondence to any given transformation curve, that is, real national income (in terms of the capital good) will be higher. For example, given the production point F (for simplicity's sake, the underlying transformation curve is not shown) the $R_{I3}R_{B3}$ isoincome will shift changing its slope as described, so that its intercept with the vertical axis will no longer be R_{I3} but, say, R'_{I3}, which is above R_{I3} (the new isoincome is not shown to avoid undue graphic complications).

---

[2]This maximum occurs when an additional unit of capital adds to the potential output of investment goods exactly as much as it adds to the investment requirements. Since the former quantity is the own marginal product of capital (the real rate of interest) and the latter the growth rate of population, we have the golden rule: consumption per head is maximized when the (real) rate of interest equals the (exogenous) growth rate of population.

[3]In the case in which both commodities are consumption goods, one of them can, in principle, be an inferior good. In the case in which only one consumption good exists, it is difficult to believe that it can be inferior and, in fact, it cannot be if its marginal utility is always positive (insatiability).

Consequently, given the *OS* line in the left-hand part of the diagram, per-capita saving will be higher, as it will no longer correspond to point H_3 but to point H_3'. By projecting this point to the A/L axis and from here to the transformation curve (not shown in the diagram) we obtain the actual equilibrium production point, denoted by F'. The locus of all such points is the actual savings (or investment supply) curve OS'.

It can now be ascertained that the point of steady-state growth equilibrium is E. It is, in fact, the only point that lies simultaneously on all the curves. As it is on the *KK* line, the stock of capital per head does not change, the transformation curve does not shift and in the following period the situation repeats itself: stock of capital, output, consumption, saving, labour force, all grow at the same proportional rate; thus the per-capita variables are unchanged and the relative price of commodities does not change.

It can also be shown that E is a stable equilibrium, so that the economic system will converge towards the steady-state growth path. Let us consider a point other than E, for example, point F. As we have seen above, it is not a general equilibrium point, and the system will move to F' on the OS' curve. But F' is below the *KK* line, so that the economy's stock of capital per head decreases, the transformation curve shrinks towards the T_2T_2 curve and the economic system converges towards E. Similarly if we take a point below E we see that the system first moves to a point on the OS' curve, but since this point is above the *KK* line, the economy's stock of capital per head increases, the transformation curve blows up towards the T_2T_2 curve and the economic system converges to E.

14.3 Extension to an Open Economy

Equipped with this graphic representation of the two-sector neoclassical growth model in a closed economy, we can tackle the problem of the relations between growth and trade in a dynamic context. To simplify the treatment we assume that the growing country is small, so that the terms of trade are given. We must distinguish two cases, according as the terms of trade $p = p_B/p_A$ are higher or lower than the autarkic price ratio.

In Fig. 14.2 we have reproduced the steady-state growth situation E from Fig. 14.1. Let us first consider the case in which the terms of trade are higher than the commodity price ratio in autarky, for example equal to the slope of the *RR* straight line. The production point shifts from E to E'; given the income *OR* there will be a saving corresponding to the ordinate of point H_E' and the consumption point will be C', so that the country will export the consumption good and import the investment good. It is clear that the new, higher per-capita saving and investment will bring about an increase in the per-capita stock of capital of the economy; the transformation curve will shift upwards and the production point will move from E' upwards and to the left along the new Rybczynski line $D'D'$. At the new commodity price ratio, the actual savings curve will be OS' (we have drawn it as a straight

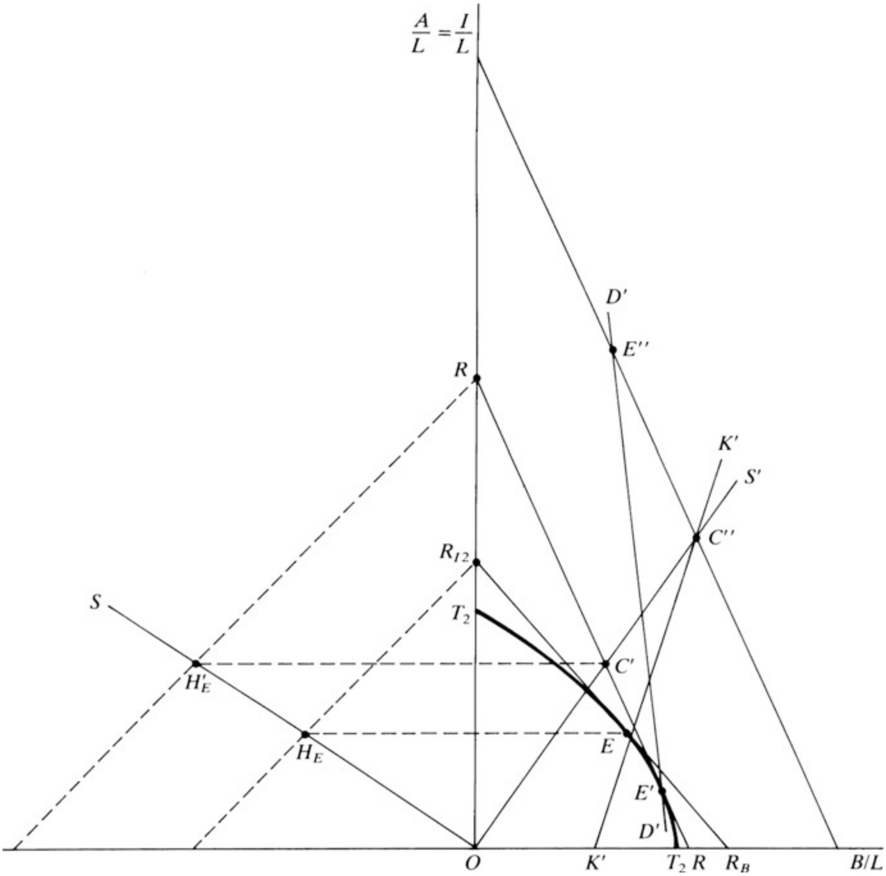


Fig. 14.2 Steady-state growth in an open economy

line for graphic convenience), whilst the investment requirements locus shifts to $K'K'$ because, for any given investment requirements, it is possible to have a higher consumption per head thanks to international trade.

It may happen that the new equilibrium will be reached at the intersection of the new saving curve with the new investment requirements curve, thus maintaining a situation of incomplete specialization. This situation may be similar, as regards the pattern of trade, to that occurring at the opening of trade (that is, the country remains an exporter of the investment good and an importer of the consumption good), or may show an inverted pattern, as in the case of Fig. 14.2, where, at point C'', the country becomes an exporter of the investment good and an importer of the consumption good (the production point is E''). Finally, the country may also specialize entirely in the production of the investment good.

The second case to be considered is when the terms of trade are lower than the relative commodity-price in autarky. In this case the production point shifts to the left of E along the transformation curve and the per-capita income measured in terms of the investment good decreases; consequently saving per head decreases and the savings curve shifts below E. The investment requirements curve shifts to the right as in the first case and for the same reasons. The country will export the investment good and import the consumption good.

As the decrease in per-capita savings brings about a decrease in the stock of capital per head, the transformation curve shrinks and the production point will shift downwards along the Rybczynski line.

In the final equilibrium, the country may be either incompletely specialized (with the same or a reversed pattern of trade with respect to that occurring at the opening of trade) or completely specialized in the production of the consumption good.

References

Brems, H. (1980). *Inflation, interest and growth* (chap. 11).
Burmeister, E., & Dobell, A. R. (1970). *Mathematical theories of economic growth.*
Findlay, R. (1984). *Growth and development in trade models*.
Gandolfo, G. (1971). *Mathematical models and methods in economic dynamics.*
Johnson, H. G. (1971a). Trade and growth: A geometric exposition.
Johnson, H. G. (1971b). The theory of trade and growth: A diagrammatic analysis.
Oniki, H., & Uzawa, H. (1965). Patterns of trade and investment in a dynamic model of international trade.
Petith, H. C. (1974). The pattern of trade in a neoclassical growth model with a general saving function.
Wan, H. Y., Jr. (1971). *Economic growth*.

Chapter 15
Endogenous Growth and Trade, Old and New

15.1 Introduction

Equilibrium growth in the basic neoclassical growth model is exogenous: the steady state path, in fact, depends on factors such as the rate of growth of the labour force and technical progress. Both are exogenous: the labour force grows according to exogenous demographic factors, and technical progress is no more than an exogenous time trend.

The theory of endogenous growth (for a general treatment see Aghion & Howitt, 1998; Barro & Sala-i-Martin, 2004; Jensen and Wong, 1998; Romer, 1994; Solow, 1992) stresses the endogenous determination of technical progress, which actually means an endogenous determination of the main source of growth (hence the name of endogenous growth theory). The basic ideas were already present in the traditional neoclassical growth theory, but in endogenous growth theory they are at the centre of the stage.

Another point considered by endogenous growth theory is the absence of decreasing returns to capital. Hence from the point of view of the interrelations with international trade, endogenous growth is often associated with the 'new' trade theories, that usually take increasing returns and imperfect competition as their points of departure (see Chap. 7)

In the 2×2 classification given in Table 15.1, the model treated in Chap. 14 is at position (1,1) in the matrix, but all other positions are theoretically possible (the names of the authors are merely exemplificative, given the host of contributions now existing: for a survey see Long & Wong, 1998), although position (1,2), namely the association of traditional growth theory with the new trade theories, has been neglected as relatively uninteresting.

In the next section we consider position (2,1), namely endogenous growth in the context of the traditional theory of international trade, and then a model falling in the (2,2) cell.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_15,

Table 15.1 Growth theories and trade theories

	International trade theory	
Growth theory	Traditional	New
Traditional	Oniki-Uzawa	–
Endogenous	Findlay	Grossman-Helpman

15.2 A Small Open Economy with Endogenous Technical Progress

The endogenization of technical progress can be performed in several ways, such as the accumulation of experience in the form of learning by doing, or the allocation of resources to R&D (Research and Development). Here we consider the second option.

The model (Findlay, 1995) is an extended Heckscher-Ohlin model with three sectors, in which growth is entirely due to technical progress while the amounts of the primary factors (capital and labour) are assumed constant. Two of the three sectors produce two final goods, say A and B, which can be traded along the lines of the Heckscher-Ohlin model; since the economy under consideration is assumed to be a small open economy, the terms of trade or relative price p_B/p_A is exogenously given by the international market. Production takes place using capital and labour under constant returns to scale and neutral technical progress.

The third sector produces a nontradable good, say Z (on non-traded goods in general see Sect. 6.6) and is the crucial one. It is the R&D sector, which can be considered as the sector that "produces" technical progress by using primary factors. More precisely, this sector employs capital and labour to provide R&D services to the traded goods sectors to increase their efficiency. Technical progress is purely 'local', as it only accrues to domestic firms, with no international spillovers.

The production of R&D services takes place under constant returns to scale, but the increase in efficiency that accrues to the tradables when there is an increase in these services is subject to diminishing marginal productivity. Since the same primary factors are used to produce the three goods, and production functions are homogeneous of the first degree, it follows that at some initial time t_0 (in which we can take the index of technological efficiency as equal to one) the relative price of the nontradable, p_Z, is also determined by the international market for traded goods (see Sect. 6.6).

We are now facing a problem of optimal allocation of resources, whose solution will yield the rate of endogenous technical progress. In fact, there is a trade-off between current and future outputs of tradable goods (on which social welfare ultimately depends). Since the amount of primary factors is given and constant through time, if more factors are allocated to the R&D sector, there will be less current output of tradables but more output of them in the future due to the higher rate of technical progress. If less factors are allocated to R&D, there will be more current output but less future output of tradables. Let v be the value of the output

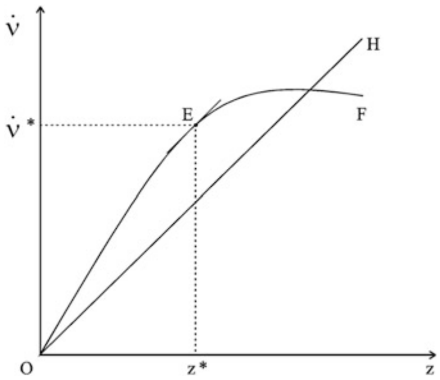


Fig. 15.1 Traditional trade theory and endogenous growth

of tradables and $\dot{\nu}$ its instantaneous change, a function of z, the per-capita output of R&D services.

As in all optimization problems involving trade-offs, we can apply the usual optimization rule that equates the marginal benefit to the marginal cost of an incremental expenditure on R&D. The marginal benefit is d$\dot{\nu}$/dz, the increment in $\dot{\nu}$ due to an increment in z. Since this benefit accrues from now to infinity, its present value is $(\mathrm{d}\dot{\nu}/\mathrm{d}z)/\delta$, where δ is a discount rate (the interest rate or the social discount rate). The marginal cost is simply p_Z, the relative price of the non-traded good. Marginal benefit and cost are equated when $(\mathrm{d}\dot{\nu}/\mathrm{d}z)/\delta = p_Z$, namely $\mathrm{d}\dot{\nu}/\mathrm{d}z = \delta p_Z$.

Leaving the mathematical treatment to the Appendix (Sect. 29.1), the result can be shown in the Fig. 15.1, taken from Findlay (1995, p. 89). The curve *OF* shows $\dot{\nu}(z)$, the increase in value of tradable output as a function of z. This curve embodies the trade-off between the reduction in current output of tradables and the enhancement of technology. As z increases, the technological improvement more than offsets the decrease in current tradable output, but only up to a certain point, after which further allocation of resources to the R&D sector will have a negative effect.

The ray *OH*, whose slope is δp_Z, shows the interest cost of z, which is $\delta p_Z z$. Marginal benefit and cost are equated when the slope of the *OF* curve (namely $\mathrm{d}\dot{\nu}/\mathrm{d}z$) equals the slope of *OH*, an equality that occurs at point E. The endogenously determined (optimal) per-capita output of R&D is z^*, which determines the rate of technological progress and hence of growth. This shows that the growth rate of the economy is endogenous.

15.3 Endogenous Growth, North-South Trade and Imitation: A New Version of the Product Cycle

Just as there is a wealth of endogenous growth models in closed economies, so there is a wealth of models of endogenous growth in open economies. These are often associated with the new theories of international trade, although this is not a necessity (see above, in the Introduction).

The voluminous literature on endogenous growth and trade is surveyed in Long and Wong (1998); in this section we present a model due to Grossman and Helpman (1991a, 1991b, 1991c), that formalizes the ideas set forth in the Hirsch-Vernon product cycle (see above, Sect. 8.3) and in Posner's technological gap (see Sect. 8.2).

Consider, for example, the product cycle of the personal computer in the 1980s and 1990s. This has been characterized not only by an increasing off-shore production in South by the Northern innovating (multinational) firm that invented the product—as predicted by the product cycle—but also by the introduction of imitations or "clones" by competitors located in NICs (newly industrialized countries) of South—as predicted by Posner's theory.

Product innovation can take place through the development of new varieties of horizontally differentiated goods or through the improvement in the quality of a set of vertically differentiated goods. Both cases are considered by Grossman and Helpman; we examine the second model because it allows the study of some additional aspects of actual North-South trade with imitation not considered in the first model. In fact, taking up again the illuminating example of the PC (Grossman & Helpman, 1991a, chap. 12), the clones of the original machine (which was based on the 8086 microprocessor) were displaced by new and superior machines developed in North, based on the 80286 microprocessor. These new machines were subsequently imitated in South and then upgraded once more by firms in North, and so on and so forth. This shows that there may be reversals in the pattern of specialization when innovative products, after becoming standardized and being copied, become obsolete due to the introduction of a higher-quality type.

The basic model considers two countries, North (that has an absolute advantage in innovation, namely new higher-quality products can only be developed there) and South (that has an absolute advantage in production costs, namely a lower wage rate, hence an absolute advantage in imitation). *Innovation* in North does of course require the allocation of resources to R&D, and is a risky process in the sense that when a firm devotes resources to R&D it has a probability of success (i.e., of developing a higher-quality product) proportional to the scale of its efforts but smaller than unity. *Imitation* in South is also treated as a risky R&D process requiring resources with an associated probability of success.

Three types of firms are distinguished:

(i) Northern *leaders*, namely firms that have exclusive ability to produce some state-of-the-art product (this is the top-of-the-line product, namely the currently highest quality of the commodity) and compete with another Northern firm (a *follower*) that can produce the second highest quality;
(ii) Northern *leaders* competing with a Southern firm that can produce the second highest quality;
(iii) Southern firms that can imitate and produce a state-of-the-art product.

In the presence of imitation threats, we must distinguish two cases. When imitation is successful, Northern leaders have the incentive to undertake research leading to innovation, namely to the development of the next generation of products so as to regain market leadership. Due to the greater accumulated knowledge, only Northern leaders do that. However, when a product has escaped imitation (let us recall that imitation activity is not always successful) also Northern followers have an incentive to undertake research leading to the development of the next generation of products, as in such a situation they stand to gain more from a research success than do leaders.

The model is rather complex, and its results can be found only by a formal analysis (see the Appendix to this chapter, Sect. 29.2). They can be summarized as follows.

In steady-state growth, two main types of equilibria may occur. In the first type, followers are relatively efficient at innovation (though less so than leaders), hence both leaders and followers engage in innovation. This equilibrium gives rise to a complex history of product cycles because at any moment the market leadership can pass from one Northern firm to another (formerly a follower) or from North to South (when imitation in South is successful and R&D in North fails to develop a higher-quality product). Product cycles go back and forth.

The second type of equilibrium (the inefficient follower case) occurs when followers have a relatively large inferiority in the research lab with respect to leaders, so that only these latter carry out R&D in the steady state. In this case the outcome is more clear-cut, as there will be alternating phases of production between North and South, with Northern firms developing new products and being market leaders until Southern firms displace them thanks to successful imitation, after which there will be another innovation by a Northern firm and so on.

Grossman and Helpman (1991a, chap. 12) also examine the consequences on world growth and trade of subsidies to R&D by either the Northern or Southern government. The results depend in an essential way on the type of equilibrium that obtains.

In the inefficient follower case, technological progress and hence growth is favourably affected by the introduction of a research subsidy by either government. Not only a research subsidy by the Northern government to its innovative firms fosters technological progress, but also a higher pace of imitation (brought about by a subsidy by the Southern government to its firms) has the same effect, causing Northern firms to increase their research efforts to regain market leadership after losing it to Southern imitators. This result is the same that can be obtained in a similar model of North-South trade with imitation in which, however, product differentiation is of the horizontal type, so that innovation consists of the development of new varieties of the product (Grossman & Helpman, 1991a, chap. 11).

Results of government intervention may however be strikingly different in the efficient follower case: "In this case an expansion in the size of South may slow down the rate of innovation in North, and policies that might be used to promote domestic productivity gains spill over abroad with adverse consequences for the foreign rates of technological progress" (Grossman & Helpman, 1991a, p. 327). As in the case of strategic policies in a static context, results of government intervention in a dynamic context are heavily model dependent, which comes as no surprise.

For other models of endogenous growth and trade see Long and Wong (1998).

References

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory.*
Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth.*
Findlay, R. (1995). *Factor proportions, trade, and growth* (chap. 3).
Grossman, G. M., & Helpman, E. (1991a). *Innovation and growth in the global economy.*
Grossman, G. M., & Helpman, E. (1991b). Endogenous product cycles.
Grossman, G. M., & Helpman, E. (1991c). Quality ladders and product cycles.
Jensen, B. S., & Wong, K.-Y. (Eds.). (1998). *Dynamics, economic growth, and international trade*.
Long, N. V., & Wong, K.-Y. (1998). Endogenous growth and international trade: A survey.
Romer, P. M. (1994). The origins of endogenous growth.
Solow, R. M. (1992). *Siena lectures on endogenous growth theory.*

Part VI
Globalization

Chapter 16
Globalization and Economic Geography

16.1 Introductory Remarks

"Globalization" is a much used and abused word. According to The American Heritage® Dictionary of the English Language (copyright © 2009 Houghton Mifflin Company), to globalize means "To make global or worldwide in scope or application." In the field of international economics, globalization means different things to different people (see, for example, Gupta ed., 1997; Stern, 2009). Some authors (see, for example, Dreher, 2006; Dreher, Gaston, & Martens, 2008) have also suggested indexes to measure the degree of globalization of the various countries, taking into account the three main dimensions of globalization (economic, social, and political). These indexes are available at http://globalization.kof.ethz.ch/

A by no means exhaustive list of the elements that make up globalization is:

(a) The increase in the share of international and transnational transactions, as measured for example by the share of world trade and world direct investment (carried out by multinational corporations) in world GNP;
(b) The integration of world markets, as measured for example by the convergence of prices and the consequent elimination of arbitrage opportunities;
(c) The growth of international transactions and organizations having a non-economic but political, cultural, social nature;
(d) An increasing awareness of the importance of common global problems (the environment, infectious diseases, the presence of international markets which are beyond the control of any single nation, etc.,)
(e) The tendency to eliminate national differences and to an increasing uniformity of cultures and institutions.

The debate on globalization usually considers the following aspects:

1. The actual degree of integration of markets;
2. Globalization as a process that undermines the sovereignty of the single states, reducing their autonomy in policy making and increasing the power of multinational corporations;

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_16,

3. The effects of globalization on world income distribution, both within and across countries;
4. The possible development of an international government to cope with global problems.

The aspects of globalization concerning multinational corporations and their foreign direct investment have been treated in Sect. 6.8.4. Here we shall take "globalization", as referred to international trade, to mean the closer integration of world markets for commodities, services, and factors, partly due to the decrease in transport and communication costs (so called "*annihilation of distance*").

The importance of transport costs and location was already stressed by Ohlin himself: the title of the 12th chapter of his treatise (Ohlin 1933) is "Interregional Trade Theory as Location Theory", where he considers the role of location and transport costs in both domestic and international trade. An early attempt at integrating location theory and international trade theory was Lösch (1954).

The topic was taken up again by Paul Krugman (1991a, p. 1), who defined *economic geography* as "the location of production in space; that is, that branch of economics that worries about where things happen in relation to one another". Under this definition, location theory is part of the much broader field of economic geography, a field that would also include international trade theory as a special case. It would then seem quite natural to observe a close integration between international trade theory and location theory in the broader context of economic geography, but this has not been the case, for several reasons examined for example by Krugman (1991a, 1993a).

The present chapter examines the relations between location of production, cost of transport, and international trade in the context of both the traditional and the new theories of international trade.

16.2 Transport Cost, Location Theory, and Comparative Advantage

Location theorists classify industries into "materials (or resource) oriented" and "market oriented" according as to whether transportation costs impose location close to the source of raw materials or to the final consumer (see, for example, Beckmann, 1987).

The original sites of the heavy industry (Pittsburgh in the United States, Birmingham in England, the Ruhr in Germany), illustrate the need for the production of iron and steel to be carried out close to the iron-ore and coal fields. Hence it is no surprise that the heavy industry arose in those countries that were well endowed with the necessary mineral resources, countries which then became exporters of the products of the heavy industry. This is perfectly in line with the standard factor proportions theory, as transport costs caused those mineral resources to be almost immobile factors.

However, after the second world war the transport revolution involving giant bulk carriers has drastically altered the situation. This has created a pool of primary resources on which all countries can draw: the most striking example is Japan, that became a top industrialized country using imported raw materials from far-away locations. Thus the relevant factor endowments are again capital (including technology and human capital), and labour rather than the endowment of primary resources, which is due to geological accidents.

These ideas have been modelled by Findlay (1995, chap. 6, sect. 6.3), who considers a three-commodity, three-factor model with constant-returns-to-scale technology. The commodities are:

1. An "all purpose" commodity (A), that can be either consumed or invested, namely added to the stock of capital, and is taken to be the *numéraire*;
2. A pure consumer good (B);
3. A raw material (Z) that is used in fixed proportions in the production of A.

The factors are:

(i) Land, or natural resources (N), specific to the production of Z;
(ii) Labour (L), used in the production of all three commodities;
(iii) Capital (K) used only for A and B.

While both N and L are in fixed supply, the supply of capital is endogenous. Commodity A is assumed to be capital-intensive with respect to B, and turns out to be also resource-intensive. To show this, we observe that commodity A—apart from the amount of N directly used to produce Z which is specific to A—indirectly requires more N with respect to B. In fact, since K embodies the part of A that has been invested (hence K indirectly embodies N), it follows that A, being capital-intensive with respect to B, indirectly requires more N.

Consider now two identical countries except for the endowment of natural resources, which is larger in country 1. Since N_1/L_1 is greater than N_2/L_2, we speak of 1 and 2 as the resource-rich and resource-poor country, respectively.

The first step is the introduction of international trade in final goods only. The raw material Z is assumed to be non traded due to prohibitive transport costs when it is in unprocessed form; these costs disappear when it is embodied in the capital-intensive final good A. With these assumptions the model behaves like the standard 2×2 Heckscher-Ohlin model, hence country 1 will export the resource-intensive commodity A and import commodity B, while country 2 will export commodity B and import commodity A. In country 1 the A sector will expand and the B sector will contract, while the opposite will take place in country 2.

Thus, when there is free trade in final goods only, what happens is a higher extraction of the raw material input in the resource-rich country; this entails an increase in the capital stock to meet the needs of the higher output of the capital-intensive exportable commodity A. In the other country the opposite will happen: the resource sector shrinks because of the reduction in the output of the import-competing commodity A, with a corresponding decrease in the long-run capital stock. As Findlay (1995, p. 168) notes, "free trade clearly *enhances* the initial difference in wealth between the two countries based on the difference in natural resource endowment."

The second step is to allow free trade in all commodities, because a transport revolution takes place so that the resource input Z can be traded at zero transport cost like the two final goods. We now have a model with three traded goods (one of which is a factor of production), and three factors (one of which is traded). Given the assumption of internationally identical technologies with constant returns to scale, if we further assume that all three commodities are produced in both countries, factor prices will be equalized.

In the long-run equilibrium, agents must have the same per capita utility level in both countries; this implies that per capita income and per capita wealth (and hence total wealth, given the assumption of identical labour force), must also be equal. Total wealth is made up of two components, the capital stock and the capitalized value of the rents from the natural resources N used to produce Z.

Commodity- and factor-price equalization implies that the price of Z increases (with respect to the pre-trade situation), in the resource-abundant country 1 (where before trade it was lower than in the resource-scarce country), and decreases in the resource-scarce country 2. The resource sector shrinks in country 2 and expands in country 1, which implies that, in the long run equilibrium, the natural-resource component of wealth is greater in country 1 than in country 2. This in turn entails a greater long-run equilibrium capital stock in country 2 than in country 1, as total wealth must be equal in both countries.

The final result is that in the long-run equilibrium country 2 may become the *exporter* of the capital-intensive commodity A.

"In other words, the possibility of sharing on equal terms in a global pool for access to the intermediate input enables the resource-poor country to build up its capital stock per head to such an extent that it leads to a *reversal* of its former comparative advantage in the labour-intensive good. [...] It is now the less naturally well endowed countries that will have a higher proportion of physical capital per capita in their portfolios and will thus *export* the capital-intensive industrial goods on the basis of imported intermediate inputs as, for example, in the case of Japan" (Findlay 1995, pp. 170, 172).

16.3 The Core-Periphery Model

The Core-Periphery model, developed in Krugman (1991c), has sparked a new and rich stream of literature known as the new economic geography (NEG). In a way analogous to the Krugman model of Sect. 9.2.1, which showed the existence of international trade in the absence of comparative advantage, the core-periphery model shows that agglomeration may emerge even in the absence or exogenous differences between locations. This is what makes this model new with respect to the preexisting literature on economic geography.

As recalled by Krugman (1991c, p. 486), many of the ideas contained in the core-periphery model had appeared in the literature since the 1950s but were not to become formalized for long time. Indeed, one of the merits of the core-periphery model is that it embodied many of these ideas into a simple and yet rigorous model. This formalization has provided a bridge between economic geography and international trade that has been crossed by many scholars. At last, as wished by Ohlin et al. (Eds.) 1977, regional economics and international economics have begun an integration process that has shed new light on many issues in both fields.

We now move to the study of the Core-Periphery model. The objective of the model is to answer the question of why and when does manufacturing become concentrated in a few regions leaving the other regions relatively undeveloped.

16.3.1 Description of the Model

Consider a world composed of two "regions" indexed by $i = 1, 2$. Assume that there are only two factors of production represented by two distinct types of labour, "farmers" and "workers": farmers are geographically immobile while workers may move between regions.[1] For simplicity, it is assumed that the world population (the sum of farmers and workers) is constant and normalized to 1. The number of farmers and workers is assumed exogenous with $(1-\gamma)$ being the number of farmers and γ the number of workers in the world economy. Farmers do not migrate and are equally distributed between the two regions, each hosting $(1-\gamma)/2$ farmers. Workers may migrate and at any point in time there are γ_i workers in each region, naturally, $\gamma_1 + \gamma_2 = \gamma$. It is convenient to compact notation and define λ_i as the share of workers residing in region i at any point in time, i.e., $\lambda_i \equiv \gamma_i/\gamma$. The world economy produces two goods, a homogenous agricultural good, A, and a differentiated manufactured good, M. The technology is identical between regions so as to eliminate any exogenous difference between them. Good A is produced with a constant return to scale technology which requires one unit of labour input (of farmers) for one unit of output. Only farmers are used in the production of A. The market for A is perfectly competitive and A is freely traded between regions. Any variety of good M is produced by use of an increasing return to scale technology characterized by a fixed and a variable input. Specifically, the labour input (of workers) per q units of output is $l = F + cq$ where F is the fixed input and cq is the variable input.[2] The market for M is characterized by monopolistic competition (see Sect. 9.2.1). Good M is traded between regions at a cost. Trade costs are assumed to be of the iceberg type already introduced in Sect. 6.3. They consist in a deterioration of the goods transported by which only a fraction $\tau \in (0, 1)$ of each unit sent from

[1] We use the terminology adopted in the early new economic geography literature. Clearly though, "region" should be understood as a geographical unit (region, country, or else) and "farmers" and "workers" as a geographically immobile and mobile factor, respectively.

[2] See Ricci (1999) for an interesting extension where the marginal labor input, c, differs between countries.

region i arrives to region j. The costs of transporting a unit of any variety of good M is therefore $(1 - \tau)$ units of the variety transported.

Consumers (farmers and workers) draw utility from the consumption of A and M. Their consumption preferences are such that they spend a share γ of their income on good M and a share $(1 - \gamma)$ on good A. For simplicity, it is assumed that the expenditure share on A is exactly equal to the share of farmers in total population. Good M is differentiated and the sub-utility derived from consumption of M is given by the form already encountered in Eq. (9.2). We recall that the key feature of these preferences is the appreciation for variety per se. The consequence of this assumption is that consumers always choose to spread any given amount of aggregate consumption on the maximum possible number of varieties. Given this appreciation for variety, it is optimal for a firm to differentiate its product from that of any other firm. Product differentiation, in turn, gives to firms a market power that they exploit by setting prices above the marginal cost. Thus, the profit maximizing price for a firm located in region i applied to consumers in the same region is $p^*_{ii} = \mu c w_i$, where w is the manufacturing wage in region i and $\mu > 1$ is the mark-up over the marginal cost cw_i. The profit maximizing price of a firm located in region i and applied to consumers in region j is $p^*_{ij} = \mu (cw_i/\tau) > p^*_{ii}$. The mark-up in p^*_{ij} is μ as in p^*_{ii} but $p^*_{ij} > p^*_{ii}$ because the marginal cost of producing for the foreign market, cw_i/τ, is higher than the marginal cost of producing for the domestic market, cw_i, reflecting the fact that to sell one unit in the foreign market the firm has to produce $1/\tau$ units. Profits, π_i, are given by $\pi_i = p_{ii}q - w_i(F + cq)$, where q is the total output produced by the firm, including the fraction that is used as transport cost. Entry into the market is assumed to be free and occurs instantaneously so that profits are zero at any point in time. We apply the zero profit condition by substituting the profit maximizing price into π_i and setting $\pi_i = 0$ to give the equilibrium quantity of output produced by any firm: $q^* = \frac{F}{c(\mu-1)}$. Firm output is the same for all firms in any country, hence we have dropped the subscript. Free trade in good A leads to equalization of the price of A between regions. Let A be the numéraire good and normalize its price to 1 so that agricultural wages are also equal to 1 in both regions.

16.3.1.1 Instantaneous Equilibrium

We begin by setting the equilibrium conditions in the factor markets. The farmers' labour market is very simple. Since it takes one farmer to produce one unit of A, each country produces a quantity A equal to $(1 - \gamma)/2$. Total demand for workers in a region is obtained by multiplying individual firm demand, $l = F + cq^*$, by the number of firms in the region, n_i, to obtain $n_i(F + cq^*)$. The total supply of workers in the region at any point in time is given by the share of workers in that region, λ_i, multiplied by the total number of workers in the world economy, γ. Therefore, the equilibrium conditions in the market for workers are:

$$n_i \left(F + cq^*\right) = \lambda_i \gamma; \qquad i = 1, 2. \tag{16.1}$$

Replacing equilibrium output $q^* = \frac{F}{c(\mu-1)}$ into Eq. (16.1) gives $n_i^* = \frac{\lambda_i \gamma(\mu-1)}{\mu F}$. Computing the total number of varieties in the world, $N = n_1 + n_2$, gives $N^* = \frac{\gamma(\mu-1)}{\mu F}$. Finally, computing region $i's$ share of manufacturing output in world output, $\frac{n_i}{N}$, gives:

$$\frac{n_i}{N} = \frac{\gamma_i}{\gamma} \equiv \lambda_i \tag{16.2}$$

Turning to the goods market, we begin by obtaining the demand for any single variety. We shall do this intuitively in three steps (see Sect. 23.2.1 for the formal derivation of the demand functions).

First, we compute total regional income,

$$E_i = (1-\gamma)/2 + \lambda_i \gamma w_i. \tag{16.3}$$

The first summand is the total farmers' income (recall that farmers' wage is equal to 1), and the second summand is the total workers' income. Second, we recall that expenditure on manufactures is γ times total income, γE_i. Third, quite intuitively, the expenditure share of a single variety in the total expenditure on manufactures depends on the price of that variety relative to the price of the other varieties. To see the latter point more specifically, let P_i be an index (think of it as an average) of the prices of all varieties. Note that the price index bears the regional subscript i. This is because the index contains the price of all varieties, domestic and foreign, and the price applied abroad is $(1/\tau)$ times the domestic price. Therefore, unless regions produce an equal number of varieties, the number of varieties on which residents of region 1 pay the transport cost is different from the number of varieties on which the residents of region 2 pay the transport cost; consequently, the price indices are different. Furthermore, the price index P_i is necessarily decreasing in $\left(\frac{n_i}{N}\right)$ since the larger is $\left(\frac{n_i}{N}\right)$, the smaller is the number of varieties on which residents of i pay transport costs. The price of a variety relative to the average price is (p_{ii}/P_i) or $\left(p_{ji}/P_i\right)$ for a domestic and a foreign variety, respectively. The demand for a single variety is decreasing in such a relative price. The exact functional form of the demand emanating from residents of region i is $\left(\frac{P_i}{p_{ii}}\right)^{\frac{1}{\mu-1}} \gamma E_i$ for any domestic variety and $\left(\frac{P_i}{p_{ji}}\right)^{\frac{1}{\mu-1}} \gamma E_i$ for any foreign variety.[3] It is not surprising to find the mark-up μ in the demand functions. After all, the mark-up reflects the market power of producers, which is related to the rigidity of the demand for any particular variety. Equilibrium in the goods market requires that

$$p_{11}^* q^* = \left(\frac{P_1^*}{p_{11}^*}\right)^{\frac{1}{\mu-1}} \gamma E_1 + \left(\frac{P_2^*}{p_{12}^*}\right)^{\frac{1}{\mu-1}} \gamma E_2, \tag{16.4}$$

[3] See Sect. 23.2.1 for a formal derivation of demand functions from S-D-S preferences.

$$p_{22}^{*}q^{*} = \left(\frac{P_1^{*}}{p_{21}^{*}}\right)^{\frac{1}{\mu-1}} \gamma E_1 + \left(\frac{P_2^{*}}{p_{22}^{*}}\right)^{\frac{1}{\mu-1}} \gamma E_2. \tag{16.5}$$

On the left hand side of each equation we have the supply and on the right hand side we have the demand. At first sight it seems as if there are many endogenous variables in Eqs. (16.4) and (16.5) but, in fact, there are only two. This is because prices are constant multiples of wages; the price indices contain only λ_i and all prices; the firm's output is only a function of the parameters; and expenditure contains only λ_i and wages as we see in Eq. (16.3). Therefore, for any given λ_i, after substituting the expressions for prices, price indices, output, and expenditure into Eqs. (16.4) and (16.5) we are left with only two endogenous variables, namely, w_1 and w_2 (see Eqs. (30.1)–(30.3) in Sect. 30.1 for details). It is now clear that for any given value of λ_i, Eqs. (16.4) and (16.5) determine w_1 and w_2; but λ_i varies over time as workers migrate. This is what we have to address next.

16.3.1.2 Dynamics

Migration flows give the model a dynamics represented by the evolution of λ_i over time. Let $\dot{\lambda}_i$ be the migration flow into region i at a point in time and take region 1 as the reference region. Migration into region 1 is determined by the real wage difference:

$$\dot{\lambda}_1 = \omega_1(\lambda_1) - \omega_2(\lambda_1) \tag{16.6}$$

where, ω_i, is the real wage in region i given by

$$\omega_i = \frac{w_i}{(P_A)^{1-\gamma}(P_i)^{\gamma}}. \tag{16.7}$$

The notation $\omega_1(\lambda_1)$ and $\omega_2(\lambda_1)$ refers to the fact that real wages are determined by Eqs. (16.4) and (16.5) and, therefore, depend on the value of λ_1. We can see at this point the dynamics of the system.[4,5] At any instant in time the value of λ_1 is given and Eqs. (16.4) and (16.5) determine nominal wages (w_i). Once nominal wages are determined so are prices, price indices, real wages and the real wage

[4]This migration mechanism implies that migration decisions are taken by comparing current real wage differentials and neglecting the future evolution of real wages. For extensions of the core-periphery model that explicitly take account of expectations on future real wage differentials see Krugman (1991b), Baldwin (2001), and Ottaviano (1999, 2001).

[5]More complex functional forms may be used for Eq. (16.6) but this simple form suffices for our expositional purposes. Note also that it suffices to write only one law of motion since $\lambda_2 = 1 - \lambda_1$ at any time.

differential. The real wage differential in turn determines the migration flow ($\overset{\bullet}{\lambda}_i$) which leads to a new value of λ_i. This new value of λ_i will determine new values of wages via Eqs. (16.4) and (16.5). The new wages give new prices, new price indices and new real wages which in turn will determine a new migration flow and so on until either all workers have moved to one region or real wages have equalized. We refer to the case where all workers are in one region as the core-periphery geographical configuration since the region where all the workers have located hosts the world's manufacturing output (the industrial core) and the other region produces only the agricultural good (the agricultural periphery). There are, obviously, two possible core-periphery configurations, one in which region 1 has the industrial core and the other one where region 2 has it. The case of equalization of real wages is instead referred to as the dispersed geographical configuration since the manufacturing output is produced in both regions. In the remainder of the section we discuss the conditions under which one has, in the long run, the core–periphery outcome or the dispersed outcome.

To understand the economic logic of the circular causation in the model, we shall consider an initial geographical configuration where the two regions are identical, $\lambda_i = 1/2$, and refer to this geographical configuration as a "symmetric configuration". In the symmetric configuration each country is an exact replica of the other. We perturb this configuration by an exogenous change in λ_i and study the mechanisms that lead to a further, this time endogenous, change in λ_i. Consider, for instance, an exogenous increase in λ_i. The direction of the endogenous change in λ_i is determined by the relative strength of three distinct economic mechanisms which we now analyse.

The first mechanism is known as the *demand linkage* and works through the effect that the exogenous change in λ_i has on expenditure. As is clear by inspection of Eq. (16.3), an increase in λ_i causes an increase in total expenditure emanating from region i and a decline in the total expenditure emanating from the other region $\left(E_i \uparrow, E_j \downarrow\right)$. Although these change have the same absolute magnitude the net effect of is an increase in demand for varieties produced in region i and a decline in demand for varieties produced in j. This is easily verified by inspection of Eqs. (16.4) and (16.5) where we should recall that $p_{ij} > p_{ii}$ and that at the symmetric equilibrium $E_1 = E_2$ and $P_1 = P_2$. We have already encountered this effect in Sect. 9.2.4 where we dubbed it "home market dominance". In the present context, the circular causation between size of demand and location of firms makes that the home market dominance gives rise to the demand linkage. Indeed, since demand increases in i and declines in j, manufacturing prices increase in i and decline in j. With the wage being a constant proportion of the price that the manufacturing wage increases in i and declines in j. Therefore, ceteris paribus, the real wage differential $\left(\omega_i - \omega_j\right)$ increases inducing further migration into region i (λ_i increases endogenously).

The second mechanism is known as the *cost of living linkage* and works as follows. The initial perturbation in λ_i brings about an increase in region $i's$ share of the total number of varieties $\left(\frac{n_i}{N} \uparrow\right)$ which, in turn, causes a decline of the

Table 16.1 Agglomeration and dispersion mechanisms in the core-periphery model

Perturbation $\lambda_i \uparrow$		
Effect	Result	Channel
$E_i \uparrow, E_j \downarrow$	$\lambda_i \uparrow$	*Demand linkage*
$\frac{n_i}{N} \uparrow \rightarrow \left(P_i \downarrow, P_j \uparrow\right)$	$\lambda_i \uparrow$	*Cost of living linkage*
$\frac{n_i}{N} \uparrow$	$\lambda_i \downarrow$	*Market crowding*

price index in region i and an increase of the price index in the other region $\left(P_i \downarrow, P_j \uparrow\right)$. Therefore, other things equal, the real wage differential $\left(\omega_i - \omega_j\right)$ increases inducing further migration into region i (λ_i increases).

The third mechanism is known as *market crowding* and works through the competition among firms for regional demand. The increase $\left(\frac{n_i}{N} \uparrow\right)$ brought about by the perturbation ($\lambda_i \uparrow$) intensifies the competition for a given amount of expenditure in region i while relaxing it in region j. Therefore, prices tend to fall in region i and tend to increase in region j. Given that the wage is a constant proportion of prices, the real wage differential $\left(\omega_i - \omega_j\right)$ declines which induces migration into region j (λ_i decreases endogenously).

The schematic representation given in Table 16.1 summarizes the causal chain of the three mechanisms.

The demand linkage and the cost of living linkage push λ_i in the same direction of the perturbation. They are agglomeration mechanisms since they push the economy towards either of the core-periphery configurations. The market crowding effect pushes λ_i in the opposite direction from the perturbation. This is a dispersion mechanism since it pushes the economy towards the symmetric configuration.

The relative strength of these three mechanisms determines the actual direction taken by λ_i after the perturbation. Whether agglomeration or dispersion prevails in the long run depends on the value of three key parameters in the model: the transport cost (represented by $1 - \tau$), the intensity of appreciation for variety per se (reflected by μ), and the share of the manufacturing sector in world output (represented by γ). In the next sub-section we shall focus on the effect of transport costs. The reason for this interest is that changes in trade costs may be taken to represent changes in the degree of market integration, which is one of the principal subjects of investigation in international trade and regional economics and is a matter of great policy relevance.

16.3.1.3 Market Integration and Industrial Localization

We now discuss the principal result of the core-periphery model. For sufficiently high trade costs, the agglomeration mechanisms are weaker than the dispersion mechanism and the dispersed configuration emerges in the long run. For sufficiently low trade costs, the balance is reversed and one of the core-periphery configurations emerges in the long run. For intermediate levels of trade costs, there exist multiple possible long run configurations. Figure 16.1 shows three representative situations.

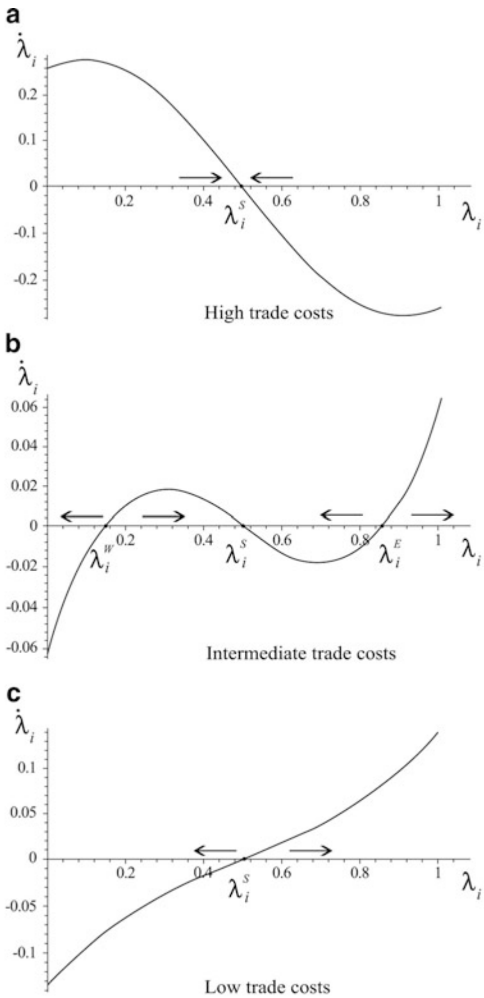


Fig. 16.1 Phase diagram of the core-periphery model

The figure plots $\dot{\lambda}_i$ against λ_i in the three cases of "high" trade costs (panel *a*), "intermediate" trade costs (panel *b*), and "low" trade costs (panel *c*).[6] The diagram in each panel is called the phase diagram, the line in the diagram is called the phase line and the arrows above the abscissa indicate the directions of motion of λ_i over time. Note that Eq. (16.6) tells us that $\dot{\lambda}_i$ is equal to the real wage difference. Therefore, the phase line also represents the real wage difference plotted

[6]The model determines the two values of τ that divide the set $(0, 1)$ into the three segments corresponding to 'high', 'intermediate' and 'low' trade costs. See Baldwin, Forslid, Martin, Ottaviano, and Robert-Nicoud (2003) for an exhaustive treatment of this matter.

against λ_i. Since the non-linearity of the model does not allow solving explicitly for the endogenous variables the phase line is obtained by numerical solutions. [7] Whenever the phase line is above the horizontal axis, the real wage is higher in region i and therefore workers will move towards that region (λ_i increases). Conversely whenever the phase line is below the horizontal axis the real wage is lower in region i and workers move towards region j (λ_i decreases).[8]

Consider first the case of high trade costs. The economy is initially in the symmetric configuration (λ_i^S) and is perturbed by an exogenous change in λ_i. No matter the direction and the size of the exogenous perturbation, the dynamics of the economy will bring λ_i back to the symmetric configuration. We conclude that for high trade costs the symmetric configuration is the only stable spatial configuration and, therefore, the world economy will be one where economic activity is dispersed.

Consider now the case of intermediate trade costs (panel b). Now the size of the exogenous perturbation matters. If the perturbation puts λ_i somewhere between λ_i^W and λ_i^S or between λ_i^S and λ_i^E, the dynamics of the economy will bring λ_i back to the symmetric configuration (the superscripts stand for East and West). If, instead, the exogenous perturbation is large such as to put λ_i between 0 and λ_i^W or between λ_i^E and 1, the dynamics of the economy will bring λ_i to 0 or 1, respectively. We conclude that, for intermediate trade cost, there are three stable spatial configurations: the two core-periphery configurations ($\lambda_i = 0$ and $\lambda_i = 1$) and the symmetric configuration ($\lambda_i = \lambda_i^S$).

Lastly, consider the case of low trade cost. In such a case the size of the perturbation does not matter. Any perturbation will bring the economy to one of the core-periphery configurations.

16.3.2 Conclusion

The core-periphery model is indeed very simple. Its simplicity has the merit of highlighting the key mechanisms that determine whether an industry agglomerates. All the mechanisms are endogenous and driven by the effect of migration on aggregate regional demand, on the price indices and on the demand for variety.

In the following sections we shall review some of the theoretical developments that have followed the core-periphery model.

---

[7] This is typical of many new economic geography models. See Ottaviano (2001) and Forslid and Ottaviano (2003) for explicitly solvable models. The appendix to this chapter provides an elementary guide to numerical solutions and calibrations.

[8] We are informally using a topological method for stability analysis. For a formal treatment of such a method see Gandolfo (2009, chap. 21, sect. 21.3.1).

16.4 Other Models

In this section, we review four variants of the core-periphery model. The first assumes the presence of a congestion force driven by the price of housing, the second introduces input-output linkages, the third highlights the role of diminishing returns to labour input in the agricultural sector and in the fourth the fixed cost is represented by a fixed input of mobile capital.

16.4.1 Housing Congestion

In this section we present the model developed by Helpman (1998) where the availability of a fixed stock of housing in each region gives rise to an additional force of dispersion. As people move into a region, housing becomes scarcer and its price increases thereby discouraging further inflow of migrants. Housing in this model stands in fact for any fixed stock of a non-tradeable resource. For clarity of exposition we refer to such a resource as housing.

16.4.1.1 Description of the Model

The world economy is composed of two regions indexed by $i = 1, 2$. Labour is the only one factor of production and the world endowment of labour is $\overline{L}$. Labour may migrate and $\lambda_i \equiv L_i/\overline{L}$ is the percentage of the world labour stock located in region i at any point in time. Consumers derive utility from consumption of a manufactured good, M, and from housing services (H). They spend a fraction γ of their income on good M and the remaining fraction on H. Profit-maximizing prices, equilibrium output, and demand functions in industry M are exactly as in the core-periphery model. So is the equilibrium condition in the labour market, $L_i = n_i\,(F + cq^*)$, from which we obtain

$$\frac{n_i}{N} = \frac{L_i}{\overline{L}} \equiv \lambda_i. \tag{16.8}$$

The only difference with the core-periphery model so far is that housing replaces good A and that there is no immobile factor of production. Each region is endowed with a constant stock of "Housing", labelled H_i. Given that consumers spend a proportion $(1 - \gamma)$ of total expenditure on housing, the equilibrium price of H is given by

$$P_i^H = \frac{(1-\gamma)\,E_i}{H_i}. \tag{16.9}$$

Total expenditure is the sum of labour income and income from local housing services $E_i = w_i \lambda_i \overline{L} + P_i^H H_i$. Replacing P_i^H in this expression gives

$$E_i = \frac{w_i \lambda_i \overline{L}}{\gamma}. \tag{16.10}$$

Equilibrium in the goods market gives conditions identical to Eqs. (16.4) and (16.5).

16.4.1.2 Dynamics

Migration flows are determined by real wage differences between regions and the real wage is

$$\frac{w_i}{\left(P_i^H\right)^{1-\gamma} (P_i)^{\gamma}}. \tag{16.11}$$

The main difference between the core-periphery model and the present model is apparent by comparing expression (16.11) with expression (16.7). In the core-periphery model, since the price of A is constant, only the price of manufactures matters for the purchasing power. In the present model the price of housing is not constant and therefore both prices matter. This gives rise to an additional circular causation mechanism since the price of housing in a region depends on the expenditure emanating from that region, as shown by expression (16.9). To understand this mechanism, consider again an exogenous perturbation that increases λ_i starting from the symmetric geographic configuration. Such a change in λ_i causes an increase in total expenditure emanating from region i and a decline in that of region j as shown by expression (16.10). Since expenditure on housing is $(1-\gamma)$ times total expenditure, the demand for housing in region i increases and decreases in region j. Therefore, P_i^H increases and P_j^H decreases, as shown by expression (16.9). As a result, ceteris paribus, we see from expression (16.11) that the real wage increases in j and declines in i, thus pushing labour to return in region i. This mechanism, which we may label the *cost of housing linkage*, is clearly a dispersion mechanism since it pushes λ_i in the opposite direction to that of the exogenous change.[9] The three mechanisms of the core-periphery model are in action in the present model too. Table 16.2 summarizes the causal chain of the four mechanisms.

[9]Since housing services are part of total consumption the cost of housing is part of the total cost of living. To keep terminology close to the literature, however, we continue to refer to the cost of living linkage as to the linkage driven by the price index of manufactures, P_i. We refer instead to the cost of housing linkage as to the linkage driven by the price of housing, P_i^H.

Table 16.2 Agglomeration and dispersion mechanisms in the housing congestion model

Perturbation $\lambda_i \uparrow$		
Effect	Result	Channel
$E_i \uparrow, E_j \downarrow$	$\lambda_i \uparrow$	*Demand linkage*
$E_i \uparrow, E_j \downarrow \rightarrow \left(P_i^H \uparrow, P_j^H \downarrow\right)$	$\lambda_i \downarrow$	*Cost of housing linkage*
$\frac{n_i}{N} \uparrow \rightarrow (P_i \downarrow, P_j \uparrow)$	$\lambda_i \uparrow$	*Cost of living linkage*
$\frac{n_i}{N} \uparrow$	$\lambda_i \downarrow$	*Market crowding*

16.4.1.3 Market Integration and Industrial Localization

Market integration gives different results from those of the core periphery model. First, if the expenditure share on housing is large, the cost of housing linkage is so strong that dispersion mechanisms always prevail on agglomeration mechanisms. In this case, the symmetric configuration ($\lambda_i = 1/2$) is the only stable spatial configuration for any value of trade costs. This means that market integration has no impact on localization of industries. If the expenditure share on housing is not very strong, agglomeration mechanisms prevail on dispersion mechanism for values of λ_i near the symmetric equilibrium, but the opposite occurs for values of λ_i far from the symmetric configuration. Another way of stating this result is that when the expenditure share on housing is small, agglomeration mechanisms prevail for low levels of agglomeration while dispersion mechanisms prevail for high degrees of agglomeration. This result is quite intuitive: as agglomeration progresses, the cost of housing linkage becomes stronger and eventually prevails on agglomeration mechanisms. The increase in the strength of the cost of housing linkage can be seen from Eqs. (16.10), (16.9), and (16.11) in the following way. Consider increasing levels of agglomeration towards region j. That means that λ_i decreases and approaches zero. Then, as we see from Eq. (16.10) expenditure in region i approaches zero and so does the price of housing as we see from Eq. (16.9). Consequently, as we see from Eq. (16.11) the real wage in i approaches infinity eventually attracting workers back to region i.

Market integration in this model has a reversed effect with respect to the core-periphery model. When trade costs are very low or zero, the cost of housing linkage prevails. Thus, for low trade cost, the symmetric configuration is the only stable spatial configuration. For high trade cost and if the expenditure share on housing is not too large, the symmetric configuration is unstable but there is one stable spatial configuration characterized by partial agglomeration on either side of the symmetric configuration (λ_i^W and λ_i^E).

These results are summarized in Fig. 16.2. Panel (a) shows the phase line in the case of high expenditure share on housing or low trade cost. Panel (b) shows instead the case of low expenditure share on housing and high trade costs. In either case the phase line approaches infinity when λ_i approaches 0 and approaches minus infinity when λ_i approaches 1. In the case depicted in Panel (b), the lateral configurations, λ_i^W and λ_i^E, are stable while the symmetric equilibrium λ_i^S is unstable and partial

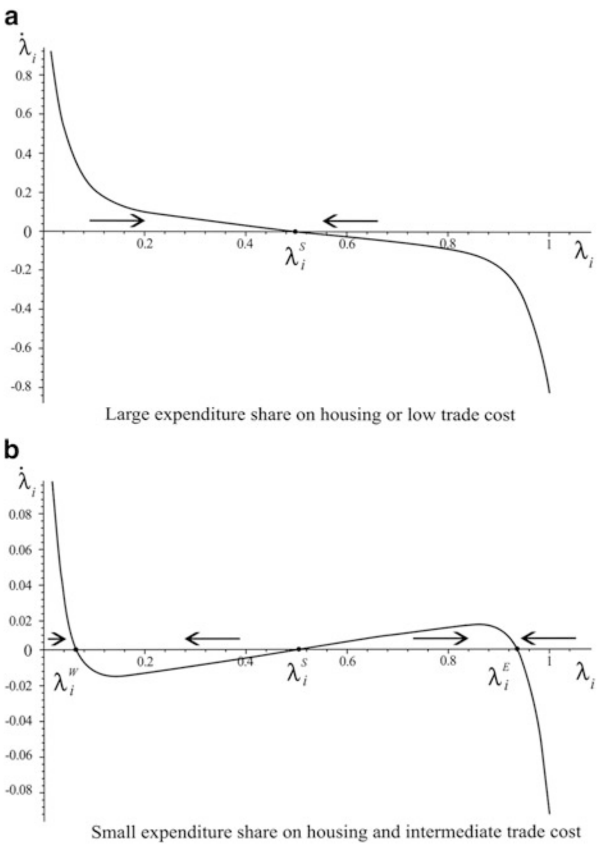


Fig. 16.2 Phase diagram under different expenditure shares on housing: Panel (**a**) high expenditure share; panel (**b**) low expenditure share

agglomeration occurs. Interestingly, the partial agglomeration configuration is one where population is unevenly distributed between regions despite the fact that the same amount of housing is available in both. As trade costs increase, the two lateral stable configurations λ_i^W and λ_i^E approach 0 and 1, respectively.

16.4.2 Input-Output Linkages

So far we have studied models where agglomeration and dispersion mechanisms were driven by migration. This may give the wrong impression that migration is a necessary assumption. We present here the model developed by Krugman and Venables (1995) where the assumption of factors migration is replaced by the presence of input-output linkages.

16.4.2.1 Description of the Model

The world is composed by two regions indexed by $i = 1, 2$. Labour (L) is immobile between regions and is used in both sectors. Production of A requires only labour and industry A is the same as in the core-periphery model. The production of any variety of good M requires instead a composite input Z produced by use of L and an aggregate of all varieties of M. Thus, part of the output of M is used as an intermediate input in the production of each variety of M.[10] To simplify matters we assume that each region is endowed with enough labour to potentially produce the world output of M and also some A. The convenience of this assumption is that both countries produce A in any geographic configuration. Thus, the price of A equalizes between countries and so does the wage in the A industry which we then normalize to unity, i.e., $w_{A1} = w_{A2} = 1$. The technology of production in M exhibits increasing returns to scale; specifically, it requires a fixed input (of Z) equal to F and c units of Z per unit of output. The requirement of Z per q units of output is therefore $F + cq$. The intermediate input Z is produced inside the firm and its technology is such γ_F is the optimal fraction of total cost represented by purchase of M and $(1 - \gamma_F)$ is the fraction representing the cost of L. The marginal cost of producing any variety increases with the wage paid in the M industry, w_{Mi}, and with the price index of manufactures, P_i. Any variety of M is traded internationally at iceberg costs. Let L_{Ai} and L_{Mi} denote employment in the A and M industry, respectively. National income is $I_i = w_{Ai}L_{Ai} + w_{Mi}L_{Mi}$. Consumers spend a fraction γ of their income on manufactures and the remaining fraction on good A. Firms spent a proportion γ_F of total cost on M but since profits are zero total costs equal total revenues. Aggregate revenue in the economy is $n_i p_{ii} q_i$. Thus, aggregate regional expenditure on M emanating from region i is

$$E_{Mi} = \gamma I_i + \gamma_F n_i p_{ii} q_i. \tag{16.12}$$

16.4.2.2 Dynamics

We begin by noting that world employment in the manufacturing industry, L_M, is constant, since world expenditure on the manufacturing industry is a constant fraction of total income, the latter being constant too. Since individual firm output is constant the total number of varieties, N, is constant too and in fixed proportion to L_M. We denote region i's share of the total employment in manufactures by $\lambda_{Mi} \equiv L_{Mi}/L_M$ and region i's share of the total number of varieties by $\eta_i = n_i/N$. Clearly, $\lambda_{Mi} = \eta_i$. Labour is assumed to move to the industry where the wage is the highest. We can therefore write the intersectoral labour flows as $\dot{\lambda}_{Mi} = w_{Mi} - w_{Ai}$, with

[10] See Alonso-Villar (2005) for an extension to this model where there are input-output linkages between two manufacturing industries and trade costs differ between industries.

Table 16.3 Agglomeration and dispersion mechanisms in the I-O linkages model

Perturbation $\eta_i \uparrow$		
Effect	Result	Channel
$E_i \uparrow, E_j \downarrow$	$\eta_i \uparrow$	*I-O demand linkage*
$P_i \downarrow, P_j \uparrow$	$\eta_i \uparrow$	*I-O marginal cost linkage*
$\frac{n_i}{N} \uparrow$	$\eta_i \downarrow$	*Market crowding*

$i = 1, 2$. The symmetry of the model implies that we may write the intersectoral labour flows in a single equation as follows:

$$\overset{\bullet}{\lambda}_{Mi} = w_{Mi}(\lambda_{Mi}) - w_{Mj}(\lambda_{Mi}), \tag{16.13}$$

where the notation $w_{Mi}(\lambda_{Mi})$ recalls that manufacturing wages depend on λ_{Mi}. We may now analyze the three mechanisms at work. Again, for clarity of exposition, we consider an exogenous perturbation to the symmetric geographic configuration, $\lambda_{Mi} = 1/2$, and study the causal chain it triggers. A change in λ_{Mi} causes an identical change in η_i and the latter sets in motion three mechanisms.

The first is a demand linkage and is conveyed by firms' expenditures on manufactures. We refer to it as the *I-O demand linkage*. An increase in η_i brings about an increase in total demand emanating from region i due to the increase in aggregate firm expenditure (see expression (16.12) and note that $n_i = \eta_i N$). A corresponding decline in aggregate expenditure takes place in region j. The net effect is an increase in demand for any variety produced in region i due to the home market dominance already discussed above and in Sect. 9.2.4. As a result, η_i increases further.

The second mechanism is a cost linkage and we refer to it as the *I-O marginal cost linkage*. An increase in η_i causes a reduction in total cost for firms located in region i, since the number of varieties on which they pay transport costs when purchasing M declines (P_i declines). Exactly the opposite happens in region j. This makes profitability higher in i and smaller in j and causes firms to enter i and exit from j, i.e., a further increase in η_i.

The third mechanism is the same *market crowding effect* studied in the core-periphery model, pushing firms towards the smaller market. The three mechanisms are summarized in Table 16.3.

16.4.2.3 Market Integration and Industrial Localization

The I-O Demand Linkage and the I-O Cost Linkage are agglomeration mechanisms while the Market Crowding effect is a dispersion mechanism. The balance between these mechanisms depends on trade costs. Similarly to the core-periphery model, for high trade costs the symmetric equilibrium is the only stable spatial configuration; for intermediate trade cost the symmetric equilibrium and the core-periphery configurations are stable; for low trade cost only the core-periphery configurations are stable. The phase diagram for this model is qualitatively identical to that in Fig. 16.1.

Table 16.4 Agglomeration and dispersion mechanisms in the model with diminishing returns to labour in A

Perturbation: $\eta_i \uparrow$		
Effect	Result	Channel
$E_i \uparrow, E_j \downarrow$	$\eta_i \uparrow$	*I-O demand linkage*
$P_i \downarrow, P_j \uparrow$	$\eta_i \uparrow$	*I-O marginal cost of linkage*
$\frac{n_i}{N} \uparrow$	$\eta_i \downarrow$	*Market crowding*
$w_i \uparrow, w_j \downarrow$	$\eta_i \downarrow$	*Labour cost linkage*

16.4.3 Diminishing Returns to Labour in the A Sector

We present here a simplified version of the model proposed by Puga (1999) and characterized by the presence of diminishing returns to labour in the agricultural industry. Recall that in the core-periphery model, the production of A requires only labour (farmers). In the present model it is assumed instead that the A sector uses land (T) and labour (L) as inputs. Factor T is used only in industry A. Industry M uses a composite input produced by use of labour and the aggregate of all varieties of M as in the I-O linkages model of Sect. 16.4.2. Since labour is perfectly mobile between sectors the wage is the same in both industries within a region and is determined by its marginal productivity in industry A. The production technology of A is such that labour marginal productivity is increasing in the land/labour ratio, T_i/L_{Ai}. This model structure generates a new dispersion mechanism channeled by the change in the marginal productivity of labour in A whenever firms move from one region to the other.[11]

To understand this mechanism, consider again a perturbation to the symmetric geographic configuration that exogenously increases the number of firms in region i. The firm moving from region j to i releases labour in j and demands labour in i. Absent migration, the new demand for labour in i must be satisfied by drawing labour from industry A. As a result the ratio T_i/L_{Ai} increases and so does w_i. Exactly the opposite happens in region j, where the labour released by the firm is employed in industry A, the ratio T_j/L_{Aj} decreases and so does w_j. The marginal cost of producing in region i increases therefore whenever a new firm enters region i and the opposite happens in region j. This *labour cost linkage* is a dispersion mechanism since the changes in marginal cost push firms to migrate in the direction opposite to that of the exogenous perturbation. This dispersion mechanism exists even for zero trade costs. The other three mechanisms in the model are the same as those already studied in Sect. 16.4.2. The four mechanisms in this model are summarized in Table 16.4.

[11] See Nocco (2005) for an extension of this model where there are endogenous differences in technology levels due to interregional knowledge spillovers.

16.4.3.1 Market Integration and Industrial Localization

When trade costs are very high, the symmetric configuration is the only stable configuration. For intermediate trade costs, partial or complete agglomeration are stable configurations. For low trade costs, the benefit from being located in a large market is outweighed by the cost of paying high wages. Thus, firms always prefer to move to the location with smaller number of firms when trade costs are low or zero. As a result, in the early stage of economic integration (i.e., from high to intermediate trade costs) agglomeration emerges but as economic integration progresses (low trade cost) the world economy returns to its dispersed initial geographic configuration.

16.4.4 Footloose Capital

We present here the model developed in Martin and Rogers (1995). This model is known as the footloose capital model, and its central assumption is that capital may migrate but profits are repatriated. In this way, the expenditure in each region is constant with respect to the migration of capital and, therefore, there is no circular causation between migration and size of the market. The assumption of profit repatriation eliminates any agglomeration force and makes the model static in nature. This notwithstanding, the footloose capital model has been developed and used in the context of the economic geography literature and we therefore review it here.

16.4.4.1 Description of the Model

The world is composed of two regions indexed by $i = 1, 2$. Two factors, capital (K) and labour (L), produce two goods.[12] Capital may migrate between regions, labour may not. Good A is produced in perfect competition with a constant returns to scale technology which requires one unit of labour input to produce one unit of output. Furthermore A is chosen as the numéraire good and its price is set to 1. Good A is traded freely between regions and, as long as both countries produce some A (as we assume to be the case), the price of A is the same between regions.[13] Given the technology in industry A wages equalize too: $w_1 = w_2 \equiv w$. Good M is differentiated and produced in monopolistic competition with an increasing returns to scale technology which requires one unit of capital as fixed input and c

[12] Capital exists in a fixed stock and cannot be accumulated. The number of varieties is constant too. See, e.g. Martin and Ottaviano (1999, 2001) and Baldwin, Martin, and Ottaviano (2001) for interesting developments which link the new economic geography to endogenous growth theory.

[13] The condition for both countries being producers of A in any geographic configuration is that industry M is small enough to fit in one country.

units of labour per unit of output. Let π_i^o be the price of capital and w the price of labour. Total cost is $\pi_i^o + wcq_i$ where q_i is total firm output. As usual, we assume the presence of iceberg trade costs in M by which only a fraction $\tau \in (0, 1)$ of each unit sent from region i arrives at region j. Profit maximizing prices depend only on marginal cost (and not on fixed cost). The domestic price is $p_{ii}^* = \mu cw$ and the foreign price $p_{ij}^* = (1/\tau)\, p_{ii}^* > p_{ii}^*$. The total profit is $\Pi = pq_i - \pi_i^o - wcq_i$. Unlike the monopolistic competition models examined earlier, the presence of positive profits cannot bring about any entry since the number of firms is determined by the stock of capital. Here, any positive profit is entirely absorbed by the price of capital; that is, $\pi_i^o = p_{ii}q_i - wcq_i$, hence the price of capital coincides with the operating profits of the firm (hence the superscript ^{o}).[14] Substituting $wc = p_{ii}/\mu$ into $p_{ii}q_i - wcq_i$ gives the price of capital as function of firm total sales, $p_{ii}q$:

$$\pi_i^o = \frac{\mu - 1}{\mu} p_{ii}q_i. \tag{16.14}$$

The structure of demand is the same as in the models studied above. Consumers derive utility from consumption of both goods and spend a fraction γ of their income on good M and the remaining fraction on good A. Let s_{Li} and s_{Ki} be, respectively, the share of world stocks of L existing in i, and the share of the world stock of K owned by residents of region i. Region i's income is:

$$E_i = s_{Li}\bar{L}w + s_{Ki}\bar{K}\pi_i \tag{16.15}$$

Capital may move across regions and K_i represents the quantity of capital present in region i. The reward to capital, π_i^o, is repatriated by the capital owners. Therefore income from capital in region i is $s_{Ki}\bar{K}\pi_i$ regardless of the localization of capital.

The sub-utility derived from consumption of M takes the S-D-S form examined in Eq. (9.2). The resulting equilibrium conditions in the goods market are the same as in Eqs. (16.4) and (16.5) with expenditure given by expression (16.15).

16.4.4.2 Equilibrium

Since it takes one unit of capital to set-up a firm, the number of firms in the world, N, equals the stock of capital $\bar{K}$ and region i's share in the total number of firms, n_i/N, equals the region's share in total capital located in it, $K_i/\overline{K}$. Labour is immobile between regions and each country is endowed with L_i units of it, the world stock of labour is constant at $\overline{L} = L_1 + L_2$. The migration of capital makes that operating profits equalize, that is,

$$\pi_1^o = \pi_2^o \tag{16.16}$$

[14] Perfect competition in the labor market makes it impossible for w to rise.

Capital migration is assumed to be instantaneous so that Eq. (16.16) holds at any time. Unlike the other models in this chapter, in the footloose capital model there is no dynamic adjustment but this is not important since this model has a unique stable equilibrium. The reason for this uniqueness and stability is, as anticipated above, that the expenditure emanating from a region is independent of the migration of capital (see Eq. 16.15).

The solution of the model yields the equilibrium profit, the size of the firm, and the distribution of capital between regions.[15] It is useful to show the solution for the share of firms (equal to the share of capital) in a region using the definition $\eta_i \equiv n_i/N$. This solution is:

$$\eta_i^* = \frac{1}{2} + \frac{1+\phi}{1-\phi}\left(\frac{E_i}{E} - \frac{1}{2}\right), \tag{16.17}$$

where ϕ is a parameter related to trade cost and ranging from 0 to 1. When $\phi = 0$, trade costs are prohibitive, when $\phi = 1$ trade costs are zero. It is immediate by inspection of expression (16.17) that the footloose capital model exhibits the home market effect already encountered in Sect. 9.2.4. In fact, since $\frac{1+\phi}{1-\phi} > 1$ whenever $0 < \phi < 1$, the larger region has a more than proportionally larger share of manufactures whenever trade is costly but not prohibitively costly. This model may be seen more as a variant of the monopolistic competition model of international trade studied in Sect. 9.2.1 than as a new economic geography model, but the assumption of migration makes it suitable to study issues related to market integration and location of industries. Furthermore, the model is particularly useful to highlight some issues related to welfare and for this reason we shall use it in Sect. 16.5 below.

16.5 Too Much or Too Little Agglomeration?

We have seen above that new economic geography models give rise to a rich set of stable geographic configurations. In this section we address the question of whether these configurations are socially optimal. There are probably as many, if not more, answers to this question as there are models. The reason is that the number, positions, and stability of the geographic configurations are often sensitive even to minor model modifications. Further, the various criteria that can be used to asses whether a geographic configuration is socially optimal often give contrasting results. We therefore do not go into a taxonomy of welfare analysis. We focus instead on a

[15] To solve the model note that aggregate operating profit in the world economy is $N\pi^o$ where the subscript i is suppressed since $\pi_1^o = \pi_2^o$. By virtue expression (16.14), $N\pi^o$ is equal to the fraction $(\mu-1)/\mu$ of world sales. World sales are equal to world expenditure which in turn is equal to world income. Furthermore, $N = \overline{K}$. Therefore we have the equation $\overline{K}\pi = \frac{\mu-1}{\mu}\gamma\left(L + \overline{K}\pi\right)$. From this equation we obtain π^* and all the other endogenous variables.

simple and insightful case where the geographic configuration determined by market forces is not socially optimal. To illustrate this case we follow Baldwin et al. (2003, sect. 11.2.4) and use the footloose capital model.

The socially optimal outcome is the one that maximizes social welfare defined here as the sum of purchasing power (indirect utility) across all individuals in the world. We assume that the "social planner", the imaginary figure who maximizes social welfare, does it by choosing η_i. The socially optimal value of η_i turns out to be

$$\eta_i^S = \frac{1}{2} + \frac{1+\phi}{1-\phi}\left(S_{Pop,i} - \frac{1}{2}\right) \tag{16.18}$$

where $S_{Pop,i} \equiv (K_i + L_i) / \left(\overline{K} + \overline{L}\right)$ is region i's share in the world population. We observe that the social planner would allocate the manufacturing industry between regions according to the regions' relative population. The larger the population in region i, the larger the share of manufacturing output allocated to that region. We also observe that the planner allocates the manufacturing activity to a region in a more than proportional relationship with the region's share of the population since $\frac{1+\phi}{1-\phi} > 1$. Baldwin et al. (2003) refer to this result as to the Social Home Market Effect paralleling the terminology used in Sect. 9.2.4.

To answer the question raised in the title of this section it suffices to compare η_i^* with η_i^S by use of expression (16.17) and (16.18):

$$\eta_i^* - \eta_i^S = \frac{1+\phi}{1-\phi}\left(\frac{E_i}{E} - S_{Pop,i}\right). \tag{16.19}$$

Expression (16.19) shows that, unless $\frac{E_i}{E} = S_{Pop,i}$, the market outcome is not optimal. There is either too much agglomeration (when $\frac{E_i}{E} > S_{Pop,i}$) or too little agglomeration (when $\frac{E_i}{E} < S_{Pop,i}$). Recalling the expressions for expenditure and population shares the condition for optimality is

$$\frac{L_i + K_i\pi^*}{\overline{L} + \pi^*\overline{K}} = \frac{L_i + K_i}{\overline{L} + \overline{K}}. \tag{16.20}$$

Therefore, except for the knife edge cases where $\pi^* = w$ or where $L_i/K_i = \overline{L}/\overline{K}$, the market outcome is not socially optimal. Further, the market allocates too many manufacturing firms to the larger region if and only if the larger region has a higher per capita income.

The analysis in this section is simple and insightful but, as recalled above, welfare results are sensitive to the model assumptions. For further welfare analysis using different models and different criteria see, e.g., Trionfetti (2001), Ottaviano, Tabuchi, and Thisse (2002), Baldwin et al. (2003), Ottaviano and van Ypersele (2005), Charlot, Gaigne, Robert-Nicoud, and Thisse (2006), and Ottaviano and Robert-Nicoud (2006).

Box 16.1 A Bird's-Eye View of Agglomeration

Satellite photographs of earth taken at night show the geographical distribution of artificial light. The presence of artificial light reveals the presence of human settlements which, in turn, implies the presence economic activity. Thus, the presence and density of artificial lights may be taken to reveal the presence and density of economic activity. Obviously, this is far from a precise way of measuring the agglomeration of economic activity, but has the advantage of revealing a lot of information at a glance.

The pictures speak clearly: even taking account of natural obstacles to human activity, such as deserts, ice, or high mountains, it is quite clear that human activity is unevenly distributed on the geographical space.

Europe and Africa

Asia and Oceania

North America

South America

Taking North America, for instance, we see that the lights are more densely present in the East than in the West, thus revealing a high concentration of economic activity in the East relative to the West. Taking smaller geographical units still reveals the presence of some agglomeration. For instance, within the East of North America human activity is concentrated in areas on the Southern shores of the Great Lakes and on the Boston-Washington strip.

The fact that economic activity is unevenly distributed does not tell us anything about the determinants of such an agglomeration. In this chapter we have studied the literature that highlights the role of endogenous agglomeration and dispersion mechanisms but there may be other and equally plausible explanations for the observed patterns of agglomeration. One such explanation is sheer chance. Think of throwing darts at a dartboard; the resulting distribution of darts will most likely exhibit some agglomeration pattern which would be the result of chance. Similarly, economic activity could settle randomly on the available land, yet such randomness could exhibit some agglomeration patterns (see Gabaix, 1999; Ellison & Glaeser, 1997). Another plausible determinant is represented by the presence of exogenous differences between location which make some of them objectively more attractive then others. Interestingly these differences in attractiveness could be such as to be relevant only at some point in history and yet such as to give rise to agglomeration patterns that persist throughout millennia. For instance, many major cities in the world today were founded near a river in ancient times. Proximity to a river was important then but its importance has faded away with time. Yet, most of these cities are at the heart of agglomerated areas still today (see Davis & Weinstein, 2002, 2008).

16.6 Conclusion

The NEG literature has evolved very rapidly but it may be argued that is still in search of a unified framework. The theoretical results obtained from the NEG are very sensitive to the models assumptions and even within a single model the sensitivity of the results to parameter values makes it difficult to draw general conclusions. This sensitivity becomes a true difficulty when it comes to drawing conclusions on welfare, or when prescribing policy recommendations, and when trying to assess the empirical validity of NEG models.

Policy research has so far addressed specific issues exploiting some, probably robust, features of NEG models. For instance, Martin and Rogers (1995) have studied the effect of infrastructure policy and Brülhart and Trionfetti (2004) have studied the effect of home-biased public procurement on international specialization and on agglomeration. Other papers have investigated some distinguishing features of NEG models that appear in tax competition. For instance, Baldwin and Krugman (2004) study the effect of agglomeration rents on tax competition. Agglomeration rents may be seen in Fig. 16.1, panels (b) and (c). The value of the phase line for $\lambda_i = 1$ (or $\lambda_i = 0$) is the real wage difference between regions. It is a rent for workers in the core in the sense that there are no endogenous economic mechanisms that erode it. This rent is taxable. The core region may tax local workers (or firms) as much as the rent without causing their departure. This possibility gives rise to a type of tax competition which would not emerge in the absence of agglomeration rents. Further works on tax competition and economic geography include Ludema and Wooton (2000), Kind, Midelfart-Knarvik, and Schjelderup (2000), Andersson and Forslid (2003), Brülhart and Jametti (2008), Brülhart and Parchet (2011), and

Brülhart, Jametti, and Schmidheiny (2012). Trionfetti (2012) studies instead the effect of public debt policies on economic geography and on tax competition.

The market structure typically used in new economic geography models is monopolistic competition. However, as noted by Neary (2001), this market structure limits the amplitude of the strategic behaviour of firms in relation to location decisions. Combes (1997) was the first to study this matter by replacing monopolistic competition with Cournot oligopoly. More recently, Combes and Lafourcade (2011) evaluate the role of competition and input–output market access in shaping the geography of economic activity. Annicchiarico, Orioli, and Trionfetti (2012) study instead the link between competition policy and market integration and their effect on firms location in the context of Cournot oligopoly.

Empirical investigation has started with some delay and still comprises only a few works. Crozet (2004) verifies the empirical validity of the cost of living linkage and estimates the parameters of the core-periphery model, finding strong support for it. One of the distinguishing features of NEG models is the presence of multiple equilibria. Thus, Davis and Weinstein (2002, 2008) search for empirical evidence of the existence multiple equilibria. The logic of their study may be understood with reference to Fig. 16.1, panel (b) where a large enough perturbation of the symmetric configuration will eventually move the economy towards a stable configuration different from the initial one. One of their findings is that even after a large and exogenous perturbation to the initial configuration (the allied bombing to Japanese cities) the economy returns to its initial geographic configuration. This means no evidence of the existence of multiple equilibria.

Redding and Sturm (2008) verify the empirical validity of the housing congestion model studied above. They study the distribution of economic activity in Western Germany before and after the sudden trade opening between East and West Germany occurred with the fall of the iron curtain. They find that after the trade opening economic activity in the former West Germany has relocated towards the former East/West border. This is coherent with the model. Brülhart, Carrère, and Trionfetti (2012) study the effect of the fall of the iron curtain on the geographical distribution of economic activity and on wage differences between Austrian regions. They find a good adherence of the model to the data especially when heterogenous preferences for locations are added to the housing congestion model. For a comprehensive appraisal of the empirical literature, see Redding (2010).

The new economic geography has come a long way since Krugman's seminal paper. Yet, as argued by Behrens and Robert-Nicoud (2010), further major breakthroughs will probably be achieved only by facing the hard questions. Two such questions concern iceberg costs and the use of numerical solutions. Iceberg costs are omnipresent and crucial but clearly a fiction. Numerical solutions are necessary in NEG models but given the large number of parameters they give rise to a large taxonomy of cases from which it is difficult to draw clear conclusions. Modeling the transport sector and moving from numerical simulation to sound calibration appear as promising lines of research.

References

Alonso-Villar, O. (2005). The effects of transport costs revisited.
Andersson, F., & Forslid, R. (2003). Tax competition and economic geography.
Annicchiarico, B., Orioli, F., & Trionfetti, F. (2012). National oligopoly and economic geography.
Baldwin, Richard E. (2001). The core-periphery model with forward-looking expectations.
Baldwin, Richard E., Forslid, R., Martin, P., Ottaviano, G. I. P., & Robert-Nicoud, F. (2003). *Economic geography and public policy*.
Baldwin, Richard E., & Krugman, P. R. (2004). Agglomeration, integration and tax harmonisation.
Baldwin, Richard E., Martin, P., & Ottaviano, G. I. P. (2001). Global economic divergence, trade and industrialisation: The geography of growth takeoffs.
Beckmann, M. J. (1987). Location of economic activity.
Behrens, K., & Robert-Nicoud, F. (2010). Tempora mutantur: In search of a new testament for neg.
Brülhart, M., Carrère, C., & Trionfetti, F. (2012). How wages and employment adjust to improved market access: Evidence from Austria.
Brülhart, M., & Jametti, M. (2008). Does tax competition tame the leviathan?
Brülhart, M., Jametti, M., & Schmidheiny, K. (2012). Do agglomeration economies reduce the sensitivity of firm location to tax differentials?
Brülhart, M., & Parchet, R. (2011). Alleged tax competition: The mysterious death of inheritance taxes in Switzerland.
Brülhart, M., & Trionfetti, F. (2004). Public expenditure, international specialisation and agglomeration.
Charlot, S., Gaigne, C., Robert-Nicoud, F., & Thisse, J.-F. (2006). Agglomeration and welfare: The core-periphery model in the light of bentham, kaldor, and rawls.
Combes, P.-P. (1997). Industrial agglomeration under cournot competition.
Combes, P.-P., & Lafourcade, M. (2011). Competition, market access and economic geography: Structural estimations and predictions for France.
Crozet, M. (2004). Do migrants follow market potentials? An estimation of a new economic geography model.
Davis, D. R., & Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity.
Davis, D. R., & Weinstein, D. E. (2008). A search for multiple equilibria in urban industrial structure.
Dreher, A. (2006). Does globalization affect growth? Evidence from a new index of globalization.
Dreher, A., Gaston, N., & Martens, P. (2008). *Measuring globalization – gauging its consequences*.
Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach.
Findlay, R. (1995). *Factor proportions, trade, and growth* (chap. 6).
Forslid, R., & Ottaviano, G. I. P. (2003). An analytically solvable core-periphery model.
Gabaix, X. (1999). Zipf's law for cities: An explanation.
Gandolfo, G. (2009). *Economic dynamics*.
Gupta, S. D. (Ed.). (1997). *The political economy of globalization*.
Helpman, E. (1998). The size of regions.
Kind, H. J., Midelfart-Knarvik, K. H., & Schjelderup, G. (2000). Competing for capital in a 'Lumpy' world.
Krugman, P. R. (1991a). *Geography and trade*.
Krugman, P. R. (1991b). History versus expectations.
Krugman, P. R. (1991c). Increasing returns and economic geography.
Krugman, P. R. (1993a). On the relationship between trade theory and location theory.
Krugman, P. R., & Venables, A. J. (1995). Globalization and the inequality of nations.
Lösch, A. (1954). *Economics of location* (English translation of 1940 German original).
Ludema, R. D., & Wooton, I. (2000). Economic geography and the fiscal effects of regional integration.

Martin, J. P., & Rogers, C. A. (1995). Industrial location and public infrastructure.
Martin, P., & Ottaviano, G. I. P. (1999). Growing locations: Industry location in a model of endogenous growth.
Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration.
Neary, J. P. (2001). Of hype and hyperbolas: Introducing the new economic geography.
Nocco, A. (2005). The rise and fall of regional inequalities with technological differences and knowledge spillovers.
Ohlin, B. (1933). *Interregional and international trade*.
Ohlin, B., Hesselborn, P., & Wijkman, P. M. (Eds.). (1977). *The international allocation of economic activity*.
Ottaviano, G. I. P. (1999). Integration, geography, and the burden of history.
Ottaviano, G. I. P. (2001). Monopolistic competition, trade, and endogenous spatial fluctuations.
Ottaviano, G. I. P., & Robert-Nicoud, F. L. (2006). The ‘Genome’ of NEG models with vertical linkages: A positive and normative synthesis.
Ottaviano, G. I. P., Tabuchi, T., & Thisse, J.-F. (2002). Agglomeration and trade revisited.
Ottaviano, G. I. P., & van Ypersele, T. (2005) Market size and tax competition.
Puga, D. (1999). The rise and fall of regional inequalities.
Redding, S. J. (2010). The empirics of new economic geography.
Redding, S. J., & Sturm, D. (2008). The costs of remoteness: Evidence from German division and reunification.
Ricci, L. A. (1999). Economic geography and comparative advantage: Agglomeration versus specialization.
Stern, R. M. (2009). *Globalization and international trade policies*, (chap. 2, sect. II).
Trionfetti, F. (2001). Public procurement, economic integration, and income inequality.
Trionfetti, F. (2012). Public debt and economic geography.

Chapter 17
Trade Integration and Wage Inequality

17.1 Introduction

A sharp increase in wage inequality has taken place from the 1990s. The inequality concerns especially the rise in the wage of skilled labour (college educated workers). Typically the wage of skilled labour is larger than the wage of unskilled labour, therefore, an increase in the relative wage of the former means an increase in wage inequality. This is why often in the literature the increase in the skill premium is synonymous with wage inequality. We adopt this terminology here whenever confusion does not arise.

A change in the skill premium may occur for various reasons, but the fact that this one occurred at a time of rising globalization makes international trade a prime suspect. In this chapter, we study the possible links between trade integration and rising skill premium. The matter is of obvious importance and has been the subject of a lively debate since the beginning of this century. The initial debate focused particularly on the appropriate use of the factor content of trade to extrapolate the effects of trade integration on the skill premium and the role of trade versus technical change as the cause of rising skill premia; see, e.g., Deardoff (2000), Krugman (2000), Leamer (2000), and Panagariya (2000). Later works explored the possibility that technical change induced by trade integration could be the cause of an increase in the skill premium: works by Acemoglu (2002, 2003), Burstein and Vogel (2012), and Costinot and Vogel (2010) belong to this group. Other works, e.g., Feenstra and Hanson (1997, 1999), Antràs and Helpman (2004), Antràs, Garicano, and Rossi-Hansberg (2006), Garicano and Rossi-Hansberg (2006) and Grossman and Rossi-Hansberg (2008) have identified in offshoring a plausible explanation for the rise in the skill premium. Other authors, such as Neary (2002), Epifani and Gancia (2008), and Dinopoulos et al. (2011) have highlighted the role of market structure and economies of scale. Krugman and Venables (1995) have proposed a model where agglomeration forces give rise to changes in the skill premium. Manasse and Turrini (2001), Yeaple (2005), Bustos (2011), Helpman, Itskhoki, and Redding

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_17,

(2010), Amiti and Davis (2011), and Unel (2010) are examples of works that have put forward the role of heterogeneity of firms and/or labour as a key element linking trade integration to the increase in the skill premium.[1]

Although this chapter focuses on the theoretical literature, it is important to mention that the empirical literature has also progressed very rapidly; for a review see, e.g., Goldberg and Pavcnik (2007) and Bernard, Jensen, Redding, and Schott (2007). Perhaps one of the most notable results is that the skill premium has increased in both skill-abundant and skill-scarce countries: a fact that runs against the convergence of relative factor prices predicted by the standard Heckscher-Ohlin model.

We begin our study precisely from the Heckscher-Ohlin model, which offers a very useful starting point to discuss the matter.

17.2 Comparative Advantage, Technical Change, and the Skill Premium

In this section factor names are unskilled labour (L) and skilled labour (S) instead of labour (L) and capital (K) as we have done in previous chapters. This is a mere relabelling which obviously leaves the Heckscher-Ohlin model unchanged. Consistently with this notation, we denote factor prices by w_L and w_S and use the convention that wage inequality is measured in terms of the relative wage of skilled labour, w_S/w_L. The phenomenon we want to explain is the rise in w_S/w_L.

17.2.1 Trade Integration

We recall that in autarky, the relative price of the skill-intensive good is lower in the skill-abundant country than in the skill-scarce country. This is indeed one way of stating that the skill-abundant country has a comparative advantage in the skill-intensive good. Trade integration brings about convergence of goods prices: the relative price of the skill-intensive good rises in the skill-abundant country and declines in the skill-scarce country. But, as we learnt in Sect. 4.1.1, there is a one-to-one correspondence between the relative price of goods and the relative

[1]The first group of studies is based on the Heckscher-Ohlin model studied in Chaps. 4 and 5. The second group of studies uses model structures related to the Krugman model studied in Sect. 9.2.1. The work by Krugman and Venables relates to models of economic geography studied in Sect. 16.4.2. The last group of studies uses models related to the work of Melitz studied in Sect. 9.2.7.

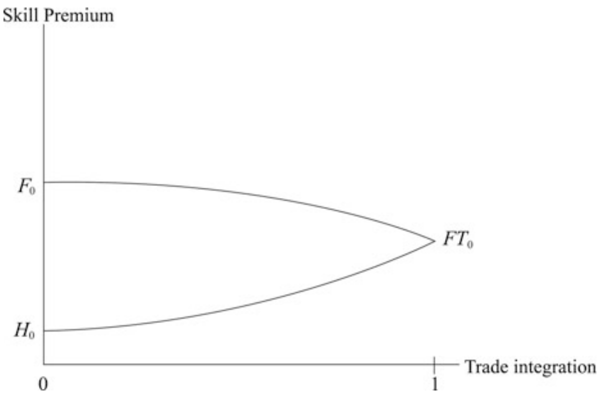


Fig. 17.1 Trade integration and the skill premium

price of factors.[2] Therefore, in the skill-abundant country, the increase in the relative price of the skill-intensive good will bring about an increase in the skill premium. Likewise, in the skill-scarce country, the decline in the relative price of the skill-intensive good will cause a fall of the skill premium. Ultimately, then, trade integration brings about an increase in the skill premium in the skill-abundant country and a decline of it in the other country. This is shown in Fig. 17.1.

Countries are labelled H and F and we use the convention that H is skill-abundant relative to F. Trade integration goes from zero (autarky) to one (free trade). The skill premium increases with trade integration in the skill-abundant country (from H_0 to FT_0) and decreases in the skill-scarce country (from F_0 to FT_0). In the literature, the evolution of the skill premium depicted in Fig. 17.1 is sometimes referred to as the Stolper-Samuelson effect[3] and we adopt this convention here.[4]

17.2.2 Technical Change

What type of technical change could explain the rise in the skill premium? In Sect. 13.5 we studied different types of technical change. With respect to the

---

[2]We assume here that the conditions that give rise to such a correspondence are satisfied.

[3]See, e.g., Epifani and Gancia (2008) and Costinot and Vogel (2010, p. 782).

[4]The Stolper-Samuelson theorem is a constitutive element of the Stolper-Samuelson effect but does not coincide with it. In fact, the Stolper-Samuelson theorem only establishes a relationship between price of goods and price of factors stemming from profit maximization, and is not necessarily related to any specific general equilibrium evolution of prices.

factor bias, we recall that technical change is Hicks-neutral when the marginal productivity of factors changes equiproportionally; conversely, technical change is biased toward a factor when the marginal productivity of that factor increases more than proportionally with respect to the marginal productivity of the other factor. With respect to the sector, we say that there is a sector-bias in the direction of the sector that benefits most from the technical change. Both the sector and the factor dimensions are potentially important in explaining the changes in the skill premium. Recall, for instance, the case of Hicks-neutral technical progress (Sect. 13.5.2) in the skill-intensive industry. At constant commodity-price ratio, this change induces an increase in the skill-premium. Conversely, a Hicks-neutral technical progress in the low-skill-intensive industry gives rise, at constant commodity-price ratio, to a decline in the skill-premium. So, at constant commodity-price ratio, the sector bias of technical change determines the consequences on the skill premium. The assumption of constant commodity-price ratio, however, is only tenable when referring to small countries. When we abandon the assumption of constant commodity-price ratio there are further consequences. Take the case of Sect. 13.5.2 again. We noted there that in general equilibrium a Hicks neutral technical progress in the skill-intensive sector brings about a decline in the relative price of the skill-intensive good which will reduce the skill premium via the Stolper-Samuelson theorem . Thus, the effects resulting from the sector bias should be weighed against the effects resulting from changes in the commodity-price ratio. This is not a marginal matter; under very common assumptions about the production and the utility functions the effect of the change in the commodity-price ratio completely neutralizes the effect due to the sector bias. More in general, the sector bias may be attenuated, neutralized or reversed by the change in the commodity-price ratio. Coming to the role of the factor bias, recall the analysis in Sect. 13.5.3 which showed that both skill-biased and low-skill biased technical change give rise, at constant commodity-price ratio, to an increase in the skill premium, as long as the technical change occurs in the skill-intensive industry. Symmetric results occur if the skill biased technical change occurs in the low-skill-intensive industry. To this, we have to add that the assumption of constant commodity-price ratio can only be applied to small countries.

The ambiguity of the results is removed if we assume a pervasive worldwide skill biased technical change. This situation is represented in Fig. 17.2 where the dashed lines represent the relationship between the skill premium and trade openness after the occurrence of skill-biased technical progress worldwide and in every sector.

17.2.3 Trade Integration and Skill-Biased Technical Change

The evolution of the skill premium when trade integration and skill biased technical change occur simultaneously is given by the sum of the distinct effects of trade integration and of skill biased technical change. This is shown in Fig. 17.2.

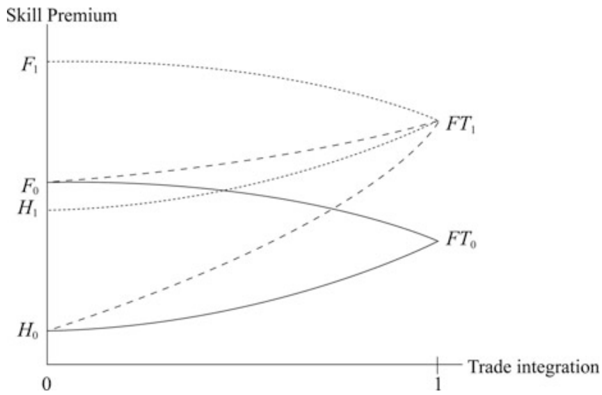


Fig. 17.2 Trade integration, technical change, and the skill premium

The skill premium in H will follow a path that goes from the autarky level H_0 to the free-trade level FT_1. The skill premium in the skill-scarce country goes from F_0 to FT_1. These paths are represented by the dotted lines. If the skill bias is sufficiently strong, as it is in the figure, point FT_1 has a larger value than F_0. In such a case, the skill premium increases in the skill-scarce country too. Otherwise the skill premium increases in H and declines in F but the decline is weaker than in the absence of skill-biased technical change.

So far, we have treated trade integration and technical change as independent causes of changes in the skill premium. We now discuss how the former may induce the latter.

17.2.4 Trade Integration Between Countries with Different Technology

The Heckscher-Ohlin model is usually treated under the assumption of identical technologies. This allows to focus on the role played by differences in factor endowments. But some studies (e.g., Acemoglu, 2003) argue that developed countries use more skill-intensive technologies than developing countries. We discuss the effect of trade integration between two countries with different technologies.

The world economy is composed of two countries, H and F, and produces two goods, Y_S and Y_L, by means of two primary factors, S and L. The countries have identical preferences over the two goods. We assume that country H uses a more skill-biased technology than country F. Specifically, though Y_S is more skill-intensive than Y_L in both countries, country H uses a more skill-intensive technology in all goods than country F. To isolate the effect of biased technology

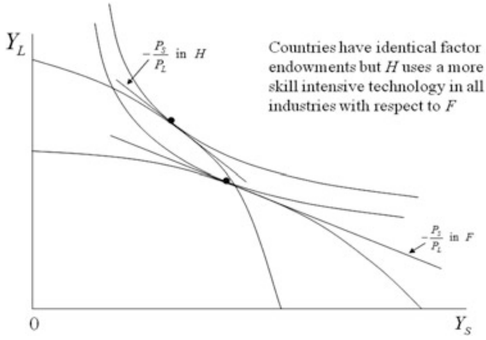


Fig. 17.3 Identical endowments and different technology

differences, we assume that countries have identical endowments. Therefore, absent the technical bias, there would be no comparative advantage. Consider the autarky situation. The relative demand for skilled labour, the skill-premium, and the relative price of the skill-intensive good are higher in H than in F since the technology is more skill-biased in H than in F. The autarky situation is represented in Fig. 17.3.

Though the two countries have identical endowments, the transformation curves are different because technologies are different.[5] The transformation curve of country H is steeper because skilled labour is absorbed by skill biased technology in H more than in F. Therefore, the intersection with the abscissa must lie closer to the origin for country H than for country F. The two curves convex to the origin represent the indifference curves. The autarky equilibria are represented by the tangency points of the indifference curve to the respective transformation curve. The relative price of the skill-intensive good is higher in H than in F as shown in the figure.

Trade integration brings about convergence of goods prices and with it, convergence of factor prices. This means that the skill premium declines in the skill-biased country and increases in the other country.[6] If we now assume that H is skill-abundant relative to F the overall effect of trade integration on the skill premium will depend on which of the two forces, technical bias or endowments, dominates. Costinot and Vogel (2010) argue that this ambiguity may explain why the overall effect of trade integration on factor allocation and factor prices tends to be small in practice.

[5]See Sect. 3.1 for the derivation of the transformation curve.

[6]Convergence of factor prices is not complete because technologies are different.

Box 17.1 Globalization and Factor Prices: A Historical Perspective

An important result of the Heckscher-Ohlin model is that trade integration, via the convergence of commodity prices, causes the convergence of absolute and relative factor prices. In this box we look at factor prices convergence during one of the most important globalization periods in history: the years between 1870 and 1939. There are two other important periods of globalization, the sixteenth century and the decades between the past and the current century. For the sixteenth century data is scant. For the most recent globalization period data abounds but we prefer to show evidence of factor prices convergence for the period 1870–1939 in honor of Eli Filip Heckscher (Stockholm, 1879–1952) and his disciple Bertil Gotthard Ohlin (Klippan, 1899—Valadalen, 1979). Both authors have lived through this period of globalization, maybe they have been inspired by it, surely they did not have the Heckscher-Ohlin theory as a reference to understand what was going on.

Before moving to the data it is worth reflecting on what we expect to find in the data. The model predicts convergence of absolute and relative factor prices. When trade is completely free the equalization of commoditiy prices brings about the equalization of absolute and relative factor prices. Thus, in free trade, $w_L^H = w_L^F$, $w_S^H = w_S^F$, and $w_S^H / w_L^H = w_S^F / w_L^F$. The first remark is that trade is never really completely free (at the very least there are trade costs to make it not totally free), so we should not expect to observe *equalization* of factor prices at the end of a globalization period. The second remark is that the model excludes any productivity differences between countries. This exclusion is made on purpose in order to focus on the role of relative factor endowments. We know that differences in technology between countries are reflected in absolute factor prices (recall Trefler's works studied in Sect. 4.6.4). Thus, we should not expect to find convergence of absolute factor prices.

Hicks-neutral productivity differences are inconsequential for relative factor prices convergence while factor-biased productivity differences may counter convergence but do not, per se, invalidate the convergence mechanism (see Sect. 17.2.4). Therefore, if we look for evidence of a robust prediction of the Heckscher-Ohlin model with respect to factor prices, we should look for evidence of relative factor prices convergence.

The period 1870–1939 was marked by the transport cost revolution, commodity and factor prices convergence. We report here some of the data and examples published in Williamson (2006, Tables 2.2, 4.1 and 4.2).

Fall in trade costs. Freight cost of American export routes fell by 45 % between 1970 and 1910. The fall in freight costs as a percentage of the rice price between Rangoon and Europe in the period between 1882 and 1914 is reported to be 75 %. Towards the end of the period the fall in transport costs slowed down but did not stop. For instance, ocean transport costs fell only by 32 % between 1920 and 1940. **Commodity price convergence.** Price convergence was remarkable. The Liverpool-Odessa percentage wheat-price gap fell by 95 % in the period 1870–1906. The London-Boston percentage wool-price gap fell by 52 % between 1870 and 1913. The London-Chicago percentage wheat-price gap fell by 72 % between 1870 and 1912. The Liverpool-Bombay percentage cotton-price gap fell by 65 % between 1873 and 1913. **Factor prices convergence.** According to the model we should observe an increase in the relative price of the relatively abundant factor. The factors taken into consideration for the period are land (whose price is the rental rate) and labour (whose price is the wage). The table reports data on the wage/rental ratio for a few countries only and for three points in time (see Williamson, 2006, for the complete set of data). It is clear from these simple statistics that factor prices have moved mostly in the direction predicted by the model. The exceptions concern the second part of the period for the Land-Abundant countries. But some of these countries (Australia, the United States, Uruguay) have received massive migration flows from Europe which have probably made them a lot less land-abundant by the end of the period.

Trends of the wage/rental ratio in Europe and the new World

	Land-abundant		Land-scarce		
Period	Australia	United States	Britain	France	Germany
1870–1874	416.2	233.6	56.6	63.5	84.4
1910–1914	100.6	101.1	102.7	99.8	100.2
1935–1939	110.5	240.1	206.5	168.2	n.a

Trends of the wage/rental ratio in the third World

	Land-abundant		Land-scarce		
	Uruguay	Siam	Egypt	Japan	Taiwan
1870–1874	1112.5	4699.1	174.3[a]	79.9[b]	
1910–1914	117.9	109.8	79.8	107.5	96.6
1935–1939	213.5	121.6	91.0	149.9	123.6

[a] Datum for 1875–1879
[b] Datum for 1885–1889

17.2.5 Trade-Induced Skill-Biased Technical Change

Acemoglu (2002) suggests that trade integration may induce skill-biased technical change. To understand the logic of his argument we present a simplified version of his model.[7]

There are two countries, H and F, endowed with constant quantities of primary factors, $\bar{S}^H$, $\bar{S}^F$, and $\bar{L}^H, \bar{L}^F$. Countries produce two goods, Y_L and Y_S, and machines. The two goods are assembled to yield a final good, Y, which is used for consumption and also as the only input in the production of machines. Goods are produced by use of machines and primary factors L (unskilled labour) and S (skilled labour). Specifically, Y_L is produced by use of L and a number N_L of different *L-complementary* machines while Y_S is produced by use of S and a number N_S of different *S-complementary* machines. Factor intensity in production is extreme since each primary factor is only used in the production of one good: Y_S is *S-intensive* and Y_L is *L-intensive*. Machines are *complementary* in the sense that they can only be used by the corresponding factor.

The production of Y_S and Y_L requires different inputs but the proportion of primary factors relative to machines for any given relative price of inputs is assumed to be the same for both goods. Furthermore, it is assumed that the marginal

[7] Acemoglu (2002) addresses a number of issues other than the effect of trade integration on factor prices. We restrict the discussion to the matter related to this chapter. Further we simplify the exposition by assuming that factors are gross substitutes. When they are gross complements trade integration does not necessarily result in skill-biased technical change.

productivity of each primary factor increases with the number of varieties of machines used by the factor. This is akin to the usual property that the marginal productivity of a factor increases as its relative use declines. Here, however, this property captures the idea that a larger number of more specific machines (tailor-made to the task) makes the primary factor more productive.

Consider the autarky equilibrium for country H (equilibrium for F is analogous) and drop the country superscript since confusion does not arise in autarky.

Primary factors are paid their marginal product. Therefore, the skill-premium, w_S/w_L, is equal to the ratio of marginal productivity, *MPS*/*MPL*, which in turn depends on the relative number of *S-complementary* machines, N_S/N_L, and on the relative supply of S, $\bar{S}/\bar{L}$. That is:

$$\frac{w_S}{w_L} = \frac{MPS}{MPL} = f\left(\underset{(+)}{\frac{N_S}{N_L}}, \underset{(-)}{\frac{\bar{S}}{\bar{L}}}\right), \tag{17.1}$$

where the notation $f\left(\frac{N_S}{N_L}, \frac{\bar{S}}{\bar{L}}\right)$ means that the skill premium depends on N_S/N_L and on $\bar{S}/\bar{L}$. The algebraic signs below these ratios indicate the direction of change of the skill premium resulting from an increase in the value of each of the ratios. An increase in N_S/N_L increases the skill-premium because it increases the relative marginal productivity of S. For the same reason, an increase in the relative supply of skilled labour, $\bar{S}/\bar{L}$, reduces the skill premium.

The markets for Y_L, Y_S, and Y are perfectly competitive. Good Y is the numéraire and p_L and p_S denote the price of Y_L and Y_S respectively. Each machine is produced by a monopolist. Machines are different but the technology of production is the same for all machines and requires a constant input of Y per unit of output.

We now discuss the link between relative goods prices, relative factor endowments, and technical change. First, the relative profit of *S-complementary* machines depends positively on the relative price of the skill-intensive good, p_S/p_L. The reason is that an increase in the relative price of this good induces an expansion of its relative production and an increase in the relative demand for *S-complementary* machines as inputs, thereby tending to increase the relative profitability of *S-complementary* machines. We refer to this mechanism as the *price effect*.

Second, the relative profitability of *S-complementary* machines increases with a rise in the relative endowment of skilled labour, $\bar{S}/\bar{L}$, precisely because of the complementarity in production. Indeed, if the input of S increases so does the input of *S-complementary* machines. We refer to this as the *market size* effect, where market means the market for machines.

Monopolists will produce *L-complementary* machines or *S-complementary* machines depending on which gives higher profits. Arbitrage between these two options assures that in equilibrium they give the same profit, i.e., $\pi_S/\pi_L = 1$.

This means that whenever there is pressure for an increase in π_S/π_L, this pressure induces an increase in the relative supply of *S-complementary* machines, restoring the equilibrium $\pi_S/\pi_L = 1$. The relative supply of *S-complementary* machines therefore depends on the relative price of goods via the price effect and on the relative number of factors via the market size effect. We write directly

$$\frac{N_S}{N_L} = f\left(\underbrace{\frac{p_S}{p_L}}_{\textit{Price effect}\,(+)}, \underbrace{\frac{\bar{S}}{\bar{L}}}_{\textit{Market size effect}\,(+)}\right), \tag{17.2}$$

where, again, the notation $f(.,.)$ means that N_S/N_L depends on the two variables in parentheses and the algebraic signs indicate the direction of the relationship.[8]

The relative price of the skill-intensive good depends negatively on its relative supply, but the latter depends positively on the relative availability of inputs (primary factors and machines). Therefore we can write

$$\frac{p_S}{p_L} = f\left(\underset{(-)}{\frac{N_S\bar{S}}{N_L\bar{L}}}\right). \tag{17.3}$$

Let us now consider free trade between H and F and resume the country superscript. Countries are identical except for factor endowments given by $\bar{S}^H/\bar{L}^H > \bar{S}^F/\bar{L}^F$. When moving from autarky to free trade—with technology unchanged—the relative price of the *S-intensive* good will increase in the *S-abundant* country. This effect may be seen in Eq. (17.3) where, in free trade, world relative endowment $\left(\bar{S}^H + \bar{S}^F\right)/\left(\bar{L}^H + \bar{L}^F\right)$ replaces the country relative endowment $\bar{S}/\bar{L}$ and the former is smaller than the latter since H is skill-abundant. So far, we have obtained the same result as in standard Heckscher-Ohlin, namely, that the relative price of the skill-intensive good increases in the skill-abundant country (and declines in the skill-scarce country) when passing from autarky to free trade, and that in free trade the relative price of goods depends on the world relative supply of factors. What is new here is that the rise in the relative price p_S/p_L increases the profitability of *S-complementary* machines so that machine producers find it optimal to increase the production of machines *complementary* to the relatively abundant factor as indicated by the price effect in expression (17.2). As a result, the relative marginal productivity of the relatively abundant factor increases and that of the relatively scarce factor declines. There is therefore a trade-induced skill-biased technical change in the skill-abundant country. The consequences on the skill premium can be immediately seen by inspection of Eq. (17.1). The skill premium increases in the *S-abundant*

[8] We should be using f, g, etc. for different functional forms but we neglect this matter to keep notation at a minimum.

country for two reasons. First, because of the usual Stolper-Samuelson effect . This is seen by replacing $\bar{S}/\bar{L}$ with $\left(\bar{S}^H + \bar{S}^L\right)/\left(\bar{L}^H + \bar{L}^L\right)$ in Eq. (17.1). Second, because of the increase in N_S/N_L induced by the price effect.

Naturally, the opposite occurs in the *L-abundant* country. Thus, the skill premium increases in the skill-abundant country and declines in the skill-scarce country.

In this model, the change in the skill premium is due to both international trade and technical change, but the latter is induced by the former.

17.2.6 A Generalization

The literature on wage inequality has gradually evolved from the two-factor model structure to a model structure with many (actually a continuum of) types of labour which differ in skill. Analogously, the traditional two-goods structure is replaced by a structure with any number of goods which differ by skill intensity. Furthermore, in these models the productivity of labour for any given skill depends on the matching between the type of labour and the type of good. These new modeling structures are better suited to the analysis of wage inequality. As an example of these structures we review a simplified version of the model in Costinot and Vogel (2010).

17.2.6.1 The Model

The world is composed of two countries, H and F. Workers differ in skills measured by s. The lowest and highest value of skills are $\underline{s}$ and $\bar{s}$, respectively. Let V_s^H and V_s^F be the endowment of workers of skill s in H and F, respectively. The definition of skill abundance in this environment where there are many different types of workers is as follows. Country H is skill-abundant if and only if

$$V_{s'}^H V_s^F \geq V_s^H V_{s'}^F \text{ for all } s' \geq s. \tag{17.4}$$

Note that this definition implies that $V_{s'}^H / V_{s'}^F \geq V_s^H / V_s^F$ if skill levels s' and s exist in both countries. This is the natural extension of the traditional definition of comparative advantage. If instead either s' or s or both do not exist in a country, then the definition implies that $\underline{s}^H \geq \underline{s}^F$ and $\bar{s}^H \geq \bar{s}^F$. In either case, country H is skill-abundant.

Figure 17.4a shows an example of the distribution of skills for two countries. Skill levels are plotted on the abscissa and the number of workers on the ordinate. Country F has workers with skill levels from one to seven. Country H has workers with skill levels from four to ten. Country F has two workers with skill level equal to one, country H has none, country F has three workers with skill level equal to two and country H has none, while country F has no workers with skill level equal to ten and country H has two, and so on. The mean skill is 7 in H and 4 in F. The shape of the distributions is symmetric around the mean; for each

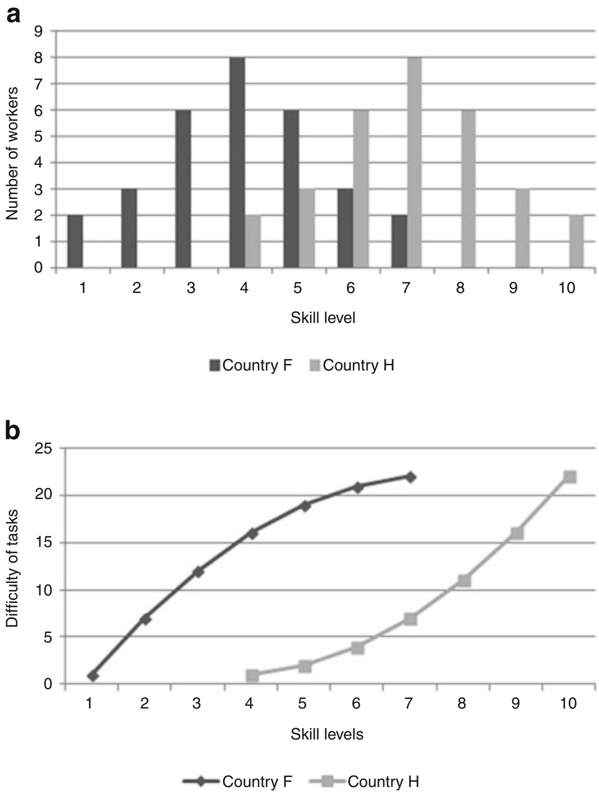


Fig. 17.4 Skills, tasks, and wages. (**a**) Skill distribution. (**b**) Wage distribution

distribution the endowment is the same for any skill level equally distant from the mean. Furthermore, the shape of the distribution is identical; each country has two workers with lowest skill-level, three workers with the second-to-lowest skill-level, etc. Naturally, neither symmetry nor identity of the distribution shape is a necessary requirement for the construction of the model, but we make these assumption here for simplicity. It is easily verified that H is skill-abundant in this example.

Total output of the economy, Y, is produced by inputting intermediate goods Y_δ, henceforth called tasks. Tasks are indexed by the difficulty of accomplishment, δ; the easiest task is $\underline{\delta}$ and the most difficult task is $\overline{\delta}$. The production function is

$$Y = \left(\sum_{\delta=\underline{\delta}}^{\overline{\delta}} (Y_\delta)^\alpha \right)^{1/\alpha}, \qquad 0 < \alpha < 1. \tag{17.5}$$

Expression (17.5) simply says that the output Y is obtained from the sum powered to $1/\alpha$ of inputs Y_δ powered to α. Tasks are produced with the following production function

$$Y_\delta = \sum_{s=\underline{s}}^{\bar{s}} a_s^\delta L_s^\delta, \qquad L_s^\delta \geqslant 0, \tag{17.6}$$

where a_s^δ is the productivity of a worker with skill s applied to task δ and L_s^δ is the endogenous labour input of skill s employed in the production of task δ. One key assumption of the model is that high skill workers have the comparative advantage in difficult tasks. This means to assume that for any $s > s'$ and any $\delta' > \delta$

$$a_{s'}^{\delta'}/a_{s'}^{\delta} > a_s^{\delta'}/a_s^{\delta}. \tag{17.7}$$

In other words, inequality (17.7) says that workers with higher skills are relatively more productive in accomplishing difficult tasks.[9]

17.2.6.2 Equilibrium

Goods and factors markets are perfectly competitive. Let p_δ denote the price of task δ, let Y serve as numéraire and let its price be set to 1. Then, profits in the final goods market are given by

$$\Pi = \underbrace{\left(\sum_{\delta=\underline{\delta}}^{\bar{\delta}} (Y_\delta)^\alpha\right)^{1/\alpha}}_{Revenues} - \underbrace{\sum_{\delta=\underline{\delta}}^{\bar{\delta}} p_\delta Y_\delta}_{Total\ costs}. \tag{17.8}$$

The final goods producers maximize profits by choosing the quantity of each intermediate input (task), Y_δ. This require the marginal revenue to equal marginal cost, which yields the following demand for each intermediate input:

$$Y_\delta = (p_\delta)^{-\alpha/(1-\alpha)}\, Y \qquad \forall \delta. \tag{17.9}$$

Intermediate goods producers maximize profits by choosing the labour input for each skill level, $L_s^\delta \geqslant 0$. In deciding the labour input, the firm compares the marginal cost of a worker, i.e., the wage w_s, with the marginal revenue produced by a worker, i.e., the value of that worker's production, $p_\delta a_s^\delta$. Employment of skill level s will be zero ($L_s^\delta = 0$) for all the s such that $p_\delta a_s^\delta < w_s$. Conversely, employment will be positive for all s such $p_\delta a_s^\delta = w_s$. Perfect competition rules out $p_\delta a_s^\delta > w_s$ in equilibrium. Thus, the conditions for profit maximization in intermediate production are

[9] This is reminiscent of Ricardian comparative advantage but here the comparative advantage is defined over worker-task pairs instead of country-good pairs.

$$p_\delta a_s^\delta - w_s \leqslant 0, \qquad \forall s \tag{17.10}$$

$$p_\delta a_s^\delta - w_s = 0, \qquad \forall s \text{ for which } L_s^\delta > 0 \tag{17.11}$$

Note that Eqs. (17.10) and (17.11), do not contain L_s^δ but employment must be compatible with them. Given the comparative advantage of workers established in (17.7), workers with the same skill level will never be employed in different tasks, nor will two workers with different skills be assigned to the production of the same task. Thus, if $L_s^\delta > 0$ then $L_s^{\delta'} = 0$ for all $\delta' \neq \delta$ and $L_{s'}^\delta = 0$ for all $s' \neq s$.[10]

Equilibrium in the market for tasks and for labour requires, respectively:

$$\underbrace{(p_\delta)^{-\alpha/(1-\alpha)}\, Y}_{Demand\ for\ task\ \delta} = \underbrace{\sum_{s=\underline{s}}^{\bar{s}} a_s^\delta L_s^\delta}_{Supply\ of\ task\ \delta} \qquad \forall \delta \tag{17.12}$$

$$V_s = \sum_{s=\underline{s}}^{\bar{s}} L_s^\delta, \qquad \forall s \tag{17.13}$$

where we recall that Y is the value of aggregate output and therefore the income of the economy.

To better understand the properties of the equilibrium, consider first a simple two-by-two simplification where $\delta = 1, 2$ and $s = 1, 2$. Then the eight equations (17.10), (17.11), (17.12), and (17.13) determine the equilibrium value of the eight endogenous variables $p_\delta, w_s, L_s^\delta$, from which all the other values of the endogenous variables are obtained. As an example let us assume an arbitrary employment allocation, for instance: $L_2^2 > 0, L_1^1 > 0$. Then, $L_1^2 = 0, L_2^1 = 0$, therefore $L_2^2 = V_2$ and $L_1^1 = V_1$; which implies $p_1 = w_1/a_1^1$ and $p_2 = w_2/a_2^2$ from (17.11). Inserting these prices into Eqs. (17.12) and (17.13); noting that Y, being the value of total output, is also national income, i.e., $Y = w_1 V_1 + w_2 V_2$, we obtain

$$\left(\frac{w_1}{a_1^1}\right)^{\alpha/(1-\alpha)} = \frac{w_1 V_1 + w_2 V_2}{a_1^1 V_1} \tag{17.14}$$

$$\left(\frac{w_2}{a_2^2}\right)^{\alpha/(1-\alpha)} = \frac{w_1 V_1 + w_2 V_2}{a_2^2 V_2}. \tag{17.15}$$

Equations (17.12) and (17.13) may be solved (numerically) for $\{w_1, w_2\}$. For this solution to be an equilibrium for the economy it must also satisfy inequality (17.10), otherwise the arbitrarily chosen employment allocation is not viable. It turns out that only the allocation $L_2^2 > 0, L_1^1 > 0$ satisfies all the equations, including inequality (17.10). Note that this employment allocation is such that the high-skill

[10] This is reminiscent of the traditional Ricardian model where specialization is complete.

worker is employed in the production of the difficult task and the low-skill worker is employed in the easy task. This property can be extended to the many-skills many-tasks context: In equilibrium, for any two workers s and s' and for any two tasks δ and δ', with $s' > s$ and $\delta' > \delta$, worker s' matches with task δ' and worker s matches with task δ.

Figure 17.4b depicts the equilibrium matching between skills and tasks for each country in autarky. Skill levels are plotted on the abscissa and the difficulty of tasks ranging from 0 to 22 on the ordinate. The matching schedule M_s shows the association of any skill level s to a task δ resulting from the competitive equilibrium. The matching schedule also reflects the wage schedule, since the wage increases with productivity and productivity increases with the skill level.

Figure 17.4 shows that in both countries, more highly-skilled workers match with more difficult tasks, but any given task is performed by more highly-skilled workers in H than in F. One way of interpreting this result is that the same task is produced with a more skill-intensive technique in H than in F. This is reminiscent of the simple two-by-two Heckscher-Ohlin model, where in autarky the S-abundant country produces both goods with more S-intensive techniques than the L-abundant country. The intuition here is practically the same: each country must make more intense use of the relatively abundant factor and produce a larger quantity of the good that is intensive in the relatively abundant factor in order to satisfy full employment conditions. Perfect competition assures that this is indeed the equilibrium outcome.

Now let us consider trade integration in tasks. Trade integration results in an increase in the skill premium in H and a decrease in F. This means that for any s and s', the skill premium $w_{s'}/w_s$ increases with trade integration in H and decreases in F. This is similar to the traditional Stolper-Samuelson effect and occurs for the same reason. Furthermore, trade integration causes skill downgrading for all tasks in H and skill upgrading for all tasks in F. Skill downgrading means that each task is performed by workers with lower skills while skill upgrading means that each task is performed by workers with higher skills. Again, this is similar to the two-by-two Heckscher-Ohlin model where the rise in the relative price of the relatively abundant factor brings about the use of techniques that are less intensive in the relatively abundant factor in every country. In terms of Fig. 17.4b, trade integration is represented by a downward shift and a clockwise rotation of the matching schedule for F, and an upward shift and a counterclockwise rotation for H.

The effects of trade integration in this model are then essentially the same as those studied in the Heckscher-Ohlin model, but here we have a more general modelling structure which allows to better understand the literature that we discuss in Sect. 17.5.

17.3 Offshoring and the Skill Premium

We have studied a model of offshoring in Sect. 6.8.5. We return to this matter here, where we focus on the effect of offshoring on the skill premium. Offshoring means relocating part of the production process abroad. This definition encompasses both multinationals (where the firm keeps ownership of the offshored activities) and foreign outsourcing (where the firm relinquishes ownership of the offshored activities). The offshored part of the production process may be material (fragmentation of the production process into different stages performed in different countries) or immaterial, typically containing knowledge or information, (assignment of different immaterial tasks to workers in different countries). In this section we describe two models of offshoring. The first is more closely related to the fragmentation of production, the second to the offshoring of knowledge output.

17.3.1 Fragmentation of Production

We follow closely Feenstra and Hanson (1997). They propose a simple and insightful model of offshoring in the context of the Heckscher-Ohlin framework. They consider a perfectly competitive industry where production requires many "activities", denoted z, with $z \in [0, 1]$, each of which performed by combining fixed quantities of skilled and unskilled labour. These quantities differ across activities. To simplify matters, let a_{Lz} and a_{Sz} denote the quantity of unskilled and skilled labour requested to produce activity z and let A_z denote the skill intensity of activity z, $A_z \equiv a_{Sz}/a_{Lz}$. We rank activities by increasing order of skill intensity so that A_z increases with z. Activities may be performed at home (country H for instance) or abroad (country F) and it is assumed that firms carry with them the home technology when delocalizing some activities abroad. Therefore, the input requirements per unit of output are the same regardless of whether an activity is performed at home or abroad. However, performing activities abroad increases the marginal cost by a factor $\chi > 1$. This is due, for instance, to the coordination and communication costs of organizing offshored activities. For firms in H, the marginal cost of performing an activity z in countries H and F is, respectively, $mc_z^H = w_L^H a_{Lz} + w_S^H a_{Sz}$ and $mc_z^F = \chi \left(w_L^F a_{Lz} + w_S^F a_{Sz} \right)$, and the ratio of marginal cost mc_z^H / mc_z^F is:

$$\frac{mc_z^H}{mc_z^F} = \frac{w_L^H + w_S^H A_z}{\chi \left(w_L^F + w_S^F A_z \right)} \tag{17.16}$$

Input requirements do not have a country superscript because the firm brings its technology when producing abroad. Firms in H perform activity z at home if $mc_z^F > mc_z^H$ and abroad if $mc_z^F < mc_z^H$. Solving

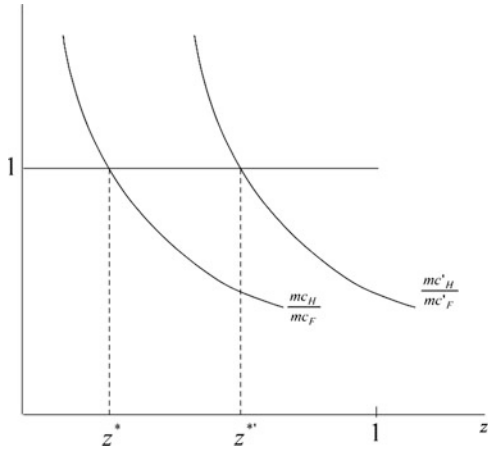


Fig. 17.5 Globalization and the partition of activities

$$\frac{mc_z^H}{mc_z^H} = 1 \tag{17.17}$$

for z gives the threshold activity z^* for which the firm is indifferent between performing it at home or abroad. If we assume that country H is skill-abundant, then $w_S^H/w_L^H < w_S^F/w_L^F$ and mc_z^H/mc_z^F is decreasing in z.

Figure 17.5 shows the determination of the equilibrium obtained by use of Eqs. (17.16) and (17.17). The intersection between the declining marginal cost ratio and the horizontal line drawn at value equal to one gives the threshold value z^*. Activities from 0 to z^* are performed in F whereas activities from z^* to 1 are performed in H. Thus, H firms perform high skill-intensive activities at home and offshore to the skill-scarce country the low skill-intensive activities. This is reminiscent of the specialization pattern occurring in the Heckscher-Ohlin model and indeed it occurs for essentially the same reasons.

Now imagine that the cost of offshoring (χ) declines. Then, as we see from Eq. (17.16), for any given z the relative marginal cost of producing at home increases. Graphically, this is represented by a movement to the right of the mc_z^H/mc_z^F curve to $mc_z^{\prime H}/mc_z^{\prime F}$ in Fig. 17.5. The threshold value of marginal cost moves to the right from z^* to $z^{*\prime}$. Activities from z^* to $z^{*\prime}$ that were produced in H are offshored to F after the decline in χ. What is the effect on the skill premium? The relative demand for skills increases at home since the newly-offshored activities are the least skill-intensive among those previously performed at home. The average skill intensity of the activities remaining at home increases and so does the relative demand for skilled labour. Therefore the skill premium increases in H. The newly-offshored activities are more skill-intensive than the activity performed in F previous to the fall in χ. The arrival of these new activities in F reduces the average skill intensity of the activities performed there and increases the relative demand for skilled labour. Therefore, the skill premium increases in F too. This is a simple but insightful story where globalization makes it possible to increase the share of offshored activities, thereby increasing the relative demand for skilled labour and the skill premium in all countries.

17.3.2 Offshoring Knowledge Output

We review the key elements of immaterial offshoring models by studying a simplified version of the model in Antràs et al. (2006). The world economy is composed by two countries, H and F, populated by agents of different skill levels, denoted z, ranging from 0 to z_H in H and from 0 to z_F in F.

In Fig. 17.6. The skill level, z, is plotted on the abscissa. Without loss of generality, it is assumed that $z_H = 1 > z_F$. The number of agents having any given value of z is measured on the left-hand ordinate. The value of the total population in each country is normalized to 1. Agents decide whether to become managers or workers. The production process requires the solution of problems of difficulty z ranging from 0 to 1 (denoting the skill level and difficulty level by the same variable simplifies notation). An agent of skill level z', whether manager or worker, is able to solve all the problems from 0 to z'. Production of the final good requires knowledge provided by a team composed of one manager and a number of workers. The team works as follows. Each worker produces knowledge output by solving problems for which he knows the solution. If a worker faces a problem he cannot solve he asks the manager. If the manager knows the solution he tells it to the worker, who produces one unit of knowledge output. If the manager does not know the solution, knowledge output is not produced. Agents are endowed with one unit of time and a manager spends $0 < h < 1$ units of time communicating with a worker, regardless of whether the manager knows the solution to the problem. A manager in a team with n workers of skill level z_p is asked for solutions for every problem of difficulty larger than z_p, and there are $\left(1 - z_p\right)$ such problems, each requiring h units of manager communication time. With n workers in a team, the manager input time is $h\left(1 - z_p\right)n$ and the manager time constraint is $h\left(1 - z_p\right)n = 1$. Therefore, a manager can supervise at most

$$n = \frac{1}{h\left(1 - z_p\right)} \tag{17.18}$$

workers of skill level z_p. The team output of a manager of skill level z_m working in a team with n workers of skill level z_p is $y_{mp} = z_m n$, or

$$y_{mp} = \frac{z_m}{h\left(1 - z_p\right)}. \tag{17.19}$$

Let w_{z_p} denote the equilibrium wage of workers with skill level z. Manager income R_{z_m} is given by the value of output minus costs, i.e., $y_{mp} - w_{z_p} n$. Thus, a manager chooses the skill level of his workers, z_p, so as to maximize R_{z_m} subject to the

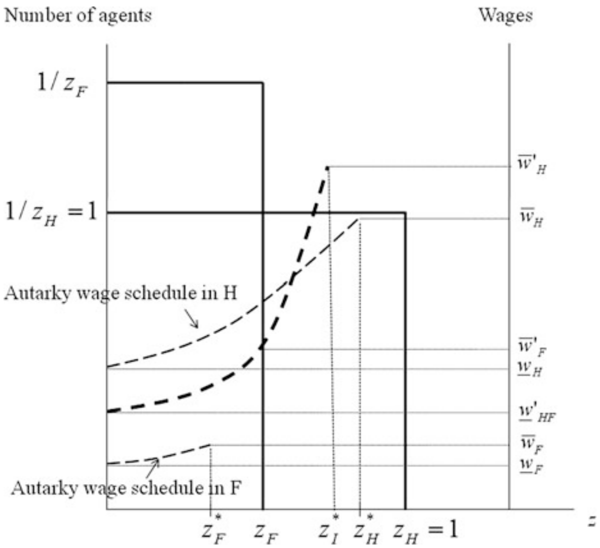


Fig. 17.6 Immaterial offshoring

time constraint $h\left(1 - z_p\right) n = 1$. Any agent of skill level z' chooses to become manager or worker by comparing the income he can get as a manager, $R_{z_m=z'}$ with the wage he can earn as a worker $w_{z_p=z'}$. This mechanism of occupational choice and the complementarity between workers and managers lead to a result in equilibrium that is pretty intuitive. The complementarity between workers and managers can be seen in the expression (17.19). It is clear from this expression that the higher the skill level of workers, z_p, the stronger the impact on team output of an increase in manager skills, z_m. Likewise, an increase in the skill level of workers is worth more to more highly-skilled managers. As a result, although $z_m > z_p$ in equilibrium, better workers will always belong to teams led by better managers.

Let us now characterize the autarky equilibria. Consider H first. The autarky equilibrium is characterized by a threshold value of skill, z^*_H, such that any agent with skill level $z = z^*_H$ is indifferent between being a manager or a worker. Thus, all agents with skill level $0 < z < z^*_H$ decide to become workers and all agents with skill level $z^*_H < z < 1$ decide to become managers. Quite naturally, the more highly-skilled agents find it optimal to become managers while the less-skilled agents find it optimal to become workers. The closed economy equilibrium in F is also characterized by a threshold value, z^*_F, but the threshold values are not the same; specifically, $z^*_H > z^*_F$. An agent with skill level $z = z^*_F$ in F is by definition indifferent between being manager or worker, but an agent in H with that same skill level is not indifferent. The reason is that he can join a team led by a manager whose skill level is higher than that of any manager in F, and there will be a manager in H with skill level higher than z^*_F willing to take him in his team. Thus, the agent in H with skill $z = z^*_F$ certainly decides to become a worker. In sum, the availability of higher skill levels in H means that the least-skilled manager in this country has higher skills than the least-skilled manager in F. This also implies that for any two workers with the same skill levels, the worker in H is led by a better manager. The thin dashed lines in Fig. 17.6 represent the wage schedules in autarky. Observe that wage inequality is larger in H than in F. Specifically, let the difference between the wage of the most highly-skilled and least-skilled workers be the measure of the skill premium (or of wage inequality). These differences are $\bar{w}_H - \underline{w}_H$ and $\bar{w}_F - \underline{w}_F$ and we see clearly that $\bar{w}_H - \underline{w}_H > \bar{w}_F - \underline{w}_F$. This is due to two reasons. The first is the occupational choice effect: since agents with higher skills decide to become workers in H ($z_H^* > z_F^*$) and since the wage is increasing in the skill level, then the skill premium will obviously be larger in H. The second reason is the complementarity effect: since any worker is led by a better manager in H than in F, any increase in z_p is worth more in H than in F and therefore gives rise to a stronger wage increase in H than in F. In other words, the slope of the wage schedule in H is larger than in F.[11]

Let us now consider globalization, in the sense that teams may be formed by agents of both countries. The threshold value in the globalized economy, z_I^*, will be between the two autarky threshold values, $z_F^* < z_I^* < z_H^*$, which means that the number of workers increases in F and the number of managers increases in H.[12] Specifically, all agents in F and agents with skills between 0 and z_I^* in H are workers while the remaining agents are managers. These changes in threshold values occur because the world economy has a distribution of skills that is between that of each of the countries. Take, for instance, the agent with skill $z = z_F^*$ in F; he is indifferent between becoming a manager or a worker in autarky but after globalization he can be led by a manager whose skill-level was not available in autarky and he will therefore no longer be indifferent: he will want to become a worker. Likewise, the agent with skill $z = z_H^*$ in H is indifferent between occupations in autarky (as a manager he would lead workers with skills $z = 0$) but after globalization he is in demand as a manager because there are many more low-skill workers in search for managers; he will decide to become a manager and will lead workers with skills $z > 0$.

These changes in the threshold values have repercussions on wage inequality, driven by the complementarity and occupational choice effects, as we now discuss. After globalization, all workers in F match with better managers who are located

[11]Incidentally, we also note that the wage schedule in H lies above that in F. This is again due to the fact that any worker in H is led by a better manager than any worker with the same skills in F. The convexity of the wage schedules is due to the complementarity between workers and managers established by expression (17.19).

[12]For clarity of exposition we have shown the case where z_I^* lies to the right of z_F, implying that all agents in F decide to become workers after globalization. Antràs et al. (2006) call this the "Low Quality Offshoring Equilibrium". The threshold value z_I^* could however fall to the left of z_F, in which case some agents in F will remain managers after globalisation. The case shown in the figure has simpler implications. The alternative case is more complex but still gives rise to an increase in the skill premium in F, while giving rise to ambiguous effects in H.

in H (offshoring of problem-solving). This is good for all $F-$workers, but the strength of the positive effect is proportional to each worker's skill; this is the complementarity effect. Furthermore, in F the number of workers increases from z_F^* to z_I^*, which pushes the skill premium upward; this is the occupational choice effect. Thus, wage inequality unambiguously increases in F. In H, among the workers that remain workers after globalization, the lower-skilled are matched with worse managers while the more highly-skilled are matched with better managers. Given the complementarity between managers and workers, this matching effect clearly tends to increase the skill premium among $H-$ workers. To counter this effect, there is the occupational choice effect represented by the decline in the number of workers in H. The effect is a priori ambiguous, but the skill premium unambiguously rises when communication costs (h) are low and when the endowment difference ($z_H - z_F$) is high. Thus, offshoring between very different countries and when the communication costs are low increases the skill premium in all countries. This is the situation represented by the thick dashed line in Fig. 17.6. This line represents the wage schedule after globalization (the same in both countries). By comparing $\bar{w}'_H - \underline{w}'_{HF}$ with $\bar{w}_H - \underline{w}_H$ and $\bar{w}'_F - \underline{w}'_{HF}$ with $\bar{w}_F - \underline{w}_F$ we see that globalization brings about an increase in the skill premium in both countries. There are also consequences at the level of nominal wages. Low-skill workers in F are matched with better managers after globalization while the opposite happens to low-skill workers in H. This explains why the thick dashed line intersects the left-hand ordinate in between the autarky wage schedules. Furthermore, the most highly-skilled workers among those who remain workers in H after globalization are matched with better managers than before, which explains the crossing of the thick dashed line with the H autarky wage schedule.

17.4 Economies of Scale and the Skill Premium

Some studies suggest that economies of scale at firm or at industry level may provide the link between trade integration and the skill premium. The types of economies of scale taken into consideration vary across authors, but they have in common that trade integration, via economies of scale, makes production more efficient. Crucially, the gain in efficiency is biased in favour of skilled labour.

We review two representative models. In the first, economies of scale come from the increased number of inputs made possible by trade liberalization. In the second, economies of scale come from the expansion of output caused by international trade and declining average costs.

17.4.1 Intermediate Inputs and Productivity

In this section we describe a simplified version of the model proposed in Epifani and Gancia (2008). Consumers derive utility from the consumption of two goods,

Y_L and Y_S. The utility function defined over these two goods takes the form already encountered in Eq. 9.2, which yields the following relative demand for Y_S:

$$\frac{Y_S^d}{Y_L^d} = \left(\frac{P_L}{P_S}\right)^{\varepsilon}, \quad \varepsilon > 1 \tag{17.20}$$

where the superscript d indicates that these are quantities demanded. The parameter ε represents the elasticity of substitution between goods. These goods are produced in perfect competition by assembling n_i intermediate inputs specific to good i $(i = L, S)$. Production of any variety of intermediate input, y_i, requires a fixed amount of labour and a constant marginal labour input. It is assumed that the production of intermediate inputs for the S industry requires only skilled labour while the production of the intermediate inputs for the L industry requires only unskilled labour. This is an extreme form of factor intensity which simplifies the exposition and has no crucial consequences on the result. Each intermediate input y_i is produced by a different firm in a monopolistic competitive market of the type studied in Sect. 9.2. The profit-maximizing price (p_i) is obtained from the condition marginal revenue = marginal cost. This condition gives $p_i = \mu w_i$, where $\mu > 0$ is the constant mark-up. The relative price of any two inputs in different industries is therefore:

$$\frac{p_S}{p_L} = \frac{w_S}{w_L}. \tag{17.21}$$

Expression (17.21) gives the Stolper-Samuelson relationship between the relative price of goods and the relative price of factors. We have already studied this relationship in Sect. 5.3 in the context of the Heckscher-Ohlin model. Here, the relationship is particularly simple because we have assumed that each input is produced with only one factor. Free entry assures zero profits and the zero profit condition determines the equilibrium size of firm output $(\bar{y})$ which turns out to be equal for all firms. Using this result, the production function of final goods may be written as:

$$Y_i = \bar{y}\,(n_i)^{\frac{\sigma_i}{\sigma_i - 1}} \tag{17.22}$$

Expression (17.22) shows the nature of economies of scale in the production of goods. As the number of intermediate inputs increases, final output Y_i expands, ceteris paribus. This property of the production function, introduced into international trade theory by Ethier (1982), captures the idea that a larger number of intermediate inputs increases the efficiency of the production process, for instance because a larger number of inputs better matches the specific needs of each production process. This property can be seen by dividing output (Y_i) by aggregate inputs (given by $\bar{y}n_i$). The result of this division is the marginal (and average) productivity of each input:

$$\frac{Y_i}{\bar{y}n_i} = (n_i)^{\frac{1}{\sigma_i - 1}} \tag{17.23}$$

which shows that the marginal productivity is increasing in the number of inputs. The relationship between number of inputs and output is not the same in both industries, however. In fact, the model assumes that $\sigma_L > \sigma_S > \varepsilon$. This means that economies of scale are stronger in the skill-intensive industry ($\sigma_L > \sigma_S$) and that the elasticity of substitution between any two inputs is larger than the elasticity of substitution between the final goods ($\sigma_L > \varepsilon$; $\sigma_S > \varepsilon$). Using (17.22) the relative supply of Y_S is:

$$\frac{Y_S^s}{Y_L^s} = \frac{(n_S)^{\frac{\sigma_S}{(\sigma_S - 1)}}}{(n_L)^{\frac{\sigma_L}{(\sigma_L - 1)}}}. \tag{17.24}$$

where the superscript s indicates quantities supplied. Profit maximization in the final good industry requires the marginal revenue to equal the marginal cost for each intermediate input, which is:

$$P_i \, (n_i)^{\frac{1}{\sigma_i - 1}} = p_i \tag{17.25}$$

Equation (17.25) is easily understood by noting that $(n_i)^{\frac{1}{\sigma_i - 1}}$ is the contribution to output of each input (its marginal productivity) and each unit of such contribution is worth P_i; therefore the left-hand side is the marginal revenue while the right-hand side is the marginal cost of each input (its price). The relative price of the L-intensive good is therefore:

$$\frac{P_L}{P_S} = \frac{p_L}{p_S} \frac{(n_S)^{\frac{1}{\sigma_S - 1}}}{(n_L)^{\frac{1}{\sigma_L - 1}}}. \tag{17.26}$$

Using (17.20) and (17.24), the equilibrium condition in the final goods market is:

$$\frac{Y_S^d}{Y_L^d} = \frac{Y_S^s}{Y_L^s} \implies \frac{P_L}{P_S} = \frac{(n_S)^{\frac{\sigma_S}{\varepsilon(\sigma_S - 1)}}}{(n_L)^{\frac{\sigma_L}{\varepsilon(\sigma_L - 1)}}} \tag{17.27}$$

Combining Eqs. (17.27) and (17.26) we obtain:

$$\frac{p_S}{p_L} = \frac{(n_L)^{\frac{\sigma_L - \varepsilon}{\varepsilon(\sigma_L - 1)}}}{(n_S)^{\frac{\sigma_S - \varepsilon}{\varepsilon(\sigma_S - 1)}}} \tag{17.28}$$

The number of varieties is determined by the equilibrium conditions in factor markets. Let $\bar{L}$ and $\bar{S}$ be the quantity of unskilled and skilled labour existing in the economy. By appropriate choice of parameter values, factor market equilibrium conditions are $n_S = \bar{S}$ and $n_L = \bar{L}$. Therefore $n_S/n_L = \bar{S}/\bar{L}$.

Now imagine a second identical economy and assume free trade. Then the number of intermediate inputs for each industry produced in the world economy is just twice the autarky number. Each good is therefore produced in free trade with twice as many inputs as in autarky. The ratio n_L/n_S remains unchanged in both countries, though the number of inputs in each good doubles, thanks to trade in intermediate inputs. The consequence on the skill premium can be seen by inspecting Eqs. (17.28) and (17.21). From (17.28) we see that p_S/p_L increases and from (17.21) we see that an increase in p_S/p_L drives up the skill premium.[13] A simple way of interpreting this result is that trade opening, through the availability of a larger number of intermediate inputs, increases the productivity of each input in Y_S relative to that of each input in Y_L. Therefore each intermediate input in Y_S is paid relatively more. Consequently, since the mark-up μ is constant, the skill premium is pulled upwards. It is worth mentioning that the relative price of the skill-intensive good declines due to the relative increase in its supply (see Eqs. 17.24 and 17.27). Nevertheless, the relative price of each input in Y_S increases. This is possible because, as shown by expression (17.25), the relative price of final goods depends positively on the relative price of intermediate inputs but negatively on the relative number of inputs.

17.4.2 Skill Biased Economies of Scale and Demand Elasticity

A possible additional link between trade integration and the skill premium may reside in the non-homotheticity of the production function. A production function is said to be homothetic when the factor intensity remains constant as output changes (with unchanged relative factor prices). When, on the contrary, the factor intensity depends on the size of output, the production function is said to be non-homothetic.[14] Clearly then, when the production function is non-homothetic, a change in output results in a change in the relative demand for factors at the firm level and also at the aggregate level, thereby affecting the skill premium. Dinopoulos et al. (2011) use this link to suggest that trade integration may drive up the skill premium because it induces a skill-biased expansion of output. There are two key assumptions in the model. First, the demand curve faced by each producer must become flatter (more elastic) when the economy moves from autarky to free trade. This induces imperfectly competitive producers to reduce the price and expand the

[13] As an example, assume $\sigma_L = 4$, $\sigma_S = 3$, $\varepsilon = 2$. Let $\bar{S} = \bar{L} = 10$ so that $n_S = n_L = 10$ in autarky and $n_S = n_L = 20$ in free trade. Therefore $w_S/w_L = 10^{1/12}$ in autarky and increases to $20^{1/12}$ in free trade.

[14] See any microeconomics textbook for further details on the concept of homotheticity.

output.[15] Second, the production function must be such that factor intensity depends on the size of firm's output. We now describe a simplified version of their model.

Each firm produces a different variety of a single consumption good. The market is in monopolistic competition but, unlike the monopolistic competition model presented in Sect. 9.2, here the perceived demand elasticity increases with trade opening. Firm output, y, is obtained from the sum powered to $1/\rho$ of *effective* factor inputs (L_e and S_e), each powered to ρ, that is:

$$y = [(L_e)^\rho + (S_e)^\rho]^{\frac{1}{\rho}} \qquad 0 < \rho < 1 \tag{17.29}$$

The *effective* factor inputs L_e and S_e are defined as:

$$L_e = y^{\phi_L} L, \tag{17.30}$$

$$S_e = y^{\phi_S} S, \tag{17.31}$$

where L and S are actual factor inputs and ϕ_L and ϕ_S are parameters whose value is between zero and one.

Expression (17.30) means that for every unit of labour input there is a corresponding effective unit equal to y^{ϕ_L}; it is as if the quantity of labour input was $y^{\phi_L} L$ instead of L. Likewise for the effective input of skilled labour defined by expression (17.31). To avoid confusion we emphasise that firms' output is not put back into the production process. The coefficients multiplying factor inputs do not represent physical input of the firm's output. They represent the efficiency of factors as depending on the quantity of output of the firm. As firm's output expands, factors become more productive according to a multiplicative term, y^{ϕ_L} and y^{ϕ_S}. Dinopoulos et al. (2011), assume that $\phi_S > \phi_L$; this means that non-homotheticity is skill-biased or, to put it differently, that any increase in output increases the effective relative input of skilled labour.[16] Firms maximize profits, which in equilibrium are driven to zero by the assumption of free entry (as in Sect. 9.2).

The equilibrium for the representative firm is characterized by two conditions: marginal revenue = marginal cost (profit maximization) and price = average cost (zero profit). Figure 17.7 represents the equilibrium for the representative firm in autarky (point A) and in free trade (point FT). The marginal revenue = marginal cost condition is not shown in the figure. The price = average cost condition is

[15]This sort of pro-competitive effect may emerge, for instance, in general equilibrium oligopoly à la Neary (2009) or Neary and Tharakan (2012), or in monopolistic competition with linear demand à la Ottaviano, Tabuchi, & Thisse (2002), or in monopolistic competition with translated additive demand à la Pollak (1971). The latter is the microfoundation chosen by Dinopoulos et al. (2011). We pass over the details of the microfoundation since they are not relevant to understanding the effect of trade opening on the skill premium.

[16]As an example consider $\phi_L = 1/4$ and $\phi_S = 3/4$. Then $S_e/L_e = \sqrt[2]{y}\,(S/L)$. Any increase in output—*ceteris paribus*—will increase the effective relative input of skilled labour.

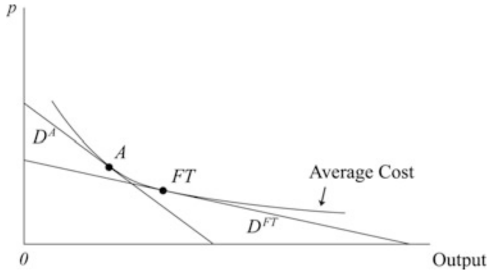


Fig. 17.7 Non homothetic technology, trade integration and firm output

represented by the tangency point between the demand curve and the average cost curve.[17]

Moving from autarky to free trade puts producers under competitive pressure. The market becomes larger (the intersection between the abscissa and the demand curve shifts to the right) but the elasticity of demand increases for any given price (the demand curve becomes flatter). Thus, each producer reacts by lowering the price and expanding the output. Here is where the non-homotheticity comes into play. The expansion of firms' output makes both factors more productive as shown by expressions (17.30)–(17.31), but under the assumption that $\phi_S > \phi_L$, the productivity of skilled labour increases more than the productivity of unskilled labour. Therefore, the relative demand for skilled labour increases and so does the skill premium in all countries.

17.5 Heterogeneous Firms, and the Skill Premium

Another channel through which trade opening may give rise to inequality is represented by the heterogeneity of firms' responses to trade opening. To study this channel, many works abandon the two-sector Heckscher-Ohlin structure and focus instead on wage inequality within a single industry. Models in these papers feature heterogeneous firms and/or heterogeneous labour and assume that the most productive firms are those that use skilled labour more intensively. Furthermore, they assume the presence of some kind of fixed exporting cost giving rise to a partition of firms between exporters and non-exporters. The triggering factor of wage inequality is trade integration, which indirectly reallocates firms' revenue from low-skill-intensive non-exporters towards the high-skill-intensive exporters. This reallocation benefits skilled labour relative to unskilled labour. Thus, we begin by studying the relationship between trade integration and revenues.

[17] Figure 17.7 represent the same equilibrium as Fig. 9.2 of Sect. 9.2.1. The only difference is in the form of the demand curve. In Sect. 9.2.1 moving from autarky to free trade induces entry of new firms but firm's output (for existing and new firms) is the same as in autarky. Here instead, moving from autarky to free trade induces firms to expand the output.

17.5.1 Trade Integration and Firm Revenue Inequality

The first step is to write the firm revenue as a function of output alone. This is intuitively a rather simple matter; since each firm faces a downward sloping demand curve, the price depends on the quantity the firm supplies to the market. Therefore, ultimately, revenue depends only on firm output. Let p be the price, y the output, and $R = py$ the revenue of the firm. Since some firms exports and some do not, we distinguish between domestic and foreign revenue. A firm has domestic revenue, R_d, and, if it exports, foreign revenue, R_f. The first depends on output sold domestically, y_d, and the second depends on output sold abroad, y_f. All works reviewed in this section adopt the type of monopolistic competition studied in Sect. 9.2, which allows revenue to be written as a function of output in the following simple way:

$$R = \underbrace{C_d \, (y_d)^{\alpha}}_{R_d} \qquad \text{if the firm does not export,} \tag{17.32}$$

$$R = \underbrace{C_d \, (y_d)^{\alpha}}_{R_d} + \underbrace{C_f \left(y_f\right)^{\alpha}}_{R_f} \qquad \text{if the firm exports,} \tag{17.33}$$

where $C_d, C_f > 0$ depend on variables that are not relevant for our purposes and α is a positive constant smaller than 1.[18] Exporting firms have larger revenue. Trade integration leads to an increase in revenue inequality for the reason that we now discuss. Following trade integration, foreign demand for domestic output increases because trade costs have declined. Thus, output produced for export increases and so does foreign revenue. This, obviously, only benefits exporting firms. Domestic demand for domestic output declines because some domestic expenditure is now reallocated to foreign varieties, since trade costs have declined. Output for the domestic market and domestic revenue therefore decline for all firms. It can be shown, however, that for exporting firms the increase in foreign revenue more than compensates for the decline in domestic revenue. Intuitively, the reason is that since exporting firms are also the most productive firms, they suffer less than other firms from the intensification of competition at home and benefit greatly from the improved access to the foreign market; their total sales, therefore, rise.[19] In sum, trade integration induces a reduction in y_d and an increase in y_f; domestic revenue declines for all firms, but exporting firms benefit from the increase in

[18]This is in fact the same α as in the utility function given by expression (9.2). It is not surprising that a parameter of the utility function ends up in the revenue function: after all, the willingness to pay for the item produced by a firm depends on the utility consumers obtain from it.

[19]Here and throughout this section we use the term "competition" somewhat loosely. In fact, what takes place in the domestic market is a market-crowding effect similar to that studied in Sect. 9.2.1.3. In that case the market crowding resulted in the exit of firms; here it results in a reduction of output and revenue.

foreign revenue. For exporting firms, total revenue rises, for non-exporting firms total revenue declines. Thus, trade integration gives rise to an increase in revenue inequality.

The mechanism that links trade integration to revenue inequality is common to all the models we study in this section. The difference between them lies in the way revenue inequality translates into wage inequality. We shall study these different ways in the next subsection.

17.5.2 Quality and Heterogeneous Fixed Inputs

A very direct way in which trade integration affects wages is proposed by Manasse and Turrini (2001). In their model, skilled labour is used as fixed input and is paid a wage equal to revenue minus variable costs. Therefore, revenue inequality translates directly into wage inequality. We present a simplified version of their model.

Goods are differentiated horizontally (by brand) and vertically (by quality). Workers differ in skills measured by s. We refer to the lowest skill as unskilled labour (s_0). Production of any variety requires one unit of skilled labour (any $s > s_0$) as a fixed input and one unit of unskilled labour per unit of output. The index s therefore identifies the skill level as well as the firm using that skill level. The quality of the variety increases with the level of skills. Thus, firms employing workers with higher skills produce higher-quality varieties. The market is in monopolistic competition and therefore the price of a variety, p_s, and the marginal cost, w_0, are in a constant proportion to each other given by:

$$\frac{p_s}{\mu} = w_0, \qquad \mu > 1, \tag{17.34}$$

where μ is the mark-up already encountered in Sect. 9.2. Marginal cost does not depend on the quality of the variety, since production only requires unskilled labour. Firms wanting to export face a fixed exporting cost in terms of one unit of unskilled labour and a variable trade cost of the iceberg type already introduced in Sect. 6.3. Firm's profit goes entirely to pay the skilled labour that made it possible to produce the variety of that particular quality. Thus, after paying all costs the firm makes zero profit. This gives the wage of skilled labour, w_s, as:

$$w_s = \underbrace{p_s y_{ds}}_{\text{Revenue}} - \underbrace{w_0 y_{ds}}_{\text{Variable cost}} \qquad \text{in non exporting firms,} \tag{17.35}$$

$$w_s = \underbrace{p_s y_s}_{\text{Revenue}} - \underbrace{w_0 y_s}_{\text{Variable cost}} - \underbrace{w_0}_{\text{Fixed exp. cost}} \qquad \text{in exporting firms,} \tag{17.36}$$

where y_{ds} is the output of a non-exporting firm and $y_s = y_{ds} + y_{fs}$ is total output of an exporting firm. Using expression (17.34) to substitute for w_0 in expressions

Fig. 17.8 Heterogeneous fixed cost and wage inequality

(17.35) and (17.36), we can write wages as a function of revenue as follows:

$$w_s = \frac{(\mu - 1)}{\mu} R_{ds} \qquad \text{in non exporting firms,} \tag{17.37}$$

$$w_s = \frac{(\mu - 1)}{\mu} \left(R_{ds} + R_{fs}\right) - w_0 \qquad \text{in exporting firms,} \tag{17.38}$$

where $R_{ds} = p_s y_{ds}, R_{fs} = p_s y_{fs}$. We can see from these two expressions that the wage increases with s since revenue increases with s (consumers like quality). The increase in wage is, however, faster for exporting firms because an increase in s increases both domestic and foreign revenues.

Turning our attention to export decisions, a firm decides to export only if it makes a non-negative profit on the foreign market. Obviously, only firms with sufficiently large revenues (high quality) find it profitable to pay the fixed exporting costs and export. Let s^* be the cut-off value of skills below which a firm decides not to export. By definition, the firm employing the worker with skill s^* makes zero profit on the foreign market. The distribution of wages among skilled workers will therefore take the shape represented by the solid line in Fig. 17.8. The wage of skilled workers increases with the level of skills, but the increase is faster for exporting firms because for these firms a marginal increase in quality increases sales at home *and* abroad.

Trade integration has two effects. First, the firm employing the worker with skill s^* is now making positive profits in the foreign market, because the reduction in trade costs has increased its foreign sales. Therefore the new cut-off value of skill, $s^{*\prime}$, will lie to the left of s^*. Second, trade integration triggers the reallocation of firms' revenues. Revenues of non-exporting firms decline because these firms suffer from stronger foreign competition. Consequently, skilled workers employed by non-exporting firms will see their wage decline. Firms which were exporters before trade integration suffer the same intensification of competition in the domestic market, but this is more than offset by the increase in foreign revenues. The wage of skilled workers in these firms increases. New exporters make larger revenues after trade integration, but they now pay the fixed exporting costs. For some of them, the increase in revenue more than compensates for the fixed costs and therefore the wage of skilled workers increases. For the other new exporters, the increase in revenue is not sufficient to allow for an increase in the wage of skilled labour. The dashed line in Fig. 17.8 shows the situation after the reduction of variable trade costs. Comparing the solid with the dashed line, we see that wages of workers in non-exporting firms decline, wages in newly exporting firms but with skills between $s^{*\prime}$ and $\bar{s}$ decline too, wages in all other firms increase.

We can now evaluate the consequences of trade integration on the skill premium. The presence of heterogenous labour means that the concept of skill premium needs to be qualified. To this purpose, let the wage ratio between any two workers, one in an exporting and the other in a non-exporting firm, be a measure of the skill premium. Inspection of Fig. 17.8 shows that trade integration induces a rise in the skill premium. For any two workers s (in a non-exporting firm) and s' (in an exporting firm), the wage ratio $w_{s'}/w_s$ is larger after trade integration than before. The effect of trade integration on inequality can also be measured by the change in relative aggregate revenue of highly-skilled workers. Highly-skilled workers means skilled workers employed by exporting firms. Since the employment share of exporting firms increases with trade integration (s^* moves to $s^{*\prime}$), trade integration implies a redistribution of total income from unskilled workers to skilled workers.

17.5.3 Endogenous Technology Adoption with Heterogeneous Workers

Yeaple (2005) develops a model where the most productive workers endogenously match with firms adopting the best technology. He suggests that trade integration induces an increase in wage inequality because it changes the matching between firms and workers in favour of highly-skilled workers to the detriment of moderately-skilled workers. To understand the logic of his argument, we now study a simplified version of his model.

The economy produces a differentiated good, Y, and a homogeneous good Z. There are three technologies indexed by $j = Z, H, L$. Technology Z is used only to produce the homogenous good Z. Good Y is produced by use of either technology H or technology L. Workers differ in skills, measured by s which takes values from 0 to infinity. Let ϕ_s^j denote the quantity of the good a worker with skill s can produce in industry j. We adopt the convention that the least-skilled worker can produce one unit of output regardless of technology, $\phi_0^j = 1$ for any j. For all other workers, ϕ_s^j depends on the skill level and on the technology the worker is using. For any given technology, a unit of high-skill labour produces more than a unit of low-skill labour, i.e., $\phi_{s'}^{j} > \phi_{s}^{j}$ for any $s' > s > 0$. For any given skill, technology H is more productive than L which, in turn, is more productive than Z, i.e., $\phi_s^H > \phi_s^L > \phi_s^Z$ for any $s > 0$. Furthermore, and crucially, highly-skilled workers have a comparative advantage in the use of the H technology relative to moderate and low-skilled workers, and moderately skilled workers have a comparative advantage in producing Y relative to low-skilled workers. Thus we have:

$$\frac{\phi_{s'}^H}{\phi_s^H} > \frac{\phi_{s'}^L}{\phi_s^L} > \frac{\phi_{s'}^Z}{\phi_s^Z} \text{ for any } s' > s > 0 \tag{17.39}$$

Good Z is produced in perfect competition, without fixed inputs, is chosen as the numéraire and its price is set to 1. Good Y is produced in monopolistic competition and requires fixed and variable inputs. Firms have to pay a fixed cost F_L to adopt technology L or a fixed cost F_H to adopt technology H. The H technology requires a higher fixed cost, $F_H > F_L$. Both fixed costs take the form of output that must be produced but cannot be sold.

Let w_s^j denote the wage of a worker with skill s using technology j. The total cost of producing y_j units of output for a firm employing workers with skill s is $\left(w_s^j/\phi_s^j\right) y_j + F_j$. Since the fixed cost takes the form of output, the unit cost (cost per unit of total output) is $\left(w_s^j/\phi_s^j\right)$. Every firm must choose one of three options: not paying any fixed cost and producing Z, paying F_L and producing Y with technology L, or paying F_H and producing Y with technology H. In equilibrium, some firms will employ low-productivity workers and produce Z, other firms will employ workers with intermediate productivity and produce Y with technology L, and the remaining firms will employ very productive workers and produce Y with technology H. The labour market is perfectly competitive, therefore wages adjust to equalize the unit cost of production for all firms using the same technology, that is:

$$\frac{w_s^j}{\phi_s^j} = \frac{w_{s'}^j}{\phi_{s'}^j} \text{ for any } s \neq s' \tag{17.40}$$

Let s_1 and s_2 be the threshold values of s above which firms adopt technology L and H respectively. A firm hiring workers of skill s_1 is indifferent between producing Z or producing Y with technology L. If it produces Y, it will have a lower unit cost but will incur fixed costs. By definition, when $s = s_1$, the lower unit cost exactly compensates for the fixed cost. Similarly, for a firm hiring workers with skill s_2, the lower unit cost of technology H exactly compensates for the higher fixed cost. Equation (17.40) characterizes the distribution of wages for any given value of s_1 and s_2.[20] The wage distribution derives from three conditions, which we now discuss in detail.

[20]The values s_1 and s_2 are determined endogenously by the equilibrium condition in the goods market and by the zero profit conditions. Since the determination of these values is not relevant for our purposes we disregard it for the sake of simplicity.

We begin by determining wages in industry Z, which are obtained as follows:

$$1 = w_s^Z/\phi_s^Z \Longrightarrow 1 = \frac{w_0^Z}{\phi_0^Z} \Longrightarrow w_s^Z = \phi_s^Z, \quad \text{for } 0 \leqslant s < s_1. \tag{17.41}$$

The first equation in (17.41), $1 = w_s^Z/\phi_s^Z$, is the price = marginal cost condition of perfectly competitive industry Z. Using this condition and $s' = 0$ in Eq. (17.40) gives the second of equations (17.41), which determines $w_0^Z = 1$ (recall that $\phi_0^Z = 1$). Then, substituting $w_0^Z = 1$ in (17.40) gives $w_s^Z = \phi_s^Z$ which is the third of equations (17.41). Wages in industry Y are determined by two arbitrage conditions. First, the wage of workers with skills s_1 must be the same in firms producing Z and in firms using technology L, otherwise these workers would not accept to be employed in either the production of Z or in the production of Y with technology L. Second, for the same reason, the wage of workers with skills s_2 must be the same in firms using technology H and technology L. The arbitrage conditions are therefore $w_{s_1}^Z = w_{s_1}^L$ and $w_{s_2}^L = w_{s_2}^H$. The first arbitrage condition and the third equation in (17.41) give the wage in L as follows:

$$w_s^L = \left(w_{s_1}^L/\phi_{s_1}^L\right)\phi_s^L \Longrightarrow w_s^L = \frac{\phi_{s_1}^Z}{\phi_{s_1}^L}\phi_s^L, \text{ for } s_1 \leqslant s < s_2 \tag{17.42}$$

To understand Eq. (17.42), the first step is to write Eq. (17.40) using $s' = s_1$ and $j = L$, which gives the first equation in (17.42). Then using $w_{s_1}^L = w_{s_1}^Z$ and the result obtained in Eq. (17.41), where we found that $w_s^Z = \phi_s^Z$, therefore $w_{s_1}^Z = \phi_{s_1}^Z$, gives w_s^L as in the second of equations (17.42). An analogous procedure applied to the arbitrage condition $w_{s_2}^L = w_{s_2}^H$ gives the wage w_s^H as shown in Eq. (17.43):

$$w_s^H = \left(w_{s_2}^H/\phi_{s_2}^H\right)\phi_s^H \Longrightarrow w_s^H = \frac{\phi_{s_1}^Z}{\phi_{s_1}^L}\frac{\phi_{s_2}^L}{\phi_{s_2}^H}\phi_s^H, \text{for } s \geqslant s_2 \tag{17.43}$$

Figure 17.9a shows the distribution of wages resulting from Eqs. (17.41) to (17.43).

The three lines (w_Z, w_L, and w_H) represent the distribution of wages for each of the three technologies. The line w_L is steeper than the line w_Z, reflecting higher productivity, but intersects the ordinate at a lower value, reflecting the lower shadow unit cost of production for any s between 0 and s_1.[21] Likewise for w_H with respect to w_L. The bold broken line represents the wage paid by firms to workers of different skills. Workers with skills between 0 and s_1 are employed by firms in the Z industry and are paid low wages. The skills of these workers are so low that if firms in industry Y employed them they would make negative profits. Workers with skills between s_1 and s_2 are employed by L technology firms and paid intermediate wages.

[21]The shadow unit cost is the unit cost that would obtain if the firm using technology L hired workers with skills between 0 and s_1.

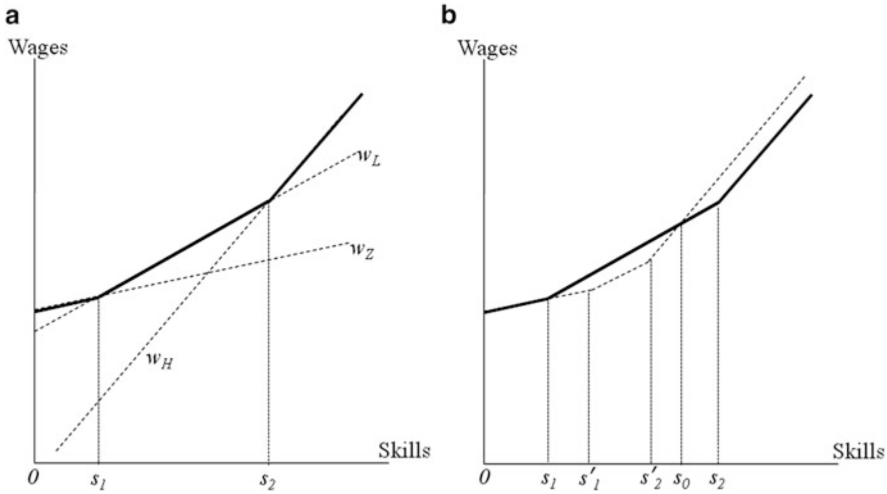


Fig. 17.9 Technology adoption and wage inequality. (**a**) Wage distribution. (**b**) Change in wage distribution

These workers are overqualified for industry Z but their skills are not high enough to give rise to non-negative profits when employed in firms using technology H. Lastly, workers with skills higher than s_2 are employed by firms adopting the most productive technology.

Consider now two identical economies opened to costly trade. Trade costs take the usual form of a fixed exporting cost F_X (for good Y only) and a variable trade cost of the iceberg type. The choice for firms in industry Y is either paying F_X and thus adding foreign sales to domestic sales, or not paying F_X and settling for domestic sales only. In the simple configurations that we discuss here, we assume that the set of firms adopting the H technology is identical to the set of firms deciding to export.[22] Therefore, in equilibrium, some firms use the L technology and do not export and the remaining firms in industry Y use the H technology and export. Since countries are identical, there is only intra-industry trade in the different varieties of Y. The consequences of trade integration on wage distribution are shown in Fig. 17.9b, where the dashed line represents a lower iceberg cost situation than the solid line.

When variable trade costs decline, the better access to foreign markets improves exporting firms' foreign revenue, while foreign competition in the domestic market deteriorates their domestic revenue. Overall, however, exporting firms' revenue increases. Therefore, the exporting firm that employed the worker with productivity s_2 is no longer indifferent between L-technology non-exporter status and H-technology exporter status; this firm will make larger profits by using the

[22]This only requires mild conditions on fixed costs.

H technology and exporting than otherwise. This is represented in Fig. 17.9b by a leftward shift of the cut-off value from s_2 to s_2' (due to a leftward shift of the line w_H in Fig. 17.9a). Non-exporting firms face fiercer competition from foreign exporters, their revenue declines and some of them succumb. Workers laid off by these firms will be employed by the Z industry (rightward shift of the cut-off value from s_1 to s_1').[23] The final situation is one in which the most highly-skilled workers gain from trade integration. Conversely, workers who are laid off by the Y industry see their wage decline, as do workers remaining in the Y industry and continuing to use the L technology. Workers remaining in the Z industry see their wage unchanged. Finally, workers with skills between s_2' and $\bar{s}$ see their wage decline: they now use a better technology but their skills are too low for them to benefit from it.

We now evaluate the consequence of trade integration on the skill premium. Again we define the skill premium as the wage ratio between any two workers, one in an exporting and the other in a non-exporting firm. Inspection of Fig. 17.9b shows that trade integration induces a rise in the skill premium. For any two workers s (in a non-exporting firm) and s' (in an exporting firm), the wage ratio $w_{s'}/w_s$ is larger after trade integration than before. Furthermore, the aggregate revenue of highly-skilled workers increases relative to the aggregate revenue of moderately-skilled workers.

17.5.4 Heterogeneous Hiring (Fixed) Costs

Helpman et al. (2010) argue that a possible source of within-industry heterogeneity is that more productive firms have stronger incentives to search and select more productive workers in costly trade than in autarky. Conversely, less productive firms have less incentive. Thus, in costly trade more than in autarky, the more productive workers end up matching with the more productive firms. Since the wage depends on average productivity at the firm level, the wage dispersion in the industry is larger in costly trade than in autarky. We now examine a simplified version of their model.

Firm output, Y, depends positively on the productivity of the firm, ϕ, on the number of workers employed by the firm, h, and on the average productivity of workers employed by the firm, $\bar{s}$. Thus, the production function of a firm is:

$$y = \phi h^{\gamma} \bar{s}, \qquad 0 < \gamma < 1. \tag{17.44}$$

A characteristic of this production function is that it is supermodular in $(\phi, \bar{s})$. Supermodularity (see, for example, Amir, 2005) refers to situations where an advantage begets further advantage. In the case of the production function (17.44),

[23]The expansion of the Y industry is consistent with the fact that the wage of workers with skills higher than $\bar{s}$ increases relative to good Y, thereby inducing an increase in the demand for this good.

supermodularity simply means that the more productive the firm is, the greater the effect of an increase in the average productivity of workers. Specifically, for any two firms with different productivity levels ϕ' and ϕ'', an equal increase in $\bar{s}$ will result in a larger output increase for the firm with higher productivity.[24]

Productivity is assigned randomly to firms, while the productivity of workers results from a search and screening activity undertaken by firms, which we now discuss.[25] Workers differ in skills measured by $s \geqslant 1$. Firms incur search and screening costs of employment. The search cost is represented by the cost of matching with workers (think of the administrative cost of opening a vacancy). To match randomly with n workers seeking a job, a firm pays bn units of the numéraire. The marginal cost of searching is therefore equal to b. The screening cost is represented by the cost of evaluating the productivity of each of the n workers sampled (think of the cost of job interviews). The screening procedure allows a firm to identify workers with skills up to a given threshold level chosen by the firm. Let $\underline{s}$ be such threshold. Screening is costly and it is assumed that screening costs are increasing in the threshold $\underline{s}$. This is plausibly justified by the need to set up more elaborate tests to identify higher skills. Specifically, it is assumed that by paying $c\underline{s}^2/2$ units of the numéraire, where $c > 0$, the firm is able to identify workers with skills lower than $\underline{s}$; the marginal cost of screening is therefore $c\underline{s}$. The number of workers actually employed by the firm, h, increases with the number of workers sampled, n, and decreases with the threshold $\underline{s}$ chosen by the firm. To be specific, assume that $h = n\,(1/\underline{s})^k$, $k > 1$.[26] Recall that $s \geqslant 1$, therefore the number of workers employed is simply a fraction $(1/\underline{s})^k < 1$ of the number of workers sampled. The average skill of workers employed by the firm, denoted by $\bar{s}$, turns out to be $\bar{s} = k\underline{s}/\,(k-1)$. Replacing $h = n\,(1/\underline{s})^k$ and $\bar{s} = k\underline{s}/\,(k-1)$ in the production function (17.44) yields:

$$y = \frac{k}{k-1}\phi n^{\gamma}\underline{s}^{1-\gamma k} \tag{17.45}$$

where it is assumed that $0 < \gamma k < 1$.

Firms operate in monopolistic competition and their revenue depends on output, as shown by expressions (17.32) and (17.33). Substituting (17.45) into y_d and y_f of expressions (17.32) and (17.33) shows that a firm's revenue depends positively on its productivity, ϕ, the number of workers sampled, n, and the threshold level $\bar{s}$.

[24]Consider, for instance, two firms with the same employment $h = 1$ and with productivity levels $\phi' = 1$ and $\phi'' = 2$. Imagine they experience the same increase in average productivity of workers, $\Delta\bar{s} = 1$. The output increases are, respectively, $\phi'\Delta\bar{s}$ and $\phi''\Delta\bar{s}$ where clearly the latter is larger than the former since $\phi'' > \phi'$.

[25]The labour market is modeled along the lines of Diamond-Mortensen-Pissarides search and matching frictions.

[26]This specification is obtained by assuming that the skill distribution is Pareto with shape parameter k and lower bound equal to one. This distribution also gives rise to the expression for the average skill of workers employed by the firm specified below.

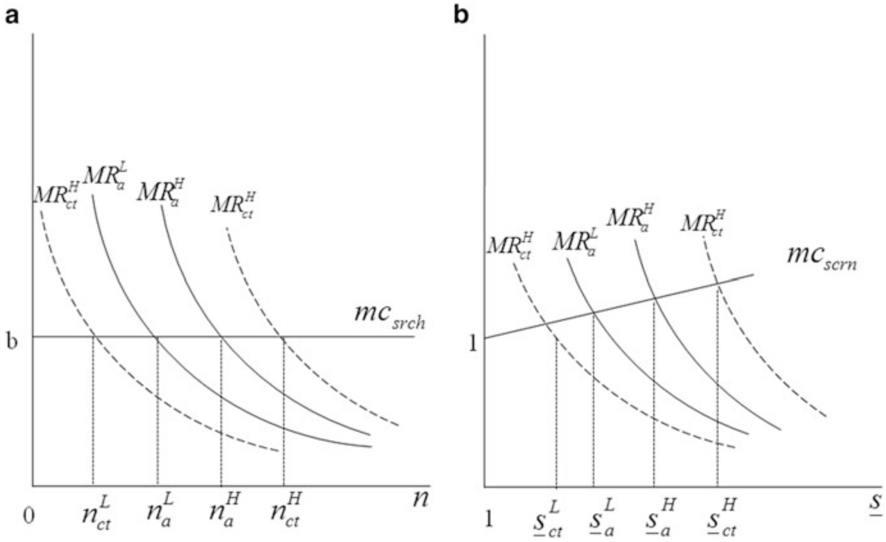


Fig. 17.10 Search and screening costs: firm's equilibrium

Figures 17.10 and 17.11 depict the autarky and costly trade equilibria. To simplify the graphical treatment, we represent the situation for only two firms: a high productivity firm, H, and a low productivity firm, L. Consider first the autarky situation represented by the solid lines. The declining solid lines in Fig. 17.10a represent the marginal revenue plotted against the number of workers sampled for the high and low productivity firms, MR_a^H and MR_a^L respectively. The H firm has a higher marginal revenue simply because it has a larger ϕ, and a worker of any given skill generates more revenue if employed by the H firm than by the L firm. Firms choose the number of workers to be sampled by equating marginal revenue to the marginal cost of searching ($mc_{srch} = b$). Thus, the equilibrium number of workers sampled is n_a^H and n_a^L with $n_a^H > n_a^L$, as shown in Fig. 17.10a. Figure 17.10b shows the marginal revenues MR_a^H and MR_a^L plotted against the threshold level of screening for any given number of workers sampled.

Supermodularity entails that a higher average productivity of workers is more valuable in firms with higher ϕ. Therefore, for any given n, the more productive firm has larger marginal revenues from screening, as represented in Fig. 17.10a. Firms choose the screening threshold $\underline{s}$ so as to equalize marginal revenue with the marginal cost of screening (mc_{scrn}).[27] The equilibrium screening thresholds are therefore $\underline{s}_a^H$ and $\underline{s}_a^L$; the more productive firm screens more severely.

[27] It is worth mentioning that the marginal revenue lines in Fig. 17.10a are plotted for given values of $\underline{s}$. The revenue lines in Fig. 17.10b are plotted for given values of n. There is therefore a relationship between the optimal number of workers sampled and the optimal level of screening. This relationship is not relevant for understanding the model and we therefore disregard it.

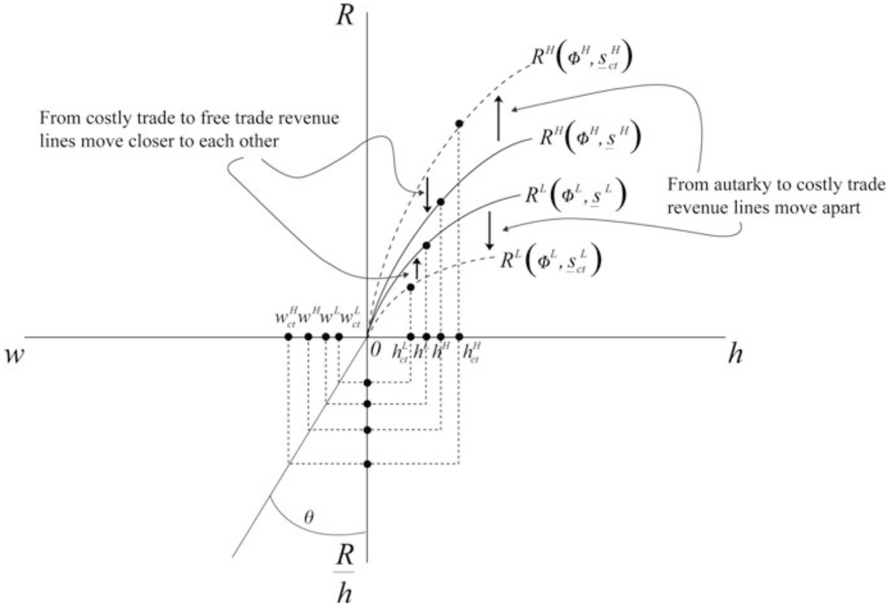


Fig. 17.11 Search and screening costs: general equilibrium

In the North-East quadrant of Fig. 17.11 we see that firm H has larger revenues for any $h > 0$. This is due to two reasons: first, H is more productive ($\phi^H > \phi^L$), and second, H has screened more severely ($\bar{s}^H_a > \bar{s}^L_a$). Although the H firm samples a larger number of workers it may end up hiring an equal or smaller number than the L firm because of more severe screening. However, under mild conditions on the parameters which we assume to hold, it turns out that $h^H_a > h^L_a$.

We now turn to wage determination. Each firm negotiates the wage with its employees. The result of this negotiation is that each party gets a constant fraction of total revenues and the individual wage is equal to the total wage bill divided by the number of workers employed in the firm. Let $\theta \in (0, 1)$ be the fraction of total revenue that goes to wages. Graphically, the revenue per worker corresponds to the size of the angle formed by a straight line emanating from the origin and reaching the dot on each of the revenue lines. For convenience of visual inspection, we plot the revenue per worker on the vertical axis of the South-East quadrant and wages on the horizontal axis of the North-West quadrant. Clearly, the H firm pays higher wages.

We can now examine the effect of trade opening from an initial situation of autarky. The costly trade situation (labelled ct) is represented by the dashed lines and is characterized by the presence of fixed and variable exporting costs. Let us assume that in costly trade, the H firm finds it profitable to export and the L firm does not. Moving from autarky to costly trade brings about a reallocation of revenues. The exporting firm loses revenue at home because of the foreign competition but this is more than offset by foreign sales. The non-exporting firm suffers the competition of foreign firms and its revenue declines. This reallocation of sales is represented in Fig. 17.10 where the marginal revenue of firm H shifts to the right and that of firm L to the left. As a result, in costly trade, firm H samples and screens more than in autarky while firm L samples and screens less. Therefore, as we see by inspecting the expression (17.45), output (and revenue) increase in firm H and decline in firm L. Graphically this means that the revenue lines move apart when passing from autarky to free trade, as shown by the dashed lines in Fig. 17.11. The consequences on wage inequality are immediately found. Taking the wage ratio between H and L as a simple measure of wage inequality, we observe that $w_{ct}^H/w_{ct}^L > w_a^H/w_a^L$; wage inequality has increased in passing from autarky to costly trade.

In sum, the model tells the following story: a reduction in trade costs makes searching and screening more profitable for highly productive firms and less profitable for less productive firms. Average labour productivity and revenue decline in less productive firms and increase in highly productive firms. Since wages are in fixed proportions to revenues, wage inequality increases. The mechanism generating wage inequality can also be seen in a different way if we observe that when the wage negotiation and employment decisions are taken, the search and screening costs have been paid and they are not retrievable. Thus, at this stage of the firm's decision, searching and screening costs are a fixed and sunk cost. In this perspective, we could say that in this model there are fixed but adjustable costs of hiring. Trade integration magnifies revenue inequality which, in turn, heightens wage inequality through heterogeneous responses of firms concerning the optimal amount of fixed hiring costs they want to pay.

From the analysis above, one might carelessly conclude that any step forward in trade integration brings about an increase in wage inequality. This would be a wrong conclusion. Indeed, when all firms export (say, in free trade), the ratio of revenues between firm H and firm L is—ceteris paribus—the same as in autarky. There is no longer a partition between exporting and non-exporting firms. In free trade, firms return to the same sampling and screening as in autarky and the wage inequality returns to the autarky level. There is, therefore, a hump-shaped relationship between trade opening and wage inequality, first increasing then decreasing until wage inequality reaches its initial level. This relationship is depicted in Fig. 17.12.

17.5.5 Heterogeneous Globalization Modes

It is well known that a large and growing share of world trade is in intermediate inputs. Amiti and Davis (2011) study the effect of globalization on wages taking account of trade in intermediates in addition to trade in consumption goods. In addition to the partition of all firms into exporters and non-exporters, their model generates an endogenous partition of all firms into importers (of intermediate inputs) and non-importers. Thus, a firm may (or may not) relate to other countries in four

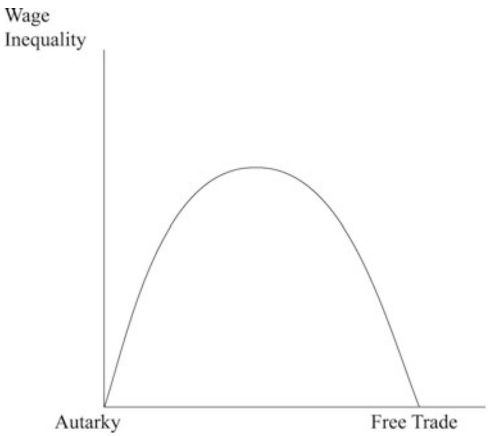


Fig. 17.12 Search and screening costs: from autarky to free trade

ways: by exporting and importing, by exporting and not importing, by importing and not exporting, or by neither exporting nor importing. This choice is endogenous and Amiti and Davis refer to it as the globalization mode chosen by the firm. They find that when firms choose different globalization modes, trade integration affects the skill premium because it magnifies the revenue inequality and/or the marginal cost inequality across firms. To understand the logic of their argument, we now describe a simplified version of their model.

Intermediate inputs are produced in each country in a fixed number of varieties. Let n_d and n_f be the number of varieties produced domestically and abroad, respectively. Producing one unit of intermediate input requires one unit of labour. All intermediates are priced at marginal cost. The production of the final good requires one unit of labour as fixed input and both labour and intermediate inputs in the production of output. An important property of the technology of production of final goods is that—ceteris paribus—the marginal cost declines as the number of varieties of intermediate inputs increases.[28]

Intermediate inputs and final goods are traded internationally at an iceberg trade cost which has two components: a component common to all firms, $\tau_X \in (0, 1)$ for the final good and $\tau_M \in (0, 1)$ for the intermediate inputs; and a component specific to each firm v, t_{Xv} and t_{Mv} for final goods and intermediate inputs, respectively. Let $\tau_{Mv} \equiv \tau_M t_{Mv}$ and $\tau_{Xv} \equiv \tau_X t_{Xv}$. So, an exporting firm will see a fraction $(1 - \tau_{Xv})$ of the good exported "melting" in transit and an importing firm will see a fraction $(1 - \tau_{Mv})$ of the intermediate good bought abroad "melting" in transit. Production, export and import are subject to the fixed costs F, F_X, and F_M respectively. Thus a firm which produces only for the domestic market and uses only domestic intermediates has a fixed cost $F_v = F$, a firm which produces for the domestic and foreign markets and uses only domestic intermediates has a fixed cost $F_v = F + F_X$,

[28]The cost function resulting from the production function in expression (17.5) has this property.

a firm which produces for the domestic market only and uses domestic and imported intermediates has a fixed cost $F_v = F + F_M$, and a firm which produces for the domestic and foreign markets and uses domestic and foreign intermediates has a fixed cost $F_v = F + F_M + F_X$.

Firms are heterogenous in three respects: productivity, ϕ_v, unit exporting cost, $(1 - \tau_X t_{Xv})$, and unit importing cost, $(1 - \tau_M t_{Mv})$. The triplet (ϕ_v, t_{Mv}, t_{Xv}) is assigned randomly to firms. Once the triplet is assigned, firms have to make three decisions: the profit-maximizing price, whether or not to produce, and the mode of globalization (import only, export only, both, or none). The market for final goods is characterized by monopolistic competition and the profit maximizing, p_v, is therefore a multiple of the firm's marginal cost, c_v:

$$p_v = \mu c_v, \tag{17.46}$$

where $\mu > 1$ is the mark-up already encountered in Sect. 9.2. The mark-up is common to all firms but the marginal cost is different in different firms (and so is the price) because firms have different productivity levels and because they may have chosen different modes of globalization. To see this, we take a closer look at marginal cost, c_v:

$$c_v = f\left(\underbrace{\phi_v}_{-}, \underbrace{w_v}_{+}, \underbrace{n_d}_{-}\right) \qquad \text{for non importing firms,} \tag{17.47}$$

$$c_v = f\left(\underbrace{\phi_v}_{-}, \underbrace{w_v}_{+}, \underbrace{n_d + \tau_{Mv} n_f}_{-}\right) \qquad \text{for importing firms.} \tag{17.48}$$

As above in this chapter, the notation $f\,(.,.,.)$ simply means that the marginal cost of any firm v depends on its productivity, ϕ_v, on the wage it pays to labour, w_v, and on the number of varieties it uses, n_d or $n_d + \tau_{Mv} n_f$. As usual, the algebraic signs below each of the three variables indicate the relationship between each of them and marginal cost. Pretty intuitively, an increase in productivity reduces marginal cost, an increase in wage increases marginal cost, and an increase in the number of varieties (moving from non-importing to importing) reduces marginal cost.

Firm revenue depends on output as in expressions (17.32) and (17.33); and firm profit, π_v, may conveniently be written in terms of revenue as:

$$\pi_v = \frac{(\mu - 1)}{\mu} R_v - F_v, \tag{17.49}$$

where we keep in mind that R_v and F_v depend on the mode of globalization and on productivity. The decision on whether or not to produce and the mode of globalization depend on the productivity level drawn by the firm. The most productive firms will find it optimal to pay $F + F_M + F_X$ in order to export and import. Slightly less productive firms will find it optimal to pay either $F + F_X$ or $F + F_M$ to be able to export or import, respectively. The other firms will produce only for the domestic market and will use only domestic inputs. Turning to wage determination, to simplify matters we assume that the wage results from a bargaining process and is equal to a constant fraction of profit:

$$w_v = \theta \pi_v, \quad 0 < \theta < 1 \tag{17.50}$$

We now turn to the effect of trade integration on wage inequalities. Consider first a trade integration in intermediate inputs (an increase in τ_M). Trade integration in intermediate inputs reduces the marginal cost of importing firms because it increases τ_{Mv} while keeping the marginal cost of the non-importing firms the same (see expression (17.47) and (17.48)). Hence the prices of the varieties produced by the importing firms decline (see expression (17.46)) and output increases. Therefore, profits and wages also increase for these firms (see expression (17.49) and (17.50)). Wage dispersion increases because the wage paid by importing firms relative to non-importing firms increases. But this is not all: the effect of trade integration is amplified because firms have different individual trade costs. An increase in the common component τ_M has a stronger impact on τ_{Mv} for firms with higher t_{Mv}. Therefore, among the importing firms, revenues, profits, and wages increase more for firms with lower import trade costs. This amplifies the wage inequality. Now let us consider trade integration in the final good (an increase in τ_X). This type of integration increases revenue inequality; the revenue of exporting firms increases and the revenue of non-exporting firms declines. Wage inequality increases because the wages paid by exporting firms increase relative to those paid by non-exporting firms. The wage dispersion is amplified by the fact that the trade cost of exports declines more for firms who have lower individual trade costs for final goods.

There is also a synergy between these two effects. To understand this synergy, consider trade integration in inputs. The revenues of importing firms increase because of the decline in the marginal cost, which makes these firms more competitive. All importing firms benefit from trade integration in inputs but the exporting firms (if there are any among the importing firms) benefit more than the non-exporters because they gain competitiveness at home and abroad.

In conclusion, both types of trade integration increase wage inequality, but they do so in different ways. A decline in output tariffs increases the wages of workers employed in exporting firms relative to those of workers employed in non-exporting firms. A decline in input tariffs raises the wages of workers employed in firms using imported inputs relative to the wages paid by firms that do not import inputs. Furthermore, there is a synergy between these effects. Lastly, wage inequality is magnified by the fact that firms have different firm-specific trade costs.

17.6 Endogenous Market Size

In the new economic geography models, trade integration may give rise to industrial agglomeration (see Chap. 16). In these models, agglomeration is often associated with wage inequality between regions. This is the case, for instance, in the core-periphery model. In this section we study the relationship between agglomeration and wage inequality by using the model developed by Krugman and Venables (1995) and studied in Sect. 16.3. In that section, the parametrization was chosen so that income inequality did not arise even in the presence of agglomeration. Here, a simple additional assumption will cause agglomeration to give rise to wage inequality. However, wage inequality arises for intermediate levels of trade costs, whereas low trade costs lead to the convergence of wages.

The basic model considers two regions, conventionally called North and South, each producing two commodities: "agricultural" goods and "manufactured" goods. Agricultural goods are produced under a constant return to scale in a perfectly competitive setting with labour as the sole input. Manufactured goods are differentiated goods produced under increasing returns to scale in a monopolistically competitive setting, using labour and a composite manufacturing intermediate good. Thus the manufacturing sector produces both final consumer goods and intermediate goods to be used as inputs.

At the beginning, no trade exists because of prohibitive transport costs, and both regions produce both kinds of goods in autarky. It is assumed that both regions are equal, in the sense that they are equally efficient in the production of both types of goods, so that neither region has any intrinsic comparative advantage in manufacturing. However, one region (say, North) has a larger manufacturing sector than the other.

Let us now assume that transport costs are gradually reduced, so that the possibility of trade in manufactured goods arises. As we know from the monopolistic competition model of international trade (see Sect. 9.2.2), there will be intra-industry trade in manufactured goods, with neither region becoming fully specialized in them. But as transport costs continue falling, a cumulative process will arise due to locational factors of the following type.

The initially larger manufacturing sector in North offers a larger market for intermediate goods, which makes this region (ceteris paribus) more advantageous for localizing the production of these goods. Such an effect is called a demand or "backward" linkage. The immediate consequence is that a greater number of intermediate goods will be produced in North than in South.

The availability of intermediate goods will then become better and better in North with respect to South, which means (again ceteris paribus) lower production costs of final goods; this effect is called a cost or "forward" linkage. Hence, further manufacturing production will be attracted to North, and so on: a trend towards the agglomeration of manufacturing in North is set in motion. There will be some critical value of transport costs below which the world economy will self-organize into a de-industrialized periphery and an industrialized core (the model thus explains the core-periphery pattern of world development). What is important to note is that this outcome is completely spontaneous, due to the self-organizing forces of the global economy (on self-organization in general see, for example, Gandolfo, 2009, chap. 25, sect. 25.6.2).

Assume that the manufacturing sector is large enough not to fit in one country.[29] Then the higher labour demand in the industrializing region (the core) will drive up real wages, while the falling demand for labour in the de-industrializing region (the periphery) will cause a decline in real wages there. In a nutshell: *globalization leads to inequality*.

However, this is not the end of the story, since a further decline in transportation costs has striking effects. In fact, the importance of being close to suppliers of intermediate goods and to markets for final goods (the backward and forward linkages) declines in concomitance with the decline in transport costs. On the other hand, the lower wage rate in the periphery is an important factor in production-cost calculations. There will be a certain threshold of transportation costs below which the lower wage rate in the periphery more than offsets the distance factor (i.e., the disadvantage of being far from suppliers and markets). Below this threshold value, manufacturing will find it profitable to relocate to the periphery. The higher labour demand there, and the lower demand for labour in the core, will bring about convergence of real wages.

Thus, after the initial formation of a core-periphery pattern, whereby globalization (due to declining transport costs) divides the world into rich and poor nations, further integration of world markets will bring about a convergence in incomes and economic structures.

17.7 Conclusion

The models studied in this chapter offer a rich set of plausible explanations for the increase in wage inequality. One of them is the Stolper-Samuelson effect, operating both directly and through skill-biased technical change. Another plausible explanation is that economies of scale affect the relative demand for factors. A third explanation hinges on firm revenue inequality. The mechanism through which revenue inequality translates into wage inequality differs in different models, but in all the models it requires some kind of heterogeneity whereby skilled labour is used relatively more in exporting firms than in non-exporting firms (either because of a skill bias in production or because of fixed costs in terms of skilled labour). Trade integration may also trigger an increase in marginal cost inequality which in turn causes wage inequality to rise. This increase in wage inequality is heightened by the export status of the firm and by the fact that firms have heterogeneous trade costs.

[29]This assumption is crucial and makes the difference between the parametrization of this model in this section and the parametrization in Sect. 16.3.

Lastly, trade integration unleashes agglomeration forces that may give rise to an increase in wage inequality across countries.

The literature on this topic is growing very rapidly and an exhaustive review is far beyond the scope of this chapter. The studies reviewed above, however, cover most of the explanations for the increase in wage inequality provided by the literature to date.

References

Acemoglu, D. (2002). Directed technical change.

Acemoglu, D. (2003). Patterns of skill premia.

Amir, R. (2005). Supermodularity and complementarity in economics: An elementary survey.

Amiti, M., & Davis, D. R. (2011). Trade, firms, and wages: Theory and evidence.

Antràs, P., Garicano, L., & Rossi-Hansberg, E. (2006). Offshoring in a knowledge economy.

Antràs, P., & Helpman, E. (2004). Global sourcing.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade.

Burstein, A., & Vogel, J. (2012). Globalization, technology, and the skill premium.

Bustos, P. (2011). Trade liberalization, exports and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms.

Costinot, A., & Vogel, J. (2010). Matching and inequality in the World economy.

Deardoff, A. V. (2000). Factor prices and the factor content of trade revisited: What's the use?

Dinopoulos, E., Syropoulos, C., Xu, B., & Yotov, Y.V. (2011). Intraindustry trade and the skill premium: Theory and evidence.

Epifani, P. & Gancia, G. (2008). The skill bias of world trade.

Ethier, W. J. (1982). National and international returns to scale in the modern theory of international trade.

Feenstra, R. C., & Hanson, G. H. (1997). Foreign direct investment and relative wages: Evidence from Mexico's Maquiladoras.

Feenstra, R. C., & Hanson, G. H. (1999). The impact of outsourcing and high-technology capital on wages: Estimates for the U.S., 1979–1990.

Garicano, L., & Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy.

Goldberg, P., & Pavcnik, N. (2007). Distributional effects of globalization in developing countries.

Grossman, G. M., & Rossi-Hansberg, E. (2008). A simple theory of offshoring.

Helpman, E., Itskhoki, O., & Redding, S. (2010). Inequality and unemployment in a global economy.

Krugman, P. R. (2000). Technology, trade and factor prices.

Krugman, P. R., & Venables, A. J. (1995). Globalization and the inequality of nations.

Leamer, E. E. (2000). What's the use of factor contents?.

Manasse, P., & Turrini, A. (2001). Trade, wages, and 'superstars'.

Neary J. P. (2002). Foreign competition and wage inequality.

Neary, J. P. (2009). International trade in general oligopolistic equilibrium, mimeo, Oxford University.

Neary, J. P., & Tharakan, J. (2012). International trade with endogenous mode of competition in general equilibrium.

Panagariya, A. (2000). Evaluating the factor-content approach to measuring the effect of trade on wage inequality.

Pollak, R. A. (1971). Additive utility function and linear engel curves.

Unel, B. (2010). Firm heterogeneity, trade, and wage inequality.

Williamson, G. J. (2006). *Globalization and the poor periphery before 1950.*

Yeaple, S. R. (2005). A simple model of firm heterogeneity and international trade, and wages.

Part VII
Appendices

Chapter 18
Appendix to Chapter 2

18.1 Maximization of World Income and the Dual Problem

We have seen in Sect. 2.3 that the theory of comparative costs can be given a modern formulation in terms of optimization (see, for example, Chipman, 1965; Dorfman, Samuelson, & Solow, 1958; Hartwick, 1979; Jones, 1961; McKenzie, 1954a,b, 1955; Takayama, 1972; Whitin, 1953). We examine here—following Takayama (1972, chap. 6)—the general case of m goods and n countries, which therefore also serves as a mathematical treatment of the generalizations examined in Sect. 2.4. The notation adopted is ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m$):

x_{ij} = quantity of good i produced in country j,
l_{ij} = constant labour-input coefficient in the production of good i in country j,
L_j = total quantity of labour existing in country j,
p_i = given international price of good i.

The problem of maximizing the value of world income (output) can then be formulated as follows

$$\begin{aligned}&\text{Max } p_1\left(\textstyle\sum_{j=1}^m x_{1j}\right) + p_2\left(\textstyle\sum_{j=1}^m x_{2j}\right) + \ldots + p_n\left(\textstyle\sum_{j=1}^m x_{nj}\right)\\ &= \text{Max } \textstyle\sum_{i=1}^n p_i\left(\textstyle\sum_{j=1}^m x_{ij}\right),\end{aligned} \tag{18.1}$$

subject to

$$\sum_{i=1}^n l_{ij}x_{ij} \le L_j, \quad j = 1, 2, \ldots, m,$$

$$x_{ij} \ge 0, \quad i = 1, 2, \ldots n; \quad j = 1, 2, \ldots, m.$$

The maximand is obviously (the value of) world output, as $\sum_{j=1}^m x_{ij}$ is nothing other than the sum of the quantities of good i produced in all countries. The first

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_18,

constraint states that the amount of labour totally employed in each country cannot exceed the amount available, and the second is the non-negativity of outputs.

We have formulated the problem directly in terms of maximization of world output; an alternative formulation—explained in the text in the simple 2×2 case—is in terms of maximization of the (value of) national output of each country separately considered. It is however also true in the general case that these two optimum problems are equivalent, for it can be shown (Takayama, 1972, pp. 172–173) that world output will be maximized if; and only if, each country maximizes its own national output.

In linear programming theory each problem has its dual, which in the case of problem (18.1) turns out to be

$$\text{Min} \sum_{j=1}^{m} w_j L_{j,} \tag{18.2}$$

subject to

$$\begin{aligned} w_j l_{ij} &\geq p_i, \quad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m; \\ w_j &\geq 0, \quad j = 1, 2, \ldots, m, \end{aligned}$$

where w_j is the unknown money wage rate (shadow price of labour) in country j. Therefore the dual problem consists in minimizing the world total labour reward (world production cost) subject both to the constraint that the price of a good cannot be greater than its production cost and to the non-negativity constraint on the wage rate.

The solution to the primal problem (18.1) yields the optimal quantities, which we shall denote by x_{ij}^*,whilst the solution to the dual problem (18.2) yields the optimal wage rates, denoted by w_j^*. At this point the fundamental duality theorem of linear programming enables us to state the following propositions (Takayama, 1972, p.162):

(a) $\sum_{i=1}^{m} p_i \sum_{j=1}^{m} x_{ij}^* = \sum_{j=1}^{m} w_j^* L_{j,}$

(b) If $\sum_{i=1}^{n} l_{ij} x_{ij}^* < L_j$, then $w_j^* = 0, j = 1, \ldots, m$,
(c) If $w_j^* > 0$, then $\sum_{i=1}^{n} l_{ij} x_{ij}^* = L_j, j = 1, \ldots, m$,
(d) If $w_j^* l_{ij} > p_i$, then $x_{ij}^* = 0, i = 1, \ldots, n; j = 1, \ldots, m$,
(e) If $x_{ij}^* > 0$, then $w_j^* l_{ij} = p_{i,} i = 1, \ldots, n; j = 1, \ldots, m$.

Proposition (a) means that (the value of) world output coincides with total factor income of the world. Propositions (b) and (c) mean that, if labour is not fully utilized in country j, then its price (money wage rate) must be zero there, whilst, on the contrary, if the money wage rate is positive in the j-th country, then all of the labour available in that country must be fully utilized. Propositions (d) and (e) mean that if the unit cost of production of good i in country j is greater than the price of this good, then good i will not be produced in country j, whilst, on the contrary, if the output of good i in country j is positive, then its unit cost of production there, will exactly equal its price.

All these propositions of course refer to the optimum point (of both the primal and the dual) and constitute an extension to the world economic system of results well known in the theory of general economic equilibrium in a closed economy (assuming that production takes place according to the same hypotheses at the basis of the Ricardian theory).

Let us note that proposition (e), apart from notational differences, is the same as Eq. (2.18), on which the treatment in the text is based. Therefore, the money wage rates—which in that treatment were assumed to be exogenously given—can actually be considered as the *shadow prices of labour*, obtained from the solution of the linear programming problem (18.2). This is so because—as we know from general equilibrium theory (see Dorfman et al., 1958)—in a system where perfect competition obtains and all agents follow a maximizing behaviour, the money wage rate(s) will turn out to be equal to such shadow price(s). See also Takayama (1972, chap. 7).

18.2 Maximization of National Income and Minimization of Real Cost

As was shown in the text in the course of the examination of the simple Ricardian example, the gains from trade can be seen from two points of view. On the one hand, they can be considered as a saving of labour (reduction in the real cost of production), obtained by importing the commodity in which the country has the smaller comparative advantage or the greater comparative disadvantage instead of producing it domestically; on the other, as an increase in the amount of commodities obtainable with the same input of labour. It follows that the optimal situation sought for can be considered both as the minimization of the real cost (in terms of labour) required to achieve a given national income (output), and as the maximization of national income (output) given the available amount of labour.

We have so far examined the latter problem; let us now examine the former, considering each country separately. In fact, as stated in Sect. 18.1, this problem is equivalent to the maximization of world income.

For this purpose, it is convenient first to rewrite the problem of national income maximization in the form

$$\text{Max} \sum_{i=1}^{n} p_i x_i, \tag{18.3}$$

subject to

$$\sum_{i=1}^{n} l_i x_i \leq L, \quad x_i \geq 0,$$

where for brevity we have dropped the subscript j, as it is understood that the optimization must be performed for each country. The problem under examination is now

$$\text{Min} \sum_{i=1}^{n} l_i x_i, \tag{18.4}$$

subject to

$$\sum_{i=1}^{n} p_i x_i \geq Y, \quad x_i \geq 0,$$

where Y is the value of any *feasible* output combination, namely

$$Y = \sum_{i=1}^{n} p_i \bar{x}_i, \text{ with } \sum_{i=1}^{n} l_i \bar{x}_i \leq L \text{ and } \bar{x}_i \geq 0, \tag{18.5}$$

where the $\bar{x}_i$'s are given quantities.

Problem (18.3) is the usual one of maximization of national income (output), whilst (18.4) is the one of minimum real cost. To avoid misunderstanding, it is as well to stress the fact that problem (18.4) must *not* be confused with the dual problem to (18.3), which would be of the type (18.2) and would consist in the minimization of the total labour reward.[1] In fact, problem (18.4) requires the minimization of the *physical quantity* of labour input and not of the total labour reward.

We now introduce some simplifying assumptions. The first is $l_i > 0$, that is, that any good, wherever produced, requires *some* labour input. The second is $p_i > 0$, that is, that any good has a positive price (free goods being ruled out). The third is that the ratios p_i/l_i are all different from each other. The stated economic meaning of these assumptions is entirely plausible.

Given these assumptions, it is possible to prove—by using a theorem by Kuhn (see Takayama, 1972, pp. 174–175)—that good i is produced in the optimum quantity for problem (18.3) *if, and only if*, it is produced in the optimum quantity for problem (18.4). It follows that the solutions of the two problems coincide, namely

[1]The dual to problem (18.3) is Min wL subject to $wl_i \geq p_i, w \geq 0$.

the output combination which maximizes national income (output) is the same as that which minimizes the real cost of producing it.

18.3 A Ricardian Model with a Continuum of Goods

The two-country, m-goods Ricardian model (see the text, Sect. 2.4.2) has been extended by Dornbusch, Fischer, and Samuelson (1977) to the case in which there is a continuum of goods and demand is present (on the problem of demand see Sect. 2.5). We treat here the essence of this model; where necessary, the original notation has been changed to conform it with the notation used in the present book.

18.3.1 Supply

In the case of a continuum of goods, we can index commodities on an interval, say [0,1], such that a commodity $0 \leq z \leq 1$ and its unit labour requirement $a_i(z)$ are associated with each point on the interval, where $i = 1, 2$ denotes the country. The relative unit labour cost (comparative cost) is then defined for each commodity as

$$A(z) = \frac{a_2(z)}{a_1(z)}, \qquad A(0) \geq A(z) \geq A(1), \qquad A'(z) < 0. \tag{18.6}$$

The assumption $A'(z) < 0$ means that commodities are ranked in order of monotonically diminishing country 1's comparative advantage, and is the continuous equivalent of the ranking performed in the discrete m-goods case, see Eq. (2.16).

The price equations (2.18) become

$$p_1(z) = a_1(z)w_1, \qquad p_2(z) = a_2(z)w_2, \tag{18.7}$$

where w_i are the wage rates in the two countries. Any commodity z will be produced in country 1 if $p_1(z) \leq p_2(z)$, or

$$a_1(z)w_1 \leq a_2(z)w_2, \tag{18.8}$$

namely

$$\omega \leq A(z), \tag{18.9}$$

where $\omega \equiv w_1/w_2$ is the relative wage rate. Similarly, for those commodities produced by country 2 we shall have

$$\omega \geq A(z). \tag{18.10}$$

For international trade to be possible it is obviously required that $A(0) < \omega < A(1)$, namely ω must lie within the two extreme comparative costs. Because of the continuity assumption there will exist a borderline good, say $\tilde{z} = \tilde{z}(\omega)$, such that country 1 will produce all commodities in the interval

$$0 \leq z \leq \tilde{z}(\omega) \tag{18.11}$$

more efficiently than country 2, while all commodities in the interval

$$\tilde{z}(\omega) \leq z \leq 1 \tag{18.12}$$

will be produced more efficiently by country 2. If we take either (18.9) or (18.11) with equality we have

$$\omega = A(\tilde{z}), \tag{18.13}$$

hence we can determine the borderline commodity as

$$\tilde{z} = A^{-1}(\omega), \tag{18.14}$$

where $A^{-1}(\omega)$ is the inverse function of $A(z)$. In other words, if we except the borderline commodity, country 1 has a comparative advantage in the commodities defined by (18.11) while country 2 has a comparative advantage in the commodities defined by (18.12).

Thus the supply side is entirely summarized by the relation (either (18.13) or (18.14): since $A'(z) < 0$, there is a one-to-one correspondence) between the relative wage rate and the borderline commodity that characterizes the efficient geographic specialization.

18.3.2 Demand

Identical tastes with homothetic utility functions are assumed. This assumption (which will also be used in the Heckscher-Ohlin model, see Chap. 4) means that the structure of demand is identical in both countries and independent of the level of income.

In the many-commodity case we would have

$$\begin{aligned} b_{1i} &= P_{1i}C_{1i}/Y_1,\ b_{2i} = b_{1i}, \\ \textstyle\sum_1^m b_{1i} &= 1. \end{aligned} \tag{18.15}$$

where Y is total income, P_i the price of good i, C_i the demand for good i, and b_i the budget share of commodity i. By analogy, in the continuum-of-goods case we have

$$b_1(z) = P_1(z)C(z)/Y > 0,\ b_2(z) = b_1(z),$$
$$\int_0^1 b_1(z)\mathrm{d}z = 1. \tag{18.16}$$

We now define the fraction of income spent in each country (and hence in the world) on those goods for which country 1 has a comparative advantage:

$$\vartheta(\tilde{z}) = \int_0^{\tilde{z}} b(z)\mathrm{d}z > 0,$$
$$\vartheta'(\tilde{z}) = b(\tilde{z}) > 0, \tag{18.17}$$

where $(0, \tilde{z})$ denotes the range of commodities that are produced more efficiently by country 1 (see the previous section). It follows that

$$1 - \vartheta(\tilde{z}) = \int_{\tilde{z}}^1 b(z)\mathrm{d}z \tag{18.18}$$

is the fraction of world income spent on the commodities for which country 2 enjoys a comparative advantage.

We have seen in the previous section that the supply side can be represented by a relation between $\tilde{z}$ and ω, the (monotonically) decreasing $A(\tilde{z})$ schedule. It is convenient to represent the demand side by a similar relation. For this purpose we begin by observing that, in any perfectly competitive setting, factors are fully employed, and the value of output coincides with total factor rewards (total income), which in turn is entirely spent in buying commodities. Thus, in our one-factor setting, total income coincides with labour income,

$$Y_i = w_i L_i, i = 1, 2. \tag{18.19}$$

Equilibrium in the market for the commodities produced by country 1 requires that the value of country 1's output, $w_1 L_1$, equals the value of world demand for country 1's products:

$$w_1 L_1 = \vartheta(\tilde{z})(w_1 L_1 + w_2 L_2). \tag{18.20}$$

If we divide through by $w_1 L_1$ and rearrange terms, we obtain

$$1 - \vartheta(\tilde{z}) = \vartheta(\tilde{z}) \frac{w_2}{w_1} \frac{L_2}{L_1},$$

from which

$$\omega = \frac{\vartheta(\tilde{z})}{1-\vartheta(\tilde{z})}\frac{L_2}{L_1} = B(\tilde{z}; L_2/L_1). \tag{18.21}$$

Thus the market clearing condition (18.20) enables us to associate with each $\tilde{z}$ a value of the relative wage rate w_1/w_2 such that market equilibrium obtains. Let us observe that the B schedule starts from the origin (when $\tilde{z} = 0, \vartheta(\tilde{z}) = 0$ by definition) and is monotonically increasing, since $\vartheta'(\tilde{z}) > 0$ as shown in (18.17).

Equation (18.21) summarizes the demand side.

18.3.3 Determination of the Relative Wage Rate and Specialization

On the supply side we have determined the borderline commodity $\tilde{z}$ and the associated hypothetical pattern of specialization by taking the relative wage rate ω as given. On the demand side we have determined the pattern of world demand by taking $\tilde{z}$ as given. Both variables can be simultaneously determined by considering the system formed by Eqs. (18.21) and (18.13) in the space $(\omega, \tilde{z})$. Since $A(\tilde{z})$ is a positive and monotonically decreasing function, while $B(\tilde{z}; L_2/L_1)$ is a positive and monotonically increasing function that starts from the origin, there will be one, and only one, positive solution.

In this way we simultaneously determine the relative wage rate and the borderline commodity, namely the unique relative wage rate $\overline{\omega}$ and borderline commodity $\overline{z}$ such that country 1 will specialize in (and export) all commodities in the range $0 \leq z < \overline{z}$, while country 2 will specialize in (and export) all commodities in the range $\overline{z} < z \leq 1$.

Given the price equations (18.7), the relative price of a commodity z produced in country 1 in terms of any commodity z'' produced in country 2 (i.e., the terms of trade) is easily determined as

$$p_1(z)/p_2(z'') = \overline{\omega} a_1(z)/a_2(z''). \tag{18.22}$$

Hence the terms of trade are endogenously determined as well.

All variables in this model are jointly determined by technology (the technical coefficients), relative factor endowments (as measured by the relative labour force), and tastes. In this sense it can be said that the present model (apart from the assumption of continuity, that merely serves to simplify the analysis) is a particular case of the general neoclassical model.

Finally, note that balance-of trade-equilibrium obtains in this model as well (see Eq. (2.8) for the original Ricardian case). In fact, Eq. (18.20) can be written in the alternative form

$$\vartheta(\tilde{z})w_2L_2 = [1 - \vartheta(\tilde{z})]w_1L_1, \tag{18.23}$$

which means that the value of country 2's demand for country 1's commodities (country 1's exports) equals the value of country 1's demand for country 2's commodities (country 1's imports).

18.3.4 Trade Policy

Let us assume that each country levies a uniform tariff rate, d_1 and d_2, on all imports. The proceeds are rebated in lump sum form.

The price of country 2's goods in country 1, and of country 1's goods in country 2 will be

$$\begin{aligned}(1 + d_1)p_2(z) &= (1 + d_1)a_2(z)w_2,\\ (1 + d_2)p_1(z) &= (1 + d_2)a_1(z)w_1,\end{aligned} \tag{18.24}$$

respectively. Hence any commodity z will be produced in country 1 if

$$a_1(z)w_1 \leq (1 + d_1)a_2(z)w_2, \tag{18.25}$$

namely

$$\frac{\omega}{1 + d_1} \leq A(z). \tag{18.26}$$

Similarly, for those commodities produced by country 2 we shall have

$$(1 + d_2)\omega \geq A(z). \tag{18.27}$$

It follows that there will now be *two* equilibrium borderline commodities, one from the point of view of country 1, the other from the point of view of country 2, between which there will be no trade. In fact, from (18.26) and (18.27) these commodities are easily seen to be

$$\begin{aligned}\bar{z}_1 &= A^{-1}(\frac{\omega}{1 + d_1}),\\ \bar{z}_2 &= A^{-1}\left(\omega(1 + d_2)\right),\end{aligned} \tag{18.28}$$

where $\bar{z}_2 < \bar{z}_1$, since the function $A^{-1}(\cdot)$ is monotonically decreasing, and $\omega(1 + d_2) > \omega/(1 + d_1)$. It is apparent that the presence of a tariff in either country (or in both) gives rise to a range of *nontraded* commodities, which are those lying in the interval between $\bar{z}_2$ and $\bar{z}_1$. Of course, equilibrium $\bar{z}_1$ and $\bar{z}_2$ are yet to be determined by the interaction between demand and technology, to which we now turn.

First, let us define variables $\lambda_1 = \lambda_1(\omega, d_1)$, $\lambda_2 = \lambda_2(\omega, d_2)$ as the fraction of country i's income spent on goods produced by the same country. Due to the range of nontraded commodities, we have

$$\begin{aligned} \lambda_1(\omega, d_1) &= \int_0^{\bar{z}_1} b(z)\mathrm{d}z, \\ \lambda_2(\omega, d_2) &= \int_{\bar{z}_2}^{1} b(z)\mathrm{d}z, \end{aligned} \tag{18.29}$$

where the arguments of λ_1, λ_2 derive from the fact that a definite integral is a function of its limits of integration. In our case these limits are, in turn, functions of the arguments given in (18.28).

Next we consider the trade balance equilibrium condition at international prices, which is no longer (18.23), but

$$(1-\lambda_1)Y_1/(1+d_1) = (1-\lambda_2)Y_2/(1+d_2), \tag{18.30}$$

where Y_i now includes lump-sum tariff rebates R_i, namely

$$\begin{aligned} Y_i &= w_i L_i + R_i \\ &= w_i L_i + d_i[(1-\lambda_i)Y_i/(1+d_i)] \end{aligned} \tag{18.31}$$

since the lump-sum tariff rebate equals the tariff rate times the fraction of income spent on imports. Hence by solving (18.31) for Y_i we have the expression

$$Y_i = w_i L_i(1+d_i)/(1+\lambda_i d_i). \tag{18.32}$$

From Eqs. (18.32) and (18.30) we get

$$\omega = \frac{(1-\lambda_2)}{(1-\lambda_1)}\frac{(1+\lambda_1 d_1)}{(1+\lambda_2 d_2)}\frac{(1+d_2)}{(1+d_1)}\frac{L_2}{L_1}, \tag{18.33}$$

where λ_1 and λ_2 are given by Eqs. (18.29). It follows that Eq. (18.33) defines an implicit relation between the four variables $\omega, d_1, d_2, L_2/L_1$, that can be solved for $\overline{\omega}$ (the equilibrium relative wage rate) in terms of the other variables, yielding

$$\overline{\omega} = \overline{\omega}(L_2/L_1, d_1, d_2). \tag{18.34}$$

Equation (18.34) can be used to perform comparative-statics exercises, in particular to check the effects of tariffs. For example, by applying the chain rule to Eq. (18.33), account being taken of Eq. (18.34), we get

$$\begin{aligned}\frac{\partial\overline{\omega}}{\partial d_1} = \Delta\Bigg\{&\left[\left(\tfrac{\partial\lambda_1}{\partial\overline{\omega}}\tfrac{\partial\overline{\omega}}{\partial d_1}+\tfrac{\partial\lambda_1}{\partial d_1}\right)d_1+\lambda_1\right][(1-\lambda_2)(1+d_2)L_2]\\ &\qquad\qquad\qquad\times[(1-\lambda_1)(1+\lambda_2 d_2)(1+d_1)L_1]\\ &+\left[\left(\tfrac{\partial\lambda_1}{\partial\overline{\omega}}\tfrac{\partial\overline{\omega}}{\partial d_1}+\tfrac{\partial\lambda_1}{\partial d_1}\right)(1+\lambda_2 d_2)(1+d_1)L_1+(1-\lambda_1)(1+\lambda_2 d_2)\right]\\ &\qquad\qquad\qquad\times[(1-\lambda_2)(1+\lambda_1 d_1)(1+d_2)L_2]\Bigg\},\end{aligned} \tag{18.35}$$

where

$$\Delta \equiv [(1-\lambda_1)(1+\lambda_2 d_2)(1+d_1)L_1]^{-2} > 0. \tag{18.36}$$

Equation (18.35) can be written as

$$\frac{\partial\overline{\omega}}{\partial d_1} = \left(\frac{\partial\lambda_1}{\partial\overline{\omega}}\frac{\partial\overline{\omega}}{\partial d_1}+\frac{\partial\lambda_1}{\partial d_1}\right)H_1 + H_2, \tag{18.37}$$

where

$$\begin{aligned}H_1 \equiv{}& \Delta\{d_1(1-\lambda_2)(1+d_2)L_2[(1-\lambda_1)(1+\lambda_2 d_2)(1+d_1)L_1]\\ &+[(1+\lambda_2 d_2)(1+d_1)L_1+(1-\lambda_1)(1+\lambda_2 d_2)][(1-\lambda_2)(1+\lambda_1 d_1)(1+d_2)L_2]\},\\ H_2 \equiv{}& \Delta\{\lambda_1(1-\lambda_2)(1+d_2)L_2[(1-\lambda_1)(1+\lambda_2 d_2)(1+d_1)L_1]\}.\end{aligned} \tag{18.38}$$

It can readily be checked that H_1, H_2 are both positive. From Eq. (18.37) we have

$$\frac{\partial\overline{\omega}}{\partial d_1} = \left(1 - H_1\frac{\partial\lambda_1}{\partial\overline{\omega}}\right)^{-1}\left(H_1\frac{\partial\lambda_1}{\partial d_1}+H_2\right). \tag{18.39}$$

To determine $\partial\lambda_1/\partial\overline{\omega}, \partial\lambda_1/\partial d_1$ we employ Eqs. (18.28) and (18.29). Using the rule for the differentiation of a definite integral with respect to a parameter, we have

$$\begin{aligned}\frac{\partial\lambda_1}{\partial\overline{\omega}} &= \left[\frac{\mathrm{d}}{\mathrm{d}\overline{z}_1}\int_0^{\overline{z}_1} b(z)\mathrm{d}z\right]\frac{\partial\overline{z}_1}{\partial\overline{\omega}} = b(\overline{z}_1)\frac{\partial\overline{z}_1}{\partial\overline{\omega}} < 0,\\ \frac{\partial\lambda_1}{\partial d_1} &= \left[\frac{\mathrm{d}}{\mathrm{d}\overline{z}_1}\int_0^{\overline{z}_1} b(z)\mathrm{d}z\right]\frac{\partial\overline{z}_1}{\partial d_1} = b(\overline{z}_1)\frac{\partial\overline{z}_1}{\partial d_1} > 0.\end{aligned} \tag{18.40}$$

To prove the signs stated in (18.40), we begin by observing that, as the function $A^{-1}(\cdot)$ is monotonically decreasing, from (18.28) it follows that $\partial\overline{z}_1/\partial\overline{\omega} < 0, \partial\overline{z}_1/\partial d_1 > 0$. We next recall that $b(\overline{z}_1) > 0$. Putting these results together proves the signs.

From these signs and the positivity of H_1, H_2, it follows that both expressions on the r.h.s. of Eq. (18.39) are positive, hence

$$\frac{\partial \overline{\omega}}{\partial d_1} > 0. \tag{18.41}$$

Thus an increase in country 1's tariff rate (or the imposition of a small tariff starting from a free trade situation) improves the imposing country's equilibrium relative wage rate and, of course, deteriorates that of the other country.

It should now be noted that, given the assumption of identical homothetic tastes across countries and no distortions, the relative wage rate $\overline{\omega}$ is a measure of the relative welfare of the representative consumer in country 1 with respect to the representative consumer in country 2. Thus we get the well-known result that the imposition of a small tariff (or a small increase in an existing tariff) is welfare-improving for the imposing country, provided of course that the other country remains a free trader.

For further analysis of this model see Dornbusch et al. (1977) and Wilson (1980).

18.4 On the Determination of the Terms of Trade

Contrary to received opinion, Negishi (1982) maintains that the *original* Ricardian theory can determine the terms of trade without any recourse to demand factors. He argues his thesis with several well-chosen textual references to Ricardo's *Principles*, which for lack of space we cannot reproduce here; we shall therefore give only the references to the pages where the reader can find the passages quoted by Negishi (1982; we have changed the page numbers to conform with Sraffa's edition of Ricardo's *Works and Correspondence*, and we have also changed Negishi's notation to conform with that adopted in this chapter). Negishi's thesis is based on the following points:

1. The money wage rate tends to the "natural" price of labour or subsistence wage rate; this is the value of the basket of commodities which enables workers to subsist and perpetuate without either increase or diminution (Ricardo, pp. 60–61; Negishi, p. 204). In our two-commodity model (A = cloth, B = wine) we have

$$w = c_A p_A + c_B p_B, \tag{18.42}$$

   where c_A and c_B are the given quantities of the commodities which make up the basket of the subsistence wage rate. Lacking indications to the contrary, Negishi assumes that c_A and c_B are identical in the two countries.
2. In the original Ricardian example (Ricardo, pp. 93–94), the coefficients representing production costs are $a_1 = 100$, $b_1 = 120, a_2 = 90, b_2 = 80$: in

other words, labour productivity in *both* the cloth and the wine industries is lower in England than in Portugal. This assumption, which seemed strange to many economists (Portugal was, in fact, a less developed country than England), is consistent with Ricardo's assumption that labour productivity is lower in more advanced countries (Ricardo, p. 93; Negishi, p. 201 and 205). Such an assumption is, however, secondary for our purpose, because what is important is the existence of a difference in comparative costs.

3. Capital consists entirely of the wage-bill; in other words, it is solely circulating capital, which takes 1 year everywhere to be re-integrated. This is a common simplifying assumption, also adopted by Negishi (p. 205); its usefulness lies in the fact that it allows us to avoid the problems arising in the Ricardian labour theory of value, when fixed capital is present. Consequently, the price of a commodity will be given by the wage bill advanced, plus profit earned on it. Thus in England (country 1) we have

$$\begin{aligned} p_{1A} &= (1+r_1)\,100w_1 = (1+r_1)\,100\,(c_A p_{1A} + c_B p_{1B})\,, \\ p_{1B} &= (1+r_1)120w_1 = (1+r_1)120\,(c_A p_{1A} + c_B p_{1B})\,, \end{aligned} \tag{18.43}$$

where r_1 is the rate of profit. If we multiply the first equation by c_A, the second by c_B, and add them, we get

$$\begin{aligned} c_A p_{1A} + c_B p_{1B} = c_A\,(1+r_1)\,100\,(c_A p_{1A} + c_B p_{1B}) \\ +c_B\,(1+r_1)\,120\,(c_A p_{1A} + c_B p_{1B})\,. \end{aligned}$$

By manipulating this equation and solving it for $1/(1+r_1)$ we obtain

$$1/\,(1+r_1) = 100c_A + 120c_B, \tag{18.44}$$

whence we see that the rate of profit is determined, once we know the composition of the wage rate and the labour input coefficients. Similarly we have, for Portugal,

$$\begin{aligned} p_{2A} &= (1+r_2)\,90w_2 = (1+r_2)\,90\,(c_A p_{2A} + c_B p_{2B})\,, \\ p_{2B} &= (1+r_2)\,80w_2 = (1+r_2)\,80\,(c_A p_{2A} + c_B p_{2B})\,, \end{aligned} \tag{18.45}$$

whence, by the same procedure as for England, we obtain

$$1/\,(1+r_2) = 90c_A + 80c_B. \tag{18.46}$$

As can be seen from (18.44) and (18.46), the rate of profit is lower in England than in Portugal.

Once international trade is opened, the prices of the same commodity are equalized in the two countries (it should be remembered that we are assuming a common monetary unit or, which amounts to the same thing, that the fixed exchange rate is set to one) and, consequently, the money wage rates are equalized (this depends on the assumption that c_A and c_B are the same in the two countries). As we know, England will certainly produce cloth and Portugal wine, so that the prices of the two commodities will be

$$\begin{aligned} p_A &= \left(1 + r_1^c\right) 100\,(c_A p_A + c_B p_B)\,, \\ p_B &= \left(1 + r_2^c\right) 80\,(c_A p_A + c_B p_B)\,, \end{aligned} \tag{18.47}$$

where r_1^c, r_2^c denote the rates of profit obtaining in the two countries *after* trade begins.

We do not know whether wine is still produced in England and cloth in Portugal: this depends on the inequalities

$$\begin{aligned} p_A &\le \left(1 + r_2^c\right) 90\,(c_A p_A + c_B p_B)\,, \\ p_B &\le \left(1 + r_1^c\right) 120\,(c_A p_A + c_B p_B)\,, \end{aligned} \tag{18.48}$$

where the strict inequality implies complete specialization, which is the natural outcome of the Ricardian model.

Let us now observe that from Eqs. (18.47), we can obtain a mathematical relationship between the rates of profit in the two countries after trade. In fact, if we denote $1/(1 + r_i^c)$ by $R_i, i = 1, 2$, and apply the usual procedure (multiply the first equation by c_A etc.) we get

$$R_1 R_2 - 100 c_A R_2 - 80 c_B R_1 = 0. \tag{18.49}$$

At this point we must examine the economic relationship between r_1^c and r_2^c after international trade begins. In the presence of perfect international capital mobility r_1^c and r_2^c should be equalized, but this is not so for Ricardo; thus we are led to Negishi's fourth point.

4. Risk, uncertainty, disinclination to quit the country of one's birth and connections, etc., lead home capitalists to be satisfied with a rate of profit lower than that in the foreign country, whence $r_1^c < r_2^c$ (Ricardo, pp. 94–95; Negishi, pp. 200–201 and 207). Let us note that this is true if the home country is England. But the same could be said of Portugal, and since it cannot be simultaneously true that $r_1^c < r_2^c$ *and* $r_2^c < r_1^c$, we are left with a problem. However, Ricardo—as can be seen by a careful reading of chap. VII of his *Principles*—had in mind a many-country world and, in any case, always reasoned from the point of view of England. We can also note that, if complete specialization does not obtain, then from Eqs. (18.47) and (18.48) (the latter are now satisfied with the equality sign), it immediately follows that $r_1^c < r_2^c$. Now, even in the case of complete specialization (which, it must be remembered, is not an instantaneous event, but a dynamic process), there will exist a time interval in which specialization is not complete and so $r_1^c < r_2^c$; it is then reasonable to assume that, when the two countries are completely specialized, the rates of profit, being those that exist historically, will still satisfy this inequality.

The importance of this point is crucial, so we give Ricardo's passage verbatim (Ricardo, 1817, p. 95):

> Experience, however, shows that the fancied or real insecurity of capital, when not under the immediate control of its owner, together with the natural disinclination which man has to quit the country of his birth and connections, and intrust himself; with all his habits fixed, to a strange government and new laws, check the emigration of capital. These feelings, which I should be sorry to see weakened, induce most men of property to be satisfied with a low rate of profits in their country, rather than seek a more advantageous employment for their wealth in a foreign nation.

These factors determine, according to Negishi, the amount of the difference existing, after trade begins, between the domestic (i.e. England's) rate of profit and the foreign (i.e. Portugal's) rate of profit; this difference can be expressed as a rate of conversion between the two profit rates, that is, by using the auxiliary variables R_i,

$$\frac{R_2}{R_1} = a, \tag{18.50}$$

where $a < 1$ by assumption, and $(1 - a)$ is a (proportional) risk coefficient for investment of capital abroad.

Equations (18.49) and (18.50) constitute a system of two equations to determine the two unknowns R_1, R_2 and, consequently, the profit rates r_1^c, r_2^c. By substituting R_2 from Eq. (18.50) into Eq. (18.49) we get

$$R_1 = 100c_A + (80/a)\,c_B, \tag{18.51}$$

whence we can easily obtain r_1^c etc.

Having thus determined the rates of profit, the terms of trade p_A/p_B can be immediately derived from Eqs. (18.47), account being taken of Eqs. (18.50),[2]

$$\frac{p_A}{p_B} = \frac{1 + r_1^c}{1 + r_2^c}\frac{100}{80} = a\frac{100}{80}. \tag{18.52}$$

It can easily be checked that Eq. (18.52) yields a value of the terms of trade included between the comparative costs. In fact, if we substitute Eq. (18.50) into Eqs. (18.47) and (18.48), we see that the admissible interval for a to satisfy them is

$$80/120 \leq a \leq 90/100, \tag{18.53}$$

[2]It should be noted that if one is interested exclusively in the terms of trade, it is sufficient to divide the first equation of (18.48) by the second and to make use of Eq. (18.50), with no need to determine r_1^c and r_2^c: what matters for this purpose is, in fact, the given relation between the two profit rates, not their actual values.

where the strict inequalities hold when specialization is complete. From Eqs. (18.52) and (18.53) it immediately follows that

$$\frac{100}{120} \leq \frac{p_A}{p_B} \leq \frac{90}{80}, \tag{18.54}$$

where the two extremes are the comparative costs. The example in Ricardo (who used the value 1 for the terms of trade) corresponds to the case $a = 0.8$.

Finally, it is interesting to note that—if we compare Eqs. (18.44), (18.46) and (18.51)—international trade in any case does *not* cause a decrease in the rate of profit with respect to the situation before trade, and certainly causes an increase in it in the case of strict inequality, which explains the tendency towards complete specialization.

It must be emphasized, in conclusion, that the *deus ex machina* of this ingenious reconstruction of the Ricardian model is Eq. (18.50): once a is known, in fact, everything can be determined. And, if a is not known exactly, but only the interval (18.53) is known, then we are in the same situation of indeterminacy as before, with the sole difference that we substitute one interval of ignorance for the other, that is, the interval in which the (indeterminate) rate of conversion between the two profit rates must lie for the interval in which the (indeterminate) terms of trade must fall. Further studies are therefore necessary to establish if and how it is possible to use the general indications contained in Ricardo's passage cited above to determine the value of a and so the value of the terms of trade. One way might be that of assuming that a is, in any given period, a historico-institutional datum (that is, having the same nature as the composition of the "subsistence" wage rate). In this way the problem would be radically solved, by simply saying that a is exogenously given like c_A and c_B. Further studies are also required to include in the model the case in which fixed capital is present.

It is however clear that by following this path it will be possible to satisfactorily determine the terms of trade within the context of the Ricardian model, without introducing demand factors.

References

Chipman, J. S. (1965). A survey of the theory of international trade, part 1, the classical theory.

Dorfman, R., Samuelson, P. A., & Solow, R. M. (1958). *Linear programming and economic analysis*.

Dornbusch, R., Fischer, S., & Samuelson, P. A. (1977). Comparative advantage, trade, and payments in a ricardian model with a continuum of goods.

Hartwick, J. M. (1979). Distribution of world income in the Ricardo-Mill model of international trade.

Jones, R. W. (1961). comparative advantage and the theory of tariffs: A multi-country multi-commodity model.

McKenzie, L. (1954a). On equilibrium in Graham's model of world trade and other competitive systems.

McKenzie, L. (1954b). Specialization and efficiency in world production.
McKenzie, L. (1955). Specialization in production and the production possibility locus.
Negishi, T. (1982). The labour theory of value in the Ricardian theory of international trade.
Ricardo, D. (1817), *On the principles of political economy and taxation* (chap. VII).
Takayama, A. (1972). *International trade: An approach to the theory* (Pt. II).
Whitin, T. M. (1953). Classical theory, Graham's theory, and linear programming.
Wilson, C. (1980). On the general structure of Ricardian models with a continuum of goods: Applications to growth, tariffs and technical change.

Chapter 19
Appendix to Chapter 3

19.1 The Transformation Curve and the Box Diagram

19.1.1 *A Formal Derivation of the Transformation Curve and Its Properties*

Let $A = f_A(L_A, K_A)$ and $B = f_B(L_B, K_B)$ be the aggregate production functions (twice continuously differentiable) of the two commodities, where $L_A + L_B = L$ (the total amount of labour existing in the economy) and $K_A + K_B = K$ (the total amount of capital existing in the economy). Consider then the following maximization problem

$$\text{Max } B = f_B(L_B, K_B) \text{ subject to } f_A(L - L_B, K - K_B) - A = 0, \tag{19.1}$$

where A is a *parametrically given* amount of commodity A.[1] The Lagrangian is

$$F = f_B(L_B, K_B) + \lambda \left[f_A (L - L_B, K - K_B) - A \right],$$

and the first-order conditions for a maximum are

$$\begin{aligned}
\frac{\partial F}{\partial L_B} &= \frac{\partial f_B}{\partial L_B} - \lambda \frac{\partial f_A}{\partial L_A} = 0, \\
\frac{\partial F}{\partial K_B} &= \frac{\partial f_B}{\partial K_B} - \lambda \frac{\partial f_A}{\partial K_A} = 0, \\
\frac{\partial F}{\partial \lambda} &= f_A (L - L_B, K - K_A) - A = 0,
\end{aligned} \tag{19.2}$$

[1]Here, as everywhere in this book, we follow the commonly adopted practice of using the same symbol to denote both the commodity and its quantity.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_19,

whence

$$\frac{\partial f_B/\partial L_B}{\partial f_B/\partial K_B} = \frac{\partial f_A/\partial L_A}{\partial f_A/\partial K_A}, \tag{19.3}$$

which states that the MRTS have to be equal, namely that the isoquants in the Edgeworth-Bowley box must be tangent.

The second-order conditions require the following bordered Hessian determinant

$$\begin{vmatrix} \frac{\partial^2 f_B}{\partial L_B^2} + \lambda \frac{\partial^2 f_A}{\partial L_A^2} & \frac{\partial^2 f_B}{\partial L_B \partial K_B} + \lambda \frac{\partial^2 f_A}{\partial L_A \partial K_A} & -\frac{\partial f_A}{\partial L_A} \\ \frac{\partial^2 f_B}{\partial K_B \partial L_B} + \lambda \frac{\partial^2 f_A}{\partial K_A \partial L_A} & \frac{\partial^2 f_B}{\partial K_B^2} + \lambda \frac{\partial^2 f_A}{\partial K_A^2} & -\frac{\partial f_A}{\partial K_A} \\ -\frac{\partial f_A}{\partial L_A} & -\frac{\partial f_A}{\partial K_A} & 0 \end{vmatrix} \tag{19.4}$$

to be positive. We assume that this condition is satisfied.

The marginal rate of transformation is

$$-\frac{\mathrm{d}B}{\mathrm{d}A} = -\frac{(\partial f_B/\partial L_B)\,\mathrm{d}L_B + (\partial f_B/\partial K_A)\,\mathrm{d}K_B}{(\partial f_A/\partial L_A)\,\mathrm{d}L_A + (\partial f_A/\partial K_A)\,\mathrm{d}K_A}, \tag{19.5}$$

where of course the two total differentials must obey the optimum conditions (19.3) as well as the constraints $\mathrm{d}L_A + \mathrm{d}L_B = 0$, $\mathrm{d}K_A + \mathrm{d}K_B = 0$. Therefore

$$\begin{aligned} -\frac{\mathrm{d}B}{\mathrm{d}A} &= -\frac{(\partial f_B/\partial K_B)\,\{[(\partial f_B/\partial L_B)/\,(\partial f_B/\partial K_B)]\,\mathrm{d}L_B + \mathrm{d}K_B\}}{(\partial f_A/\partial K_A)\,\{[(\partial f_A/\partial L_A)\,/\,(\partial f_A/\partial K_A)]\,\mathrm{d}L_A + \mathrm{d}K_A\}} \\ &= \frac{\partial f_B/\partial K_B}{\partial f_A/\partial K_A} = \frac{\partial f_B/\partial L_B}{\partial f_A/\partial L_A}. \end{aligned} \tag{19.6}$$

Consider now the marginal costs of A and B, MC_A and MC_B. We know from microeconomics that

$$\begin{aligned} MC_A &= \frac{p_L}{\partial f_A/\partial L_A} = \frac{p_K}{\partial f_A/\partial K_A}, \\ MC_B &= \frac{p_L}{\partial f_B/\partial L_B} = \frac{p_K}{\partial f_B/\partial K_B}, \end{aligned} \tag{19.7}$$

where p_L and p_K are factor prices, and that—under perfect competition—

$$MC_A = p_A, \;\; MC_B = p_B. \tag{19.8}$$

From (19.7) and (19.6) it follows that

$$-\frac{\mathrm{d}B}{\mathrm{d}A} = \frac{MC_A}{MC_B}, \tag{19.9}$$

and from (19.8) and (19.9) we have

$$-\frac{\mathrm{d}B}{\mathrm{d}A} = \frac{p_A}{p_B}. \tag{19.10}$$

These results have been commented on in Sect. 3.1. It goes without saying that the marginal rate of transformation could also be defined as $-(\mathrm{d}A/\mathrm{d}B)$, in which case the only modification is to take the reciprocal of the various fractions.

From (19.10) it follows that—provided the transformation curve is not linear and does not change its curvature—a unique point on the transformation curve corresponds to any price ratio p_B/p_A, so that the outputs of A and B can be considered as single-valued functions of this ratio.

19.1.2 Convexity or Concavity of the Transformation Curve

We now turn to examine the convexity or concavity of the transformation curve. From (19.2) and (19.6), we obtain

$$-\frac{\mathrm{d}B}{\mathrm{d}A} = \lambda, \tag{19.11}$$

in other words, the marginal rate of transformation is equal to the Lagrange multiplier. Therefore

$$\frac{\mathrm{d}^2B}{\mathrm{d}A^2} = -\frac{\partial\lambda}{\partial A}. \tag{19.12}$$

If we consider (19.2) as a set of implicit functions, we can express the variables L_B, K_B, λ as differentiable functions of the parameters A, L, K, provided that the Jacobian determinant of the set with respect to the variables is not zero at the equilibrium point (implicit function theorem: see, for example, Gandolfo, 2009, Pt. III, chap. 20). As is always the case in maximization problems, this Jacobian coincides with the Hessian involved in the second order conditions, and so the condition required by the implicit function theorem is automatically satisfied.

Therefore we can differentiate totally the first-order conditions with respect to A and obtain

$$\begin{aligned}
\left(\frac{\partial^2 f_B}{\partial L_B^2} + \lambda\frac{\partial^2 f_A}{\partial L_A^2}\right)\frac{\partial L_B}{\partial A} + \left(\frac{\partial^2 f_B}{\partial L_B \partial K_B} + \lambda\frac{\partial^2 f_A}{\partial L_A \partial K_A}\right)\frac{\partial K_B}{\partial A} - \frac{\partial f_A}{\partial L_A}\frac{\partial\lambda}{\partial A} &= 0,\\
\left(\frac{\partial^2 f_B}{\partial K_B L_B} + \lambda\frac{\partial^2 f_A}{\partial K_A L_A}\right)\frac{\partial L_B}{\partial A} + \left(\frac{\partial^2 f_B}{\partial K_B^2} + \lambda\frac{\partial^2 f_A}{\partial K_A^2}\right)\frac{\partial K_B}{\partial A} - \frac{\partial f_A}{\partial K_A}\frac{\partial\lambda}{\partial A} &= 0,\\
-\frac{\partial f_A}{\partial L_A}\frac{\partial L_B}{\partial A} - \frac{\partial f_A}{\partial K_A}\frac{\partial K_B}{\partial A} &= 1,
\end{aligned} \tag{19.13}$$

whence, solving for $\partial\lambda/\partial A$,

$$\frac{\partial\lambda}{\partial A}=\frac{\left(\frac{\partial^2 f_B}{\partial L_B^2}+\lambda\frac{\partial^2 f_A}{\partial L_A^2}\right)\left(\frac{\partial^2 f_B}{\partial K_B^2}+\lambda\frac{\partial^2 f_A}{\partial K_A^2}\right)-\left(\frac{\partial^2 f_B}{\partial L_B\partial K_B}+\lambda\frac{\partial^2 f_A}{\partial L_A\partial K_A}\right)^2}{|H|}, \tag{19.14}$$

where $|H|$ is the Hessian determinant (19.4). Since $|H|$ is positive, the sign of $\partial\lambda/\partial A$ depends on the sign of the numerator of the fraction. In general, this sign is indeterminate, even if we assume that marginal productivities are positive and decreasing. In fact, since $\lambda > 0$ by (19.2), the first two terms in the numerator of (19.14) are negative, so that their product is positive, and from this we must subtract the third term, which is also positive, if we rule out the exceptional case of its being zero.

19.1.3 Homogeneous Production Functions and Transformation Curve

This indeterminacy can be eliminated if we assume that the production functions are homogeneous of the first degree (constant returns to scale).

Since the properties of these functions are widely used in this and following chapters, we list them here for the reader's convenience, though they are well-known.

Given the twice-differentiable function

$$Y=F\left(X_1,X_2\right),$$

it is said to be (positively) homogeneous of degree $n>0$ if, for any $\mu>0$,

$$F\left(\mu X_1,\mu X_2\right)=\mu^n F\left(X_1,X_2\right).$$

It is easy to see that returns to scale are increasing, constant, decreasing, respectively for $n \gtreqless 1$. Let us consider the case of a homogeneous function of the first degree. Such a function has the following properties:

(*a*) *Intensive form:* $Y = X_2F(X_1/X_2,1)=X_2 f(X_1/X_2)$,
$Y = X_1F(1,X_2/X_1)=X_1 g(X_2/X_1)$,

which allows us, for example, to express output per head as a function of capital per head only. If we let X_1 denote capital, X_2 labour, y output per head, and ρ capital per head, from *(a)* we have $y=F(\rho,1)$.

(*b*) *Radiality* :
$$\frac{\partial Y}{\partial X_1} = f'(X_1/X_2)=g(X_2/X_1)-(X_2/X_1)g'(X_2/X_1),$$
$$\frac{\partial Y}{\partial X_2} = f(X_1/X_2)-(X_1/X_2)f'(X_1/X_2)=g'(X_2/X_1),$$

which means that marginal productivities are functions of the input ratio alone, so that the isoquants have an identical slope along any ray starting from the origin.

(c) Euler's theorem: $\frac{\partial Y}{\partial X_1}X_1 + \frac{\partial Y}{\partial X_2}X_2 = Y$,
so that output is exhausted if—as is the case under perfect competition—each factor's reward in real terms equals its marginal productivity.

(d) Relations between second-order pure and mixed derivatives:

$$\frac{\partial^2 Y}{\partial X_1^2} = -\frac{X_2}{X_1}\frac{\partial^2 Y}{\partial X_1 \partial X_2}, \quad \frac{\partial^2 Y}{\partial X_2^2} = -\frac{X_1}{X_2}\frac{\partial^2 Y}{\partial X_1 \partial X_2}.$$

This property means that, if we assume $\partial^2 Y/\partial X_i^2 < 0$ (decreasing marginal productivities), then $\partial^2 Y/\partial X_1 \partial X_2 = \partial^2 Y/\partial X_2 X_1 > 0$, namely an increase in a factor has a *positive* effect on the marginal productivity of the *other* factor.

Let us now define what is meant by a "well-behaved" production function. Mathematically, it is a first-degree homogeneous function with positive first-order partial derivatives and negative second-order pure partial derivatives, that—with reference to the intensive form—further satisfies the Inada-Uzawa conditions:

$$\begin{aligned} &F(0,1) = 0, \quad \lim_{\rho\to\infty} F(\rho,1) = \infty, \\ &F'(0,1) = \infty, \lim_{\rho\to 0} F'(\rho,1) = \infty. \end{aligned}$$

Going back to our problem, property *(d)* enables us to write

$$\frac{\partial^2 f_i}{\partial L_i^2} = -\frac{K_i}{L_i}\frac{\partial^2 f_i}{\partial L_i \partial K_i}, \quad \frac{\partial^2 f_i}{\partial K_i^2} = -\frac{L_i}{K_i}\frac{\partial^2 f_i}{\partial L_i \partial K_i}, \quad i = A, B. \tag{19.15}$$

By using these relations we can rewrite the numerator of (19.14), after some manipulations, in the form

$$\begin{aligned} &\lambda \frac{\partial^2 f_A}{\partial L_A \partial K_A}\frac{\partial^2 f_B}{\partial L_B \partial K_A}\left(\frac{K_B L_A}{K_A L_B} + \frac{K_A L_B}{K_B L_A} - 2\right) \\ &= \lambda \frac{\partial^2 f_A}{\partial L_A \partial K_A}\frac{\partial^2 f_B}{\partial L_B \partial K_B}\frac{(K_B L_A - K_A L_B)^2}{K_A L_B K_B L_A} \\ &= \lambda \frac{\partial^2 f_A}{\partial L_A \partial K_A}\frac{\partial^2 f_B}{\partial L_B \partial K_B}\frac{(\varrho_B - \varrho_A)^2}{\varrho_A \varrho_B}, \end{aligned} \tag{19.16}$$

where $\varrho_A \equiv K_A/L_A, \varrho_B \equiv K_B/L_B$ are the factor intensities in the two sectors. Now, since we have assumed decreasing marginal productivities, it follows from (19.15) that the second order mixed partial derivative must be positive in both sectors, and so (19.16) is a *positive* quantity, barring out the exceptional case of $\varrho_A = \varrho_B$. Therefore, $\partial\lambda/\partial A \geq 0$ and consequently $\mathrm{d}^2 B/\mathrm{d}A^2 \leq 0$. This proves that, with constant returns to scale, the transformation curve is concave to the origin, except for the case of equal factor intensities in the two sectors (in which case it is a straight line).

It can also be shown that, when the production functions are homogeneous of degree higher than the first (increasing returns to scale), the numerator of (19.14) is equal to (19.16) *plus* several other terms; these terms have different signs so that the numerator under consideration can be either negative or positive as well as change its sign. Therefore, with increasing returns to scale, the transformation curve can be either convex or concave to the origin as well as change its curvature; for further details see Sect. 3.5 and Herberg (1969).

19.2 A Simple Closed Economy

19.2.1 The Basic Model

The following model is derived from Kemp (1964, 1969b):

$$\begin{aligned} A &= L_A g_A(\varrho_A), \\ B &= L_B g_B(\varrho_B), \\ g'_A &= p g'_B, \\ g_A - \varrho_A g'_A &= p\left(g_B - \varrho_B g'_B\right), \\ L_A + L_B &= L, \\ \varrho_A L_A + \varrho_B L_B &= K, \\ I_A &= A + pB, \\ A^D(I_A, p) &= A, \\ B^D(I_A, p) &= B. \end{aligned} \tag{19.17}$$

The first two equations are the production functions which, thanks to the assumption of first-degree homogeneity, can be written in the intensive form $L_i g_i(\varrho_i) \equiv L_i f_i(1, \varrho_i), \quad i = A, B$, where $\varrho_i \equiv K_i/L_i$ are factor intensities in the two sectors.

Since $g'_A \equiv \partial f_A/\partial K_A$, $g'_B \equiv \partial f_B/K_B$, the third equation states that the value of the marginal product of capital, measured in terms of commodity A taken as *numéraire* ($p \equiv p_B/p_A$), is equal in both sectors. In fact, under perfect competition the value of the marginal product of a factor must be equal to that factor's reward, which in turn must be equal in each sector. The fourth equation expresses the same condition for the marginal product of labour, since

$$g_A - \varrho_A g'_A \equiv \partial f_A/\partial L_A, \quad g_B - \varrho_B g'_B \equiv \partial f_B/\partial L_B.$$

The fifth and sixth equations state that both factors are fully employed; L and K are the given total amounts existing.

The seventh equation defines real aggregate income expressed in terms of the first commodity. The last two equations are the equilibrium conditions on the markets for the commodities; the aggregate demand for each commodity is assumed to depend on aggregate income and on the relative price.

The model has nine equations and only eight unknowns $(A, B, \varrho_A, \varrho_B, L_A, L_B, I_A, p)$. However, one of the two demand = supply equations is not independent, for either one may be derived from the other if we take account of the budget restraint $A + pB = A^D + pB^D$ (see Eq. (3.8), where the outputs of A and B have been called S_A and S_B).

19.2.2 The Supply Side of the Model

Let us now consider the subset consisting of the first six equations of the model, which define the supply side of the economy. We see that it includes seven unknowns $(A, B, \varrho_A, \varrho_B, L_A, L_B, p)$ so that—assuming that its Jacobian determinant with respect to the first six variables is not zero—we can use the implicit function theorem and express the first six variables as continuously differentiable function of the seventh (p); this proves rigorously that the supplies of A and B are functions of p, as shown heuristically in the text. It follows that I_A is ultimately a function of p only, and, consequently, that A^D and B^D can be expressed as general equilibrium functions of p only, as explained verbally in Sect. 3.2.2 (see also below, Sect. 19.2.3).

If we differentiate totally the first six equations with respect to p we obtain

$$\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}p} &= \frac{\mathrm{d}L_A}{\mathrm{d}p} g_A + L_A g_A' \frac{\mathrm{d}\varrho_A}{\mathrm{d}p},\\
\frac{\mathrm{d}B}{\mathrm{d}p} &= \frac{\mathrm{d}L_B}{\mathrm{d}p} g_B + L_B g_B' \frac{\mathrm{d}\varrho_B}{\mathrm{d}p},\\
g_A'' \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} &= g_B' + p g_B'' \frac{\mathrm{d}\varrho_B}{\mathrm{d}p},\\
-\varrho_A g_A'' \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} &= g_B - \varrho_B g_B' - p \varrho_B g_B'' \frac{\mathrm{d}\varrho_B}{\mathrm{d}p},\\
\frac{\mathrm{d}L_A}{\mathrm{d}p} + \frac{\mathrm{d}L_B}{\mathrm{d}p} &= 0,\\
\frac{\mathrm{d}\varrho_A}{\mathrm{d}p} L_A + \varrho_A \frac{\mathrm{d}L_A}{\mathrm{d}p} + \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} L_B + \varrho_B \frac{\mathrm{d}L_B}{\mathrm{d}p} &= 0,
\end{aligned} \tag{19.18}$$

from which we can compute the derivatives $\mathrm{d}A/\mathrm{d}p$, $\mathrm{d}B/\mathrm{d}p$, $\mathrm{d}\varrho_A/\mathrm{d}p$, $\mathrm{d}\varrho_B/\mathrm{d}p$, $\mathrm{d}L_A/\mathrm{d}p$, $\mathrm{d}L_B/\mathrm{d}p$. We are interested in

$$\begin{aligned}
\frac{\mathrm{d}A}{\mathrm{d}p} &= \frac{L_A g_B^2 p}{g_A'' (\varrho_B - \varrho_A)^2} + \frac{g_A^2 L_B}{p^2 g_B'' (\varrho_B - \varrho_A)^2},\\
\frac{\mathrm{d}B}{\mathrm{d}p} &= -\frac{L_A g_B^2}{g_A'' (\varrho_B - \varrho_A)^2} - \frac{g_A^2 L_B}{p^3 g_B'' (\varrho_B - \varrho_A)^2},
\end{aligned} \tag{19.19}$$

where of course $\varrho_B \neq \varrho_A$. Since g''_A and g''_B are negative by the assumption of decreasing marginal products, it follows from (19.19) that $\mathrm{d}A/\mathrm{d}p < 0$, $\mathrm{d}B/\mathrm{d}\,p > 0$, namely the supply of B increases, and the supply of A decreases, as p increases. It also follows from (19.19) that

$$\frac{\mathrm{d}B}{\mathrm{d}p} = -\frac{1}{p}\frac{\mathrm{d}A}{\mathrm{d}p}, \tag{19.20}$$

whence

$$-\frac{\mathrm{d}A}{\mathrm{d}B} = p, \tag{19.21}$$

as already shown in (19.10) and in Sect. 3.1. An alternative way of arriving at (19.20) is to start from the transformation curve, $B = h(A)$, whence

$$\frac{\mathrm{d}B}{\mathrm{d}p} = \frac{\mathrm{d}B}{\mathrm{d}A}\frac{\mathrm{d}A}{\mathrm{d}p},$$

and since $\mathrm{d}B/\mathrm{d}A = -1/p$ from (19.10), we have

$$\frac{\mathrm{d}B}{\mathrm{d}p} = -\frac{1}{p}\frac{\mathrm{d}A}{\mathrm{d}p}.$$

19.2.3 The Demand Side of the Model

Let us now consider the demand side of the model. If we differentiate I_A with respect to p, we obtain

$$\frac{\mathrm{d}I_A}{\mathrm{d}p} = \frac{\mathrm{d}A}{\mathrm{d}p} + B + p\frac{\mathrm{d}B}{\mathrm{d}p} = B, \tag{19.22}$$

because $\mathrm{d}A/\mathrm{d}p + p\ (\mathrm{d}B/\mathrm{d}p) = 0$ by (19.20). Therefore we can compute the total derivative of each demand function with respect to p:

$$\begin{aligned} \frac{\mathrm{d}A^D}{\mathrm{d}p} &= \frac{\partial A^D}{\partial I_A}\frac{\mathrm{d}I_A}{\mathrm{d}p} + \frac{\partial A^D}{\partial p} = \frac{\partial A^D}{\partial I_A}B + \frac{\partial A^D}{\partial p}, \\ \frac{\mathrm{d}B^D}{\mathrm{d}p} &= \frac{\partial B^D}{\partial I_A}\frac{\mathrm{d}I_A}{\mathrm{d}p} + \frac{\partial B^D}{\partial p} = \frac{\partial B^D}{\partial I_A}B + \frac{\partial B^D}{\partial p}. \end{aligned} \tag{19.23}$$

We assume that these demand functions are well behaved, namely $\partial A^D/\partial I_A > 0$, $\partial B^D/\partial I_A > 0$ (no inferior goods), and $\partial A^D/\partial p > 0$, $\partial B^D/\partial p < 0$ (normal price-effect: remember that p is p_B/p_A, so that $\partial A^D/\partial p > 0$ means $\partial A^D/\partial\,(1/p) < 0$). It follows that $\partial A^D/\partial p > 0$, as shown in Fig. 3.5b. The sign of $\mathrm{d}B^D/\mathrm{d}p$ remains indeterminate, for this derivative is the sum of a positive and a negative term; however, it can be shown that $\mathrm{d}B^D/\mathrm{d}p$ must be negative at least in the neighbourhood of the equilibrium point. In fact, differentiation of the budget constraint yields

$$\frac{\mathrm{d}A^D}{\mathrm{d}p} + B^D + p\frac{\mathrm{d}B^D}{\mathrm{d}p} = B,$$

whence

$$\frac{\mathrm{d}A^D}{\mathrm{d}p} + p\frac{\mathrm{d}B^D}{\mathrm{d}p} = B - B^D. \tag{19.24}$$

Now, $B - B^D = 0$ at the equilibrium point, so that $\mathrm{d}B^D/\mathrm{d}p < 0$ since $\mathrm{d}A^D/\mathrm{d}p > 0$. Also note that $\mathrm{d}B^D/\mathrm{d}p$ must a fortiori be negative below the equilibrium point, namely when $B - B^D < 0$.

19.3 International Trade and Offer Curves

19.3.1 The Equilibrium Conditions: The Offer Curve and Its Slope

Consider the excess demands for commodities A and B in country 1 and the budget constraint (Walras' law)

$$\begin{aligned} E_{1A}(p) &= A_1^D(I_{1A}, p) - A_1(p), \\ E_{1B}(p) &= B_1^D(I_{1A}, p) - B_1(p), \\ E_{1A}(p) + pE_{1B}(p) &= 0, \end{aligned} \tag{19.25}$$

which are written as functions of p only, because I_{1A} is a function of p as shown in Sect. 19.2.2. If the economy is closed, the equilibrium conditions (19.17) require that $E_{1A} = E_{1B} = 0$. If we introduce country 2, we have the relations

$$\begin{aligned} E_{2A}(p) &= A_2^D(I_{2A}, p) - A_2(p), \\ E_{2B}(p) &= B_2^D(I_{2A}, p) - B_2(p), \\ E_{2A}(p) + pE_{2B}(p) &= 0, \end{aligned} \tag{19.26}$$

where the terms of trade p must be the same in both countries as shown in the text.

International equilibrium requires that the world demands for the two commodities are equal to the respective world supplies, namely

$$\begin{aligned} E_{1A}(p) + E_{2A}(p) &= 0, \\ E_{1B}(p) + E_{2B}(p) &= 0. \end{aligned} \tag{19.27}$$

These conditions are not independent, for either one can be derived from the other given the two countries' budget constraints. By using these constraints, international equilibrium can also be expressed as

$$\begin{aligned} E_{1A}(p) &= pE_{2B}(p), \\ pE_{1B}(p) &= E_{2A}(p), \end{aligned} \tag{19.28}$$

where of course either one depends on the other. Hence p (the terms of trade) is determined, and by substituting back we determine all the other variables.

The graphic counterpart of (19.27) is Fig. 3.6; the graphic counterpart of (19.28) is Viner's terms-of-trade diagram (Viner, 1937, p. 362; see also Mosak, 1944, p. 77 and Kemp, 1964, pp. 62–63).

The derivation of the offer curve from the excess demand functions has already been shown in Sect. 3.4.1 and Fig. 3.7. Here we wish to demonstrate that *the offer curve is not necessarily well behaved* (namely monotonically increasing and concave to its import axis) *even if we assume underlying normal supply and demand schedules*.

Let us assume, as in the text, that country 1 wishes to import commodity A and to export commodity B, namely

$$E_{1A}(p) > 0, \quad E_{1B}(p) < 0, \tag{19.29}$$

so that we can write this country's offer curve as

$$-E_{1B} = G_1(E_{1A}), \tag{19.30}$$

where the minus sign serves to make $-E_{1B}$ a positive quantity. Similarly,

$$E_{2A}(p) < 0, \quad E_{2B}(p) > 0, \tag{19.31}$$

and

$$-E_{2A} = G_2(E_{2B}). \tag{19.32}$$

For our purpose it suffices to consider one offer curve, say that of country 1. We have

$$G_1' = \frac{\mathrm{d}(-E_{1B})}{\mathrm{d}E_{1A}} = -\frac{\mathrm{d}E_{1B}/\mathrm{d}p}{\mathrm{d}E_{1A}/\mathrm{d}p}. \tag{19.33}$$

By differentiating (19.25) and using (19.23) and (19.19) it can be seen that

$$\begin{aligned} \frac{\mathrm{d}E_{1A}}{\mathrm{d}p} &= \frac{\partial A_1^D}{\partial I_{1A}} B_1 + \frac{\partial A_1^D}{\partial p} - \frac{\mathrm{d}A_1}{\mathrm{d}p} > 0, \\ \frac{\mathrm{d}E_{1B}}{\mathrm{d}p} &= \frac{\partial B_1^D}{\partial I_{1A}} B_1 + \frac{\partial B_1^D}{\partial p} - \frac{\mathrm{d}B_1}{\mathrm{d}p} \gtrless 0. \end{aligned} \tag{19.34}$$

The ambiguity of the sign of $\mathrm{d}E_{1B}/\mathrm{d}p$ derives from the fact that $\mathrm{d}B_1^D/\mathrm{d}p$ has an ambiguous sign. Note that since $\mathrm{d}B_1^D/\mathrm{d}p < 0$ at the equilibrium point, as shown above, then $\mathrm{d}E_{1B}/\mathrm{d}p < 0$ at the equilibrium point, and so $G_1' > 0$ at the origin. But, as we move away from the origin the sign becomes indeterminate, since we are considering $E_{1B} < 0$ namely $B_1 - B_1^D > 0$, and Eq. (19.24) shows that $\mathrm{d}A_1^D/\mathrm{d}p > 0$ and $\mathrm{d}\, B_1^D/\mathrm{d}p > 0$ are perfectly compatible with the budget constraint.

We have thus proved that the offer curve, although increasing at the origin, need not be increasing everywhere, notwithstanding normal demand and supply functions for both commodities.

As regards its convexity or concavity, the sign of $G_1'' = \mathrm{d}^2(-E_{1B})/\mathrm{d}E_{1A}^2$ is also indeterminate, for it involves the second derivatives of the demand and supply functions with respect to p, which are indeterminate.

The conclusion is that cases such as that depicted in Fig. 3.11 (multiple equilibria) as well as cases in which the offer curves are decreasing over some interval cannot be ruled out on the basis of normal underlying demand and supply functions.

19.3.2 Relationships Between the Various Elasticities

Let us finally examine the relations between the elasticity of the offer curve, the elasticity of import demand and the elasticity of export supply. We first examine country 1's elasticities.

The elasticity of the offer curve is defined—see Eq. (3.18)—as

$$e_1 = \frac{\mathrm{d}\,(-E_{1B})}{\mathrm{d}E_{1A}} \frac{E_{1A}}{(-E_{1B})} = \frac{\mathrm{d}E_{1B}}{\mathrm{d}E_{1A}} \frac{E_{1A}}{E_{1B}}. \tag{19.35}$$

The (total) price-elasticity of import demand is defined as

$$\xi_1 = \frac{\mathrm{d}E_{1A}}{\mathrm{d}\,(1/p)} \frac{1/p}{E_{1A}} = \frac{\mathrm{d}E_{1A}}{\mathrm{d}\,(1/p)} \frac{1}{pE_{1A}} = -\frac{\mathrm{d}E_{1A}}{\mathrm{d}p} \frac{p}{E_{1A}}, \tag{19.36}$$

and the (total) price-elasticity of export supply is defined as

$$\varepsilon_1 = \frac{\mathrm{d}\,(-E_{1B})}{\mathrm{d}p} \frac{p}{(-E_{1B})} = \frac{\mathrm{d}E_{1B}}{\mathrm{d}p} \frac{p}{E_{1B}}, \tag{19.37}$$

where the adjective total serves to remind us that when p changes I_A changes as well (as a function of p), so that the quantity change includes both effects.

Since $E_{1A} = -pE_{1B}$ from the budget constraint, we can write (19.36) as

$$\xi_1 = \frac{\mathrm{d}E_{1A}}{\mathrm{d}\,(-E_{1B}/E_{1A})} \frac{E_{1B}}{-E_{1A}^2} = \left(\frac{\mathrm{d}E_{1B}}{\mathrm{d}E_{1A}} \frac{E_{1A}}{E_{1B}} - 1 \right)^{-1}. \tag{19.38}$$

It follows from (19.38) and (19.35) that

$$\xi_1 = (e_1 - 1)^{-1}, \quad e_1 = \frac{1+\xi_1}{\xi_1}. \tag{19.39}$$

Similarly we can write (19.37) as

$$\varepsilon_1 = \frac{\mathrm{d}E_{1B}}{\mathrm{d}\,(-E_{1A}/E_{1B})}\,\frac{-E_{1A}}{E_{1B}^2} = \left(\frac{\mathrm{d}E_{1A}}{\mathrm{d}E_{1b}}\,\frac{E_{1B}}{E_{1A}} - 1\right)^{-1}. \tag{19.40}$$

It follows from (19.40) and (19.35) that

$$\varepsilon_1 = \left(\frac{1}{e_1} - 1\right)^{-1} = \frac{e_1}{1-e_1}. \tag{19.41}$$

Therefore

$$1 + \xi_1 + \varepsilon_1 = 0. \tag{19.42}$$

Similarly it can be shown that

$$\xi_2 = (e_2 - 1)^{-1}, e_2 = \frac{1+\xi_2}{\xi_2}, \varepsilon_2 = \frac{e_2}{1-e_2}, 1 + \xi_2 + \varepsilon_2 = 0, \tag{19.43}$$

where

$$\begin{aligned} e_2 &= \frac{\mathrm{d}\,(-E_{2A})}{\mathrm{d}E_{2B}}\,\frac{E_{2B}}{(-E_{2A})}, \qquad \xi_2 = \frac{\mathrm{d}E_{2B}}{\mathrm{d}p}\,\frac{p}{E_{2B}}, \\ \varepsilon_2 &= \frac{\mathrm{d}\,(-E_{2A})}{\mathrm{d}\,(1/p)}\,\frac{(1/p)}{(-E_{2A})} = -\frac{\mathrm{d}E_{2A}}{\mathrm{d}p}\,\frac{p}{E_{2A}}. \end{aligned} \tag{19.44}$$

Consequently,

$$(1+\xi_1+\varepsilon_1) + (1+\xi_2+\varepsilon_2) = (1+\xi_1+\xi_2) + (1+\varepsilon_1+\varepsilon_2) = 0. \tag{19.45}$$

19.4 Stability

19.4.1 Terms-of-Trade Adjustment

In describing the determination of international equilibrium (Sect. 3.3) we assumed that p moves according to the pressure of world excess demands. The mathematical counterpart of this assumption is the following differential equation

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \psi\left[E_{1B}(p) + E_{2B}(p)\right] = \psi\left[E_{2B}(p) - \frac{1}{p}E_{1A}(p)\right], \tag{19.46}$$

where ψ is a sign-preserving function and $\psi'[0] \equiv v > 0$. To examine *local* stability, we expand the r.h.s. of (19.46) in Taylor's series at the equilibrium point and neglect all terms of order higher than the first, thus obtaining

$$\begin{aligned}\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} &= v\left(\frac{\mathrm{d}E_{2B}}{\mathrm{d}p} + \frac{1}{p^2}E_{1A} - \frac{1}{p}\frac{\mathrm{d}E_{1A}}{\mathrm{d}p}\right)\overline{p} \\ &= v\frac{E_{2B}}{p}\left(\frac{\mathrm{d}E_{2B}}{\mathrm{d}p}\frac{p}{E_{2B}} + \frac{E_{1A}}{pE_{2B}} - \frac{p}{pE_{2B}}\frac{\mathrm{d}E_{1A}}{\mathrm{d}p}\right)\overline{p},\end{aligned} \tag{19.47}$$

where $\overline{p} \equiv p - p_E$ denotes the deviations from equilibrium, and all derivatives, etc., are evaluated at the equilibrium point. Since $E_{1A} = pE_{2B}$ at the equilibrium point, by using the definitions of the price-elasticities of import demands—see (19.36) and (19.44)—we obtain

$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = v\frac{E_{2B}}{p}\left(1 + \xi_1 + \xi_2\right)\overline{p}. \tag{19.48}$$

Since $E_{2B} > 0$ by assumption—see (19.31)—the necessary and sufficient stability condition is

$$1 + \xi_1 + \xi_2 < 0. \tag{19.49}$$

Condition (19.49) is sometimes referred to as the "*Marshall-Lerner condition*". We completely agree with Kemp (1964, p. 70; 1969b, p. 84 fn. 5) when he writes that "Never were adjectives so incongruously applied. Marshall [...] developed a quite different stability condition; and Lerner was concerned neither with a barter nor with a dynamical economy". We only add that this denomination is even more incongruous when it is applied to the stability condition concerning the foreign exchanges.

19.4.2 Quantity Adjustment

Let us now examine the stability of equilibrium when the variables which adjust themselves are the quantities of exports, as Marshall believed; the two behaviour assumptions have been described in Sect. 3.4.2. The mathematical counterpart of *behaviour assumption I* is the following system of differential equations

$$\begin{aligned}\frac{\mathrm{d}(-E_{1B})}{\mathrm{d}t} &= \varphi_1\left[-E_{1B}(p) - (-E_{1B})\right], \\ \frac{\mathrm{d}(-E_{1A})}{\mathrm{d}t} &= \varphi_2\left[-E_{2A}(p) - (-E_{2A})\right],\end{aligned} \tag{19.50}$$

where φ_1, φ_2 are sign-preserving functions with $\varphi_1'[0] \equiv v_1 > 0, \varphi_2'[0] \equiv v_2 > 0$; the quantities $-E_{1B}$ and $-E_{2A}$ are the actual quantities of exports, whereas $-E_{1B}(p)$ and $-E_{2A}(p)$ are the *desired* quantities of exports at the *current* terms of trade p.

The linearization of system (19.50) at the equilibrium point is rather long-winded, and we refer the reader to Gandolfo (1971, p. 305) for the details. The result is

$$\begin{aligned} -\frac{\mathrm{d}E_{1B}}{\mathrm{d}t} &= (1+\varepsilon_1)\left[E_{1B} - E_{1B}(p_E)\right] - p_E\varepsilon_1\left[E_{2A} - E_{2A}(p_E)\right], \\ -\frac{\mathrm{d}E_{2A}}{\mathrm{d}t} &= -\frac{\varepsilon_2}{p_E}\left[E_{1B} - E_{1B}(p_E)\right] + (1+\varepsilon_2)\left[E_{2A} - E_{2A}(p_E)\right], \end{aligned} \tag{19.51}$$

where ε_1, ε_2 are defined in (19.37) and (19.44), and the units of quantity in both countries have been chosen so as to make $v_1 = v_2 = 1$. The characteristic equation of the differential equation system (19.51) is

$$\begin{vmatrix} 1+\varepsilon_1+\lambda & -p_E\varepsilon_1 \\ -\varepsilon_2/p_E & 1+\varepsilon_2+\lambda \end{vmatrix} = \lambda^2 + (2+\varepsilon_1+\varepsilon_2)\lambda + (1+\varepsilon_1+\varepsilon_2) = 0, \tag{19.52}$$

whose roots are $-1, -(1+\varepsilon_1+\varepsilon_2)$. Thus the movement will be monotonic, and will converge, if, and only if

$$1+\varepsilon_1+\varepsilon_2 > 0. \tag{19.53}$$

By using the relations between the ε and e elasticities—see (19.41) and (19.43)—the necessary and sufficient stability condition can be written as

$$\frac{1-e_1e_2}{(1-e_1)(1-e_2)} > 0. \tag{19.54}$$

Let us now consider *behaviour assumption II*, which gives rise to the following differential equation system

$$\begin{aligned} \frac{\mathrm{d}(-E_{1B})}{\mathrm{d}t} &= \varphi_1\left[G_1(E_{1A}) - (-E_{1B})\right], \\ \frac{\mathrm{d}(-E_{2A})}{\mathrm{d}t} &= \varphi_2\left[G_2(E_{2B}) - (-E_{2A})\right], \end{aligned} \tag{19.55}$$

where φ_1, φ_2 are sign-preserving functions with $\varphi_1'[0] \equiv s_1 > 0, \varphi_2'[0] \equiv s_2 > 0$; the quantities $-E_{1B}$ and $-E_{2A}$ are the *actual* quantities of exports, whereas $G_1(E_{1A})$ and $G_2(E_{2B})$ are the *desired* quantities of exports corresponding to the *current* quantities of imports E_{1A} and E_{2B} respectively.

The linear approximation to system (19.54) is (for details of the procedure see Gandolfo, 1971, pp. 308–309):

$$
\begin{aligned}
-\frac{\mathrm{d}E_{1B}}{\mathrm{d}t} &= \left[E_{1B} - E_{1B}\left(p_E\right)\right] - p_E e_1 \left[E_{2A} - E_{2A}\left(p_E\right)\right], \\
-\frac{\mathrm{d}E_{2A}}{\mathrm{d}t} &= \frac{e_2}{p_E}\left[E_{1B} - E_{1B}\left(p_E\right)\right] + \left[E_{2A} - E_{2A}\left(p_E\right)\right],
\end{aligned}
\tag{19.56}
$$

where e_1, e_2 are defined in (19.35) and (19.44), and the units of quantity in both countries have been chosen so as to make $s_1 = s_2 = 1$.The characteristic equation of this linear differential equation system is

$$
\begin{vmatrix} 1+\lambda & p_E e_1 \\ e_2/p_E & 1+\lambda \end{vmatrix} = \lambda^2 + 2\lambda + (1 - e_1 e_2) = 0, \tag{19.57}
$$

whose roots are $-1 \pm \sqrt{e_1 e_2}$. The movement can be either monotonic or oscillatory according to whether $e_1 e_2 \gtrless 0$. The necessary and sufficient stability condition is

$$
e_1 e_2 < 1. \tag{19.58}
$$

Note that when $e_1 e_2 < 0$ this condition is certainly satisfied, so that possible oscillatory movements are necessarily convergent.

19.5 Duality Approach

19.5.1 The Jones Model

We describe here the general equilibrium model due to Jones (1965), that will be put to use in subsequent chapters to prove some of the standard theorems in international trade theory.

Let $a_{ij}, i = K, L, j = A, B$, denote the quantity of factor i required to produce a unit of commodity j. Then we have

$$
\begin{aligned}
a_{LA}A + a_{LB}B &= L, \\
a_{KA}A + a_{KB}B &= K, \\
a_{LA}p_L + a_{KA}p_K &= p_A, \\
a_{LB}p_L + a_{KB}p_K &= p_B.
\end{aligned}
\tag{19.59}
$$

These equations emphasize the dual relations between factor endowments and commodity outputs (first two equations, which derive from full factor employment), and between commodity prices and factor prices (last two equations, which derive from competitive equilibrium). Since, in general, the input coefficients a_{ij} are variable, Eq. (19.59) must be supplemented by four equations to determine these coefficients. Such equations derive from the firm's optimization procedure. In fact, with constant returns to scale, the input coefficients depend solely upon the factor-price ratio: as can be seen from property (a) in Sect. 19.1.3, the input coefficients depend solely on the factor ratio which in turn—see property (b)—is uniquely determined, independently of the scale of production, by the factor-price ratio according to the cost minimization procedure. Therefore

$$a_{ij} = a_{ij}\left(\frac{p_L}{p_K}\right), \quad i = K, L; \quad j = A, B, \tag{19.60}$$

which are the four equations that we need. The eight Eqs. (19.59) and (19.60) describe the production side of the model, and make it possible to determine the eight unknowns a_{ij}, A, B, p_L, p_K given the four parameters L, K, p_A, p_B. Let us now write the total differentials of Eq. (19.59):

$$\begin{aligned}
A\mathrm{d}a_{LA} + a_{LA}\mathrm{d}A + B\mathrm{d}a_{LB} + a_{LB}\mathrm{d}B \qquad &= \mathrm{d}L,\\
A\mathrm{d}a_{KA} + a_{KA}\mathrm{d}A + B\mathrm{d}a_{KB} + a_{kB}\mathrm{d}B \qquad &= \mathrm{d}K,\\
p_L\mathrm{d}a_{LA} + a_{LA}\mathrm{d}p_L + p_K\mathrm{d}a_{KA} + a_{KA}\mathrm{d}p_K \; &= \mathrm{d}p_A,\\
p_L\mathrm{d}a_{LB} + a_{LB}\mathrm{d}p_L + p_K\mathrm{d}a_{KB} + a_{KB}\mathrm{d}p_K \; &= \mathrm{d}p_{B.}
\end{aligned} \tag{19.61}$$

If we denote the relative changes by an asterisk (namely, $a^*_{LA} \equiv \mathrm{d}a_{LA}/a_{LA}$, $A^* \equiv \mathrm{d}A/A$, etc.), we can rewrite Eq. (19.61), after simple manipulations,[2] in the form

$$\begin{aligned}
\lambda_{LA}A^* + \lambda_{LB}B^* &= L^* - \left(\lambda_{LA}a^*_{LA} + \lambda_{LB}a^*_{LB}\right),\\
\lambda_{KA}A^* + \lambda_{KB}B^* &= K^* - \left(\lambda_{KA}a^*_{KA} + \lambda_{KB}a^*_{KB}\right),\\
\theta_{LA}p^*_L + \theta_{KA}p^*_K &= p^*_A - \left(\theta_{LA}a^*_{LA} + \theta_{KA}a^*_{KA}\right),\\
\theta_{LB}p^*_L + \theta_{KB}p^*_K &= p^*_B - \left(\theta_{LB}a^*_{LB} + \theta_{KB}a^*_{KB}\right),
\end{aligned} \tag{19.62}$$

where $\lambda_{LA} \equiv a_{LA}A/L$, $\lambda_{LB} \equiv a_{LB}B/L$ denote the fractions of the labour force used in sector A and in sector B respectively; by the first equation in (19.59) these fractions must add up to one, $\lambda_{LA} + \lambda_{LB} = 1$. Similarly, the sum of $\lambda_{KA} \equiv a_{KA}A/K$ and $\lambda_{KB} \equiv a_{KB}B/K$ must be equal to one. The θ's denote the factor shares in each sector: $\theta_{LA} \equiv a_{LA}p_L/p_A$, $\theta_{KA} \equiv a_{KA}p_K/p_A$ and so on; by the last two equations in (19.59) these shares must add up to one, namely $\theta_{LA} + \theta_{KA} = 1$, $\theta_{LB} + \theta_{KB} = 1$.

If the input coefficients are fixed, $a^*_{ij} \equiv 0$, and so Eq. (19.62) are greatly simplified. But in the general case of variable coefficients we need four additional equations to determine the four a^*_{ij}. These are

---

[2]Consider for example the first equation and divide both sides by L, obtaining

$$\frac{A}{L}\mathrm{d}a_{LA} + a_{LA}\frac{1}{L}\mathrm{d}A + \frac{B}{L}\mathrm{d}a_{LB} + a_{LB}\frac{1}{L}\mathrm{d}B = \frac{\mathrm{d}L}{L}.$$

Then multiply and divide the first term on the left by a_{LA} and so on; the result is

$$\frac{a_{LA}A}{L}\cdot\frac{\mathrm{d}a_{LA}}{a_{LA}} + \frac{a_{LA}A}{L}\cdot\frac{\mathrm{d}A}{A} + \frac{a_{LB}B}{L}\cdot\frac{\mathrm{d}a_{LB}}{a_{LB}} + \frac{a_{LB}B}{L}\cdot\frac{\mathrm{d}B}{B} = L,$$

which is the first equation in (19.62).

$$\begin{aligned}
\theta_{LA}a_{LA}^* + \theta_{KA}a_{KA}^* &= 0,\\
\theta_{LB}a_{LB}^* + \theta_{KB}a_{KB}^* &= 0,\\
\frac{a_{KA}^* - a_{LA}^*}{p_L^* - p_K^*} &= \sigma_{A,}\\
\frac{a_{KB}^* - a_{LB}^*}{p_L^* - p_K^*} &= \sigma_B.
\end{aligned} \tag{19.63}$$

The first two equations are easily derived by the usual cost minimization procedure. For any given output level the entrepreneur minimizes costs, treating factor prices as given. In other words, the entrepreneur chooses the input coefficients so as to minimize unit costs. The first order condition is, for commodity A

$$\mathrm{d}\left(a_{LA}p_L + a_{KA}p_K\right) = p_L \mathrm{d}a_{LA} + p_K \mathrm{d}a_{KA} = 0.$$

Dividing by p_A and expressing the changes in relative terms, we obtain the first equation in (19.63); we obtain the second equation in a similar way.

The third and fourth equations in (19.63) define the elasticity of substitution between factors in each sector by using the fact that, in equilibrium, the slope of the isoquant (the marginal rate of technical substitution) in each sector is equal to the ratio of factor prices, so that the proportional change in the marginal rate of substitution (which appears in the denominator of the formula defining the elasticity of substitution) can be expressed as $p_L^* - p_K^*$.

We can use Eq. (19.63) to express the proportional changes in the input coefficients in terms of the proportional changes in factor prices, namely

$$\begin{aligned}
a_{Lj}^* &= -\theta_{Kj}\sigma_j\left(p_L^* - p_K^*\right),\\
a_{Kj}^* &= \theta_{Lj}\sigma_j\left(p_L^* - p_K^*\right).
\end{aligned} \quad j = A, B \tag{19.64}$$

These expressions can be substituted in Eq. (19.62); the result is

$$\begin{aligned}
\lambda_{LA}A^* + \lambda_{LB}B^* &= L^* + \delta_L\left(p_L^* - p_K^*\right),\\
\lambda_{KA}A^* + \lambda_{KB}B^* &= K^* - \delta_K\left(p_L^* - p_K^*\right),\\
\theta_{LA}p_L^* + \theta_{KA}p_K^* &= p_A^*,\\
\theta_{LB}p_L^* + \theta_{KB}p_K^* &= p_B^*,
\end{aligned} \tag{19.65}$$

where $\delta_L \equiv \lambda_{LA}\,\theta_{KA}\sigma_A + \lambda_{LB}\theta_{KB}\sigma_B$, $\delta_K \equiv \lambda_{KA}\theta_{LA}\sigma_A + \lambda_{KB}\theta_{LB}\sigma_B$; note that δ_L and δ_K are zero in the case of fixed coefficients.

To close the model, the demand side must be introduced. To keep things as simple as possible, it is assumed that community taste patterns are homothetic and that no difference exists between the taste patterns of workers and capitalists. Therefore the ratio of the quantities demanded of A and B depends solely upon the commodity price-ratio:

$$\frac{A^D}{B^D} = h\left(\frac{p_A}{p_B}\right). \tag{19.66}$$

If we differentiate Eq. (19.66) and express the changes in relative form, we obtain

$$\frac{A^D}{B^D}\left(A^{D^*} - B^{D^*}\right) = h'\frac{p_A}{p_B}\left(p_A^* - p_B^*\right), \tag{19.67}$$

and if we use the definition of the elasticity of substitution between the two commodities on the demand side,

$$\sigma_D \equiv -\frac{\dfrac{\mathrm{d}\left(A^D/B^D\right)}{A^D/B^D}}{\dfrac{\mathrm{d}\left(p_A/p_B\right)}{p_A/p_B}} = h'\frac{B^D}{A^D}\frac{p_A}{p_B}, \tag{19.68}$$

we arrive at

$$A^{D^*} - B^{D^*} = -\sigma_D\left(p_A^* - p_B^*\right). \tag{19.69}$$

Equation (19.69) gives directly the change in the ratio of outputs consumed; to obtain the change in the ratio of outputs produced, we subtract the second equation in (19.65) from the first, which gives

$$(\lambda_{LA} - \lambda_{KA})\,A^* + (\lambda_{LB} - \lambda_{KB})\,B^* = \left(L^* - K^*\right) + (\delta_L + \delta_K)\left(p_L^* - p_K^*\right). \tag{19.70}$$

Now, from the fact that $\lambda_{LA} + \lambda_{LB} = \lambda_{KA} + \lambda_{KB} = 1$ it follows that $\lambda_{LA} - \lambda_{KA} = \lambda_{KB} - \lambda_{LB}$ and so

$$A^* - B^* = \frac{L^* - K^*}{\lambda_{LA} - \lambda_{KA}} + \frac{\delta_L + \delta_K}{\lambda_{LA} - \lambda_{KA}}\left(p_L^* - p_K^*\right). \tag{19.71}$$

Similarly, by subtracting the fourth equation in (19.65) from the third and noting that $\theta_{LA} - \theta_{LB} = \theta_{KB} - \theta_{KA}$ (since $\theta_{LA} + \theta_{KA} = \theta_{LB} + \theta_{KB} = 1$), we have

$$p_L^* - p_K^* = \frac{1}{\theta_{LA} - \theta_{LB}}\left(p_A^* - p_B^*\right). \tag{19.72}$$

Substitution of (19.72) into (19.71) gives

$$A^* - B^* = \frac{L^* - K^*}{\lambda_{LA} - \lambda_{KA}} + \sigma_S\left(p_A^* - p_B^*\right), \tag{19.73}$$

where $\sigma_S \equiv \dfrac{\delta_L + \delta_K}{(\lambda_{LA} - \lambda_{KA})(\theta_{LA} - \theta_{LB})}$ represents the elasticity of substitution between the commodities on the supply side (along the transformation curve).

Finally observe that, in equilibrium, demand equals supply, so that $A^{D^*} = A^*, B^{D^*} = B^*$. Therefore from (19.73) and (19.69) we obtain the change in the commodity price-ratio, which turns out to be

$$p_A^* - p_B^* = -\frac{1}{(\lambda_{LA} - \lambda_{KA})(\sigma_S + \sigma_D)} \left(L^* - K^*\right). \tag{19.74}$$

Consequently the change in the ratio of commodities produced is

$$A^* - B^* = \frac{1}{\lambda_{LA} - \lambda_{KA}} \frac{\sigma_D}{\sigma_S + \sigma_D} \left(L^* - K^*\right). \tag{19.75}$$

This completes the description of the equations of change of the model. As we said above, these will be put to use in the following chapters, in order to derive some important theorems in the theory of international trade; see Sects. 20.1–20.3, 21.1–21.3, 22.3, and 22.6; in some of these, further properties and extensions of the model are also examined.

19.5.2 Revenue Functions and Expenditure Functions

The duality approach is usually presented in terms of revenue functions (or GNP functions) and expenditure functions rather than in terms of the equations described in the previous section. These alternative presentations are, however, equivalent. The revenue and expenditure functions, in fact, respectively summarize the production side and the consumption side of the economy. Here we simply show their basic nature and relate it to the Jones model explained in the previous section. For an in-depth treatment see Dixit and Norman (1980); Sgro (1986); Woodland (1982), Diewert (1974, 1982) and Cornes (1992).

Let us begin with the production side. We have seen in the previous section that the production equations determine the eight unknowns a_{ij}, A, B, p_L, p_K given the four parameters L, K, p_A, p_B. We can thus write the *value of GNP*, or *revenue function*, as

$$R = p_A A(p_A, p_B; L, K) + p_B B(p_A, p_B; L, K). \tag{19.76}$$

Since the production equations have been derived from an optimization process, R is clearly an optimal value.

If we now consider the (optimal) *unit cost functions*

$$\begin{aligned} c_A &= a_{LA} p_L + a_{KA} p_K, \\ c_B &= a_{LB} p_L + a_{KB} p_K, \end{aligned} \tag{19.77}$$

we get the *total cost function*

$$\begin{aligned} C &= c_A A + c_B B \\ &= R \\ &= p_L(a_{LA} A + a_{LB} B) + p_K(a_{KA} A + a_{KB} B) \\ &= p_L L + p_K K, \end{aligned} \tag{19.78}$$

since $c_A = p_A, c_B = p_B$, and $a_{LA} A + a_{LB} B = L, a_{KA} A + a_{KB} B = K$ by the production Eq. (19.59). This expresses the fact that the (maximum) production revenue equals the (minimum) cost of production.

Equations (19.78) are implicit in Jones's model, and their properties of change have already been examined in the previous section.

As regards the demand side, the *expenditure function* $E = f(u, \mathbf{p})$ is defined as the minimum required expenditure as a function of the utility target to be achieved (u) and the vector of money prices ($\mathbf{p}$) faced by the consumer. Given the dual nature of the consumer optimization problem (minimum expenditure to achieve a given utility level, or maximum achievable utility for a given money income), the value of E also expresses the money income Y just sufficient to achieve the utility level.

The expenditure function is related to the standard Hicksian compensated demand functions through *Shephard's lemma*, which shows that $\partial E/\partial p_j$, the partial derivative of the expenditure function with respect to any price p_j, coincides with the compensated demand function for commodity $j, q_j^d = q_j(u, \mathbf{p})$. At the optimum point, the compensated and uncompensated demand functions coincide, hence we have

$$\partial E/\partial p_j = q_j(u, \mathbf{p}) = x_j(Y, \mathbf{p}), \tag{19.79}$$

where $x_j^d = x_j(Y, \mathbf{p})$ is the standard uncompensated (or Marshallian) demand function. Since the expenditure function is homogeneous of the first degree in all prices, its partial derivatives are homogeneous of degree zero with respect to prices, namely the demand function is a function of relative rather than absolute prices, a well known result.

In the case of homothetic utility functions, the expenditure function takes on the particularly simple form

$$E = ue(\mathbf{p}), \tag{19.80}$$

i.e., similar to the cost function for firms with constant-returns-to scale technology; in fact, the expenditure function is homogeneous of degree one with respect to the utility level.

References

Cornes, R. (1992). *Duality and modern economics*.
Diewert, W. E. (1974). Applications of duality theory.
Diewert, W. E. (1982). Duality approaches to microeconomic theory.
Dixit, A. K., & Norman, V. (1980). *Theory of international trade: A dual, general equilibrium approach* (chaps. 1–3).
Gandolfo, G. (1971). *Mathematical methods and models in economic dynamics* (Pt. II, chap. 9, sect. 30.
Gandolfo, G. (2009). *Economic dynamics* (Pt. II, chap. 19, sect. 19.3; Pt. III, chap. 20).
Herberg, H. (1969). On the shape of the transformation curve in the case of homogeneous production functions.
Jones, R. W. (1965). The structure of simple general equilibrium models.
Kemp, M. C. (1964). *The pure theory of international trade* (Pt. II).
Kemp, M. C. (1969b). *The pure theory of international trade and investment* (Pt. I).
Mosak, J. L. (1944). *General-equilibrium theory in international trade*.
Sgro, P. M. (1986). *The theory of duality and international trade*.
Viner, J. (1937). *Studies in the theory of international trade*.
Woodland, A. D. (1982). *International trade and resource allocation*.

Chapter 20
Appendix to Chapter 4

20.1 Factor-Intensity Reversals

We have shown in the text the crucial importance of the strong factor-intensity assumption (i.e., absence of factor-intensity reversals); here we examine formally the conditions under which reversals are present or absent. Let us begin by establishing the relationship between capital intensity and relative price of factors; for this purpose we employ the equilibrium conditions that state the equality between the value of marginal productivity of a factor and its price (this must be equal in both sectors). With the symbology introduced in Eq. (19.17), we have

$$\begin{aligned} g_A' \qquad &= p g_B' = p_K, \\ g_A - \varrho_A g_A' &= p\left(g_B - \varrho_B g_B'\right) = p_L, \end{aligned} \tag{20.1}$$

whence dividing the second equation by the first

$$\frac{p_L}{p_K} = \frac{g_i - \varrho_i g_i'}{g_i'} = \frac{g_i}{g_i'} - \varrho_i, \quad i = A, B. \tag{20.2}$$

Since g_i and g_i' are functions of ϱ_i, Eq. (20.2) expresses a relation between p_L/p_K and ϱ_i.This relation is increasing monotonically: in fact,

$$\frac{\mathrm{d}\left(p_L/p_K\right)}{\mathrm{d}\varrho_i} = \frac{\left(g_i'\right)^2 - g_i'' g_i}{\left(g_i'\right)^2} - 1 = -\frac{g_i'' g_i}{\left(g_i'\right)^2}, \tag{20.3}$$

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_20,

whence, by the inverse-function differentiation rule,[1]

$$\frac{\mathrm{d}\varrho_i}{\mathrm{d}\,(p_L/p_K)} = -\frac{\left(g_i'\right)^2}{g_i'' g_i}. \tag{20.4}$$

Since $g_i > 0$, $g_i'' < 0$ by the assumption of positive but decreasing marginal productivities, the derivatives (20.3) and (20.4) are positive. Equation (20.2) give rise to the curves drawn in Fig. 4.2.

We must now find the conditions under which these curves do or do not intersect (presence or absence of factor-intensity reversals, respectively). Interesting conditions are provided by the following theorem:

If the elasticity of substitution between factors is constant in each sector, no (one) factor-intensity reversal will occur when this elasticity is the same in (different between) the two sectors.

It should be remembered that the elasticity of substitution is defined as

$$\sigma_i = \frac{\mathrm{d}\varrho_i/\varrho_i}{\mathrm{d}MRTS/MRTS} = \frac{\mathrm{d}\varrho_i/\varrho_i}{\mathrm{d}\,(p_L/p_K)/(p_L/p_K)}, \tag{20.5}$$

where *MRTS* is the marginal rate of technical substitution along an isoquant, equal to the factor-price ratio in the situation of equilibrium.

From (20.5) we obtain

$$\frac{\mathrm{d}\varrho_i}{\varrho_i} = \sigma_i \frac{\mathrm{d}q}{q}, \tag{20.6}$$

where q denotes the factor-price ratio. Now, if we assume that σ_i is a constant, we can integrate throughout and obtain

$$\varrho_i = C_i q^{\sigma_i}, \tag{20.7}$$

where C_i depends on the arbitrary constants of integration. Then, if $\sigma_A = \sigma_B$, from Eq. (20.7) it follows that

$$\frac{\varrho_A}{\varrho_B} = C, \qquad C \equiv C_A/C_B, \tag{20.8}$$

and so either ϱ_A will always be greater than ϱ_B (if $C > 1$) or vice versa (if $C < 1$): no factor-intensity reversal can occur. This is the case represented in Fig. 4.2a.

[1] For a complete proof that (20.2) is a one-to-one correspondence between ϱ_i and p_L/p_K see, for example, Gandolfo (1971, Appendix III, sect. 7, §7.5).

If, on the contrary, $\sigma_A \neq \sigma_B$ (for example we assume $\sigma_A > \sigma_B$), from (20.7) we get

$$\frac{\varrho_A}{\varrho_B} = C q^{\sigma_A - \sigma_B}, \quad C \equiv C_A / C_B. \tag{20.9}$$

Since the function $C q^{\sigma_A - \sigma_B}$ is increasing monotonically from zero to infinity, a unique value of q will exist, call it q^*, such that $\varrho_A / \varrho_B \lesseqgtr 1$ for $q \lesseqgtr q^*$. There will thus be one, and only one, factor intensity reversal, as is the case in Fig. 4.2b.

It is important to note that, when the elasticity of substitution is variable, the integration allowing the passage from (20.6) to (20.7) can no longer be performed, so that, in general, any number of reversals can occur.

As a typical example of production functions never giving rise to factor-intensity reversals, we recall the Cobb-Douglas function, $Y = HK^{\alpha} L^{1-\alpha}$ which has a constant elasticity of substitution equal to one, whilst the *CES* function, $Y = \left[\alpha K^{-\beta} + \gamma L^{-\beta}\right]^{-(1/\beta)}$, has a constant elasticity of substitution equal to $1/\left(1+\beta\right)$, and so can give rise to a reversal when the parameter β is different between the two sectors.

We must now demonstrate the one-to-one correspondence between the relative price of goods and the relative price of factors in the absence of factor-intensity reversals. This amounts to showing that there exists a monotonic relationship between the two variables if and only if no factor-intensity reversal occurs.

Let us consider the equilibrium conditions given in (19.59), namely

$$\begin{aligned} p_B &= a_{LB} p_L + a_{KB} p_K, \\ p_A &= a_{LA} p_L + a_{KA} p_K. \end{aligned} \tag{20.10}$$

If we divide the first equation by the second we get

$$\frac{p_B}{p_A} = \frac{q a_{LB} + a_{KB}}{q a_{LA} + a_{KA}}, \quad q \equiv p_L / p_K, \tag{20.11}$$

whence, by differentiation with respect to q (remember that the coefficients a_{ij} are functions of q through the optimization procedure), we obtain

$$\begin{aligned} &\frac{\mathrm{d}\left(p_B / p_A\right)}{\mathrm{d}q} \\ &= \frac{\left(a_{LB} + q a'_{LB} + a'_{KB}\right)\left(q a_{LA} + a_{KA}\right) - \left(a_{LA} + q a'_{LA} + a'_{KA}\right)\left(q a_{LB} + a_{KB}\right)}{\left(q a_{LA} + a_{KA}\right)^2}, \end{aligned} \tag{20.12}$$

where $a'_{ij} \equiv \mathrm{d}a_{ij} / \mathrm{d}q$. Now, from the optimum conditions,

$$\begin{aligned} p_L \mathrm{d}a_{LA} + p_K \mathrm{d}a_{KA} &= 0, \\ p_L \mathrm{d}a_{LB} + p_K \mathrm{d}a_{KB} &= 0, \end{aligned}$$

and so

$$\begin{aligned} qa'_{LA} + a'_{KA} &= 0, \\ qa'_{LB} + a'_{KB} &= 0. \end{aligned} \tag{20.13}$$

Thanks to (20.13), expression (20.12) simplifies to

$$\frac{\mathrm{d}\,(p_B/p_A)}{\mathrm{d}q} = \frac{a_{LB}a_{KA} - a_{LA}a_{KB}}{(qa_{LA} + a_{KA})^2} = a_{LA}a_{LB}\frac{\varrho_A - \varrho_B}{(qa_{LA} + a_{KA})^2}, \tag{20.14}$$

from which it can readily be seen that the derivative of the relative price of goods with respect to the relative price of factors is either always positive or always negative if and only if ϱ_A is either always greater or always smaller than ϱ_B, that is, if and only if no factor intensity reversal occurs. When, on the contrary, one or more reversals are present, the derivative (20.14) will change its sign (passing through zero) one or more times, and so the relation between p_B/p_A and p_L/p_K will be no longer monotonic. In Fig. 4.5a we have represented this relation when $\varrho_A > \varrho_B$ everywhere, whilst Fig. 4.5b represents the case of one factor-intensity reversal ($\varrho_B > \varrho_A$ initially, and then $\varrho_A > \varrho_B$).

20.2 Proof of the Fundamental Theorem

The basic Heckscher-Ohlin proposition to be proved is that a country abundant in a factor has a production bias in favour of the commodity which uses that factor more intensively. In what follows we are going to use the physical definition of factor abundance.

If we consider the full-employment relations (see Sect. 19.5)

$$\begin{aligned} a_{KA}A + a_{KB}B &= K, \\ a_{LA}A + a_{LB}B &= L, \end{aligned} \tag{20.15}$$

and divide through by L, we obtain

$$\begin{aligned} a_{KA}A/L + a_{KB}B/L &= K/L, \\ a_{LA}A/L + a_{LB}B/L &= 1. \end{aligned} \tag{20.16}$$

By solving this linear system we can express A/L and B/L in terms of the remaining quantities, namely

$$A/L = \frac{a_{LB}K/L - a_{KB}}{a_{KA}a_{LB} - a_{KB}a_{LA}}, \quad B/L = \frac{a_{KA} - a_{LA}K/L}{a_{KA}a_{LB} - a_{KB}a_{LA}}, \tag{20.17}$$

whence

$$\frac{A/L}{B/L} = \frac{A}{B} = \frac{a_{LB}K/L - a_{KB}}{a_{KA} - a_{LA}K/L}. \tag{20.18}$$

Equation (20.18) expresses the output ratio (A/B) in terms of the factor endowment ratio (K/L), given the technical coefficients a_{ij}. These coefficients depend on the factor-price ratio but, given this, are constant for any output level owing to the assumption of constant returns to scale. Therefore, for any factor-price ratio we can compute the derivative

$$\frac{\mathrm{d}(A/B)}{\mathrm{d}\,(K/L)} = \frac{a_{KA}a_{LB} - a_{LA}a_{KB}}{(a_{KA} - a_{LA}K/L)^2} = a_{LA}a_{LB}\frac{\varrho_A - \varrho_B}{(a_{KA} - a_{LA}K/L)^2}, \tag{20.19}$$

which will have an unambiguous sign thanks to the assumption of no factor intensity reversal; this assumption enables us to state that either ϱ_A is always greater than ϱ_B or ϱ_B is always greater than ϱ_A *independently of* the factor-price ratio. If we assume, as in the text, that commodity A is capital intensive, the derivative under consideration turns out to be positive, that is, the greater the factor endowment ratio (K/L) the higher the output of A relative to B, and vice versa. Since the production functions are assumed to be internationally identical, the above result holds for both countries; this proves the basic proposition, which can be used as a lemma in the proof of the fundamental theorem exactly as in the text.

We have also stated in the text—see Fig. 4.7—that in the pre-trade equilibrium situation, assuming that country 1 is capital abundant relative to country 2 ($\varrho_1 > \varrho_2$) and commodity A is capital-intensive relative to B ($\varrho_A > \varrho_B$), for any given A/B ratio the marginal rate of transformation, and so the relative price of goods (p_B/p_A), is higher in country 1 than in country 2. To show this it is sufficient to observe that, with no factor-intensity reversal, $\varrho_1 > \varrho_2$ implies $(p_L/p_K)_1 > (p_L/p_K)_2$—see Fig. 4.8—so that, by (20.1), $(p_B/p_A)_1 > (p_B/p_A)_2$ since by assumption $\varrho_A > \varrho_B$. It can also be seen that $(p_B/p_A)_1 = (p_B/p_A)_2$ when $\varrho_1 = \varrho_2$, so that no international trade can take place when the relative factor endowments of the countries coincide.

We conclude by observing that, as we have said in Sect. 1.2, the Heckscher-Ohlin model stresses the difference in factor endowments as the basis for trade, whilst the Ricardian theory emphasizes the differences in technology. however, Ford (1982) has argued that under certain conditions the two theories are in fact equivalent. This has generated considerable controversy (see, for example, Ford, 1985; Lloyd, 1985); for a balanced exposition of the issues involved we refer the reader to Neary (1985a), who argues that "what is at stake is not the *logical* but the *observational* equivalence of the two theories".

20.3 The Factor-Price-Equalization Theorem

Let us take up Eqs. (20.1) again (these postulate the absence of factor-intensity reversals) and rewrite them in the form

$$\begin{aligned} g_A'(\varrho_A) - p g_B'(\varrho_B) &= 0, \\ \left[g_A(\varrho_A) - \varrho_A g_A'(\varrho_A)\right] - p\left[g_B(\varrho_B) - \varrho_B g_B'(\varrho_B)\right] &= 0. \end{aligned} \tag{20.20}$$

We have here a set of two implicit functions in three variables: p, ϱ_A, ϱ_B. By using the implicit-function theorem, we can express ϱ_A and ϱ_B as single-valued and differentiable functions of p if the Jacobian of (20.20) with respect to ϱ_A, ϱ_B is different from zero. This Jacobian turns out to be

$$\begin{vmatrix} g_A'' & -p g_B'' \\ -\varrho_A g_A'' & p \varrho_B g_B'' \end{vmatrix} = p g_A'' g_B'' (\varrho_B - \varrho_A), \tag{20.21}$$

which is different from zero if and only if no factor-intensity reversal occurs. Thus, if we assume absence of reversals it follows that a unique value of ϱ_A and a unique value of ϱ_B will correspond to the relative price of commodities, p, determined as a consequence of international trade. By substituting these values in (20.1), the values of p_K and p_L can be uniquely determined. Now, given the assumption of internationally identical production functions, Eqs. (20.20) and, therefore, the single-valued relations between ϱ_A and p and between ϱ_B and p are identical in both countries; similarly identical are Eqs. (20.1). Therefore, as p is the same given the assumptions of free trade and no transports costs, the absolute prices of factors will be equalized between countries.

Alternatively, we could have used the one-to-one relation between the commodity price ratio and the factor price ratio demonstrated at the end of Sect. 20.1 and then the one-to-one relation between the relative price of factors and ϱ_i demonstrated at the beginning of the same section.

It should be noted, in conclusion, that the condition on the Jacobian ensures univalence only *locally*, that is, in the neighbourhood of the equilibrium point; for the conditions for global univalence see Gale and Nikaidô (1965) .

Further considerations on the FPE theorem are contained in Sect. 20.5.

20.4 A Brief Outline of the Generalizations of the Heckscher-Ohlin Model

The attempts at extending the Heckscher-Ohlin theorem and the factor price equalization theorem to the general multi-commodity, multi-factor, multi-country case, have given rise to an immense literature which it would be impossible to deal with here. Therefore we do no more than focus on what we feel are some of the most important points, referring the reader to the surveys by Chipman (1966) and

Ethier (1984) for the rest. In Sects. 20.5 and 20.6 we shall consider in more depth two important aspects of these generalizations.

Jones (1956) formulated the "chain proposition" in the many-commodity, two-factor, two-country model, whereby if the goods are ranked in order of factor intensities, then all of a country's exports must lie higher in this list than all its imports. Bhagwati (1972) showed this proposition to be incorrect, if factor-price equalization obtains. Deardoff (1979) gave a formal proof of the non-factor-price equalization case (in which the chain proposition is true) and provided an extension to the many-country case, showing that all of the exports of a country more abundant in a factor will be at least as intensive in that factor as each of the exports of all countries less abundant in that factor.

The reader will note that these extensions remain within the context of the *two-factor* assumption. In fact, except for special cases, the concept itself of factor intensity can no longer be clearly defined when there are many factors.

These difficulties have led to a search for an alternative formulation of the Heckscher-Ohlin theorem, which should be more or less equivalent to the original one in the $2 \times 2 \times 2$ case *and* be capable of easy generalization. Such a formulation (called the *factor-content* version of the Heckscher-Ohlin theorem) exists, and refers to the factors embodied in the goods traded internationally, instead of the goods themselves. In the simple $2 \times 2 \times 2$ case this formulation states that *each country is a net exporter of the (services of the) country's more abundant factor and a net importer of the (services of the) other factor*.

This is the path followed, for example, by Vanek (1968), who used the same basic assumptions as in the original theorem and assumed, in addition, factor-price equalization and productive diversification. We use "productive diversification" in Chipman's sense (1966, p. 21). The precise assumption of (Vanek, 1968, p. 750) is: "specialization (in the two-country world) in no more than $m - n$ products" where m is the number of products, n the number of factors, and $m \geq n$.

With these assumptions Vanek achieved interesting results in the context of a two-country model, but with any number of goods and factors. Let us denote by ${}^1V_i, {}^2V_i$ the endowments of the i-th factor ($i = 1, 2, \ldots, r$) in countries 1 and 2 respectively. Now, if the relation

$$\frac{{}^1V_1}{{}^2V_1} \geq \frac{{}^1V_2}{{}^2V_2} \geq \ldots \geq \frac{{}^1V_r}{{}^2V_r}, \tag{20.22}$$

holds with at least one strict inequality, then free international trade in commodities brings about the following consequences (amongst others):

(a) Country 1 is a net exporter of the services of factors $1, 2, \ldots, j$, with $j < r$, and a net importer of the services of factors $j + 1, \ldots, r$;
(b) j can be determined if we know the vector of factor prices;
(c) Knowing this vector, we can also compute exactly the net amounts of the services of the factors traded internationally.

These are interesting results (which can be extended to the case of more than two countries: see Horiba 1974), but are obtained at the cost of a serious limitation, that is, the assumption that factor price equalization obtains. What was an important result, demonstrated as another theorem in the original version of the theory, now becomes a basic assumption like, say, the international identity of production functions etc.

Other writers have tried to do without this very restrictive assumption, but only at the cost of introducing other and perhaps equally restrictive ones (see, for example, Harkness, 1978, 1983). Brecher and Choudhri (1982) have succeeded in proving the validity of the factor-content version of the Heckscher-Ohlin theorem without the assumption of factor price equalization or other restrictive assumptions, but only in the two-factor multi-commodity model of a two-country world.

Deardoff (1982), in the general case of the multi-factor multi-commodity multi-country model and without recourse to the assumption of factor price equalization or other restrictive assumptions, has proved that both the factor-content and the commodity version of the Heckscher-Ohlin theorem are valid in an *average* sense. More precisely, as regards the factor-content version, he has shown that the simple correlation between the vector containing the autarky factor prices (which inversely reflect the abundance of those factors: Deardoff is using the economic definition of abundance) of all countries and factors and the vector containing the net exports by each country of (the services of) each factor, arranged in the same order, is negative. The interpretation of this result is that countries will on average tend to be net exporters of their abundant factors and net importers of their scarce factors. As regards the commodity version, Deardoff shows that the "*comvariance*"[2] among the vector containing a measure of factor abundance (for each factor and country), the vector containing a measure of factor intensity, and the vector of net exports at world prices, is positive. The economic interpretation is that exported goods must *on average* use the relatively abundant factors relatively intensively, and imported goods must *on average* use the relatively scarce factors relatively intensively. This important result generalizes the Heckscher-Ohlin theorem as an explanation of the pattern of commodity trade in an "average" sense.

For results similar to Deardoff's, see Ethier (1982), Dixit and Woodland (1982), Helpman (1984a), and Svensson (1984).

An alternative approach to the general case is also possible, which consists in aggregating a higher dimensional model so as to obtain a model which exhibits all the properties of the two-by-two model (provided that suitable restrictions are imposed); for this line of research see, for example, Neary (1984, 1985b), and references therein.

Another important point is the generalization of the factor price equalization theorem. It is perhaps worth mentioning, in passing, that the debate on this generalization—beginning with an incorrect conjecture by Samuelson (1953)—has given origin to a new mathematical theorem, the Gale and Nikaidô (1965) on the global univalence of mappings.

[2]This is a term used by Deardoff (1982, p. 690) to denote a generalization (that he suggested) of the concept of covariance when one needs to correlate three variables symmetrically.

Three cases must be distinguished in examining factor price equalization in the general case.

1. The number of commodities is *equal* to the number of factors. In this case, if complete productive diversification obtains and the cost functions are *globally* invertible,[3] then—independently of the factor endowments of the various countries—the equalization of commodity prices will involve the equalization of factor prices.
2. The number of commodities is *smaller* than the number of factors. In this case the determination of factor prices depends not only on the (international) prices of commodities (assumed to be known), but also on the trading countries' factor endowments. Generally speaking, the difference in these endowments causes the non-equalization of factor prices. In other words, this equalization, though not impossible, is extremely unlikely.
3. The number of commodities n is *greater* than the number of factors r. In this case the determination of factor prices depends only on the prices of r commodities, but we do not know which the r commodities are. Thus to be sure that factor prices will be equalized, the global invertibility conditions must be verified for all square $r \times r$ submatrices drawn from the Jacobian of the system relating the vector of commodity prices to the vector of factor prices.

For further analysis of factor-price equalization see Deardoff (1994), Feenstra (2004).

Recent theoretical research on the generalizations of the Heckscher-Ohlin model has concentrated on the role of factor mobility, a topic dealt with in Sect. 6.8.1 and its appendix.

20.5 The Factor Price Equalization Set

Assume that there are N goods indexed by g and M factors indexed by j and that $N \geq M \geq 2$. Let $\mathbf{p}_v^*$ be the row vector of the integrated equilibrium factors price and let $a_{jg}\left(\mathbf{p}_v^*\right)$ be the unitary demand function for factor j in the production of good g. Let $\mathbf{\Delta}^*$ be the M by N technology matrix

$$\mathbf{\Delta}^* \equiv \begin{bmatrix} a_{11}\left(\mathbf{p}_v^*\right) & \dots & a_{1N}\left(\mathbf{p}_v^*\right) \\ \dots & \dots & \dots \\ a_{M1}\left(\mathbf{p}_v^*\right) & \dots & a_{MN}\left(\mathbf{p}_v^*\right) \end{bmatrix}. \tag{20.23}$$

[3] The optimum conditions will give a differentiable mapping $\mathbf{p} = \mathbf{g}(\mathbf{w})$, where $\mathbf{p}$ is the vector of commodity prices and $\mathbf{w}$ is the vector of factor prices. Global invertibility (or univalence) ensures that the inverse mapping $\mathbf{w} = \mathbf{g}^{-1}(\mathbf{p})$ exists uniquely, namely a unique vector of factor prices corresponds to any vector of commodity prices exactly as a unique vector of commodity prices corresponds to any vector of factor prices; note that as we are considering *global* univalence, the conditions stated by the Gale-Nikaidô theorem must be satisfied.

Let $\mathbf{Z}^*$ be the integrated equilibrium output vector whose elements are Z_g^* and let $\mathbf{V}$ be the vector of world endowments whose elements are V_j. Lastly, let $\mathbf{Z}_i$ be the output vector and $\mathbf{V}_i$ the endowment vector of country i $(i = 1, 2)$ whose elements are V_{ji}. The factor price equalization set, denoted Φ, is defined as follows:

$$\Phi \equiv \{\mathbf{V}_i \mid \mathbf{\Delta}^* \mathbf{Z}_i = \mathbf{V}_i, \quad \mathbf{0} \leqslant \mathbf{Z}_i \leqslant \mathbf{Z}^*, \ i = 1, 2\}. \tag{20.24}$$

In words, the FPE set is the set of all possible endowment vectors $\mathbf{V}_i$ such that factors market clear at the integrated equilibrium factors price, and such that the output vector $\mathbf{Z}_i$ is between the null vector and the output vector of the integrated equilibrium. Expression (20.24) defines the set we are searching for. Now we turn to finding it starting from the integrated equilibrium. We begin by assigning arbitrarily to each country a non negative share of the integrated equilibrium output. These shares are denoted λ_{gi}. We then compute the factors endowment needed to produce the arbitrarily chosen quantity of output when the unitary factors input are those of integrated equilibrium; this is:

$$\mathbf{V}_i = \mathbf{\Delta}^* diag\left(\lambda_{gi}\right) \mathbf{Z}^*. \tag{20.25}$$

From linear algebra notation we recall that $diag\left(\lambda_{gi}\right)$ is a diagonal matrix whose elements are λ_{gi}. Let δ_g^* be the g-th column vector of $\mathbf{\Delta}^*$ and let $\mathbf{E}_g$ denote the sectorial employment vector of the integrated equilibrium, $\mathbf{E}_g = \delta_g^* Z_g$. It is now clear that the endowment vector obtained in expressions (20.25) satisfies the requirements of the FPE set defined in (20.24). Therefore

$$\Phi = \{\mathbf{V}_i \mid \exists \lambda_{gi} \geqslant 0, \sum_{i=1}^{2} \lambda_{gi} = 1 \ \forall g, \mathbf{V}_i = \sum_{g=1}^{N} \lambda_{gi} \mathbf{E}_g \ \forall \ i\} \tag{20.26}$$

In words, the FPE set is the set of endowment vectors obtained from all the convex combinations of the integrated equilibrium sectorial employment vectors. The FPE set in Sect. 4.3.2 has been constructed geometrically following expression (20.26), i.e., by finding the surface identified by all the convex combinations of the integrated equilibrium sectorial employment vectors.

20.6 The Heckscher-Ohlin-Vanek Theorem

Assume that there are N goods indexed by g and M factors indexed by j and that $N \geq M \geq 2$. Assume free trade and incomplete specialization. All variables are computed at the free trade equilibrium. Let $\mathbf{p}_v$ be the row vector of factor prices and let s_i be country i's share in world gross domestic product,

$$s_i \equiv \frac{\mathbf{p}_v \mathbf{V}_i}{\mathbf{p}_v \mathbf{V}}. \tag{20.27}$$

Let $\mathbf{C}_i$ and $\mathbf{C}$ be country i's and world's column vectors of consumption. With identical and homothetic preferences and trade balance equilibrium, $\mathbf{C}_i = s_i \mathbf{C}$. Let $\mathbf{T}_i$ be the vector of net exports (exports minus imports); net exports are equal to output minus consumption, thus, $\mathbf{T}_i = \mathbf{Z}_i - \mathbf{C}_i$. Let $\mathbf{F}_i$ denote the factor content of trade vector. The elements of $\mathbf{F}_i$ are the quantity of each factor services needed to produce the net exports of country i at the integrated equilibrium factor prices, $\mathbf{F}_i \equiv \mathbf{\Delta}^*\mathbf{T}_i$. Naturally, some elements of $\mathbf{F}_i$ are positive and some are negative. Free trade equilibrium in goods market requires $\mathbf{C} = \mathbf{Z}$ and equilibrium in factors market requires $\mathbf{\Delta} \boldsymbol{Z}_i = \mathbf{V}_i$. Therefore,

$$\mathbf{F}_i \equiv \mathbf{\Delta}^*\mathbf{T}_i = \mathbf{\Delta}^*\mathbf{Z}_i - \mathbf{\Delta}^*\mathbf{C}_i = \mathbf{V}_i - \mathbf{\Delta}^*\mathbf{C}_i = \mathbf{V}_i - s_i \mathbf{\Delta}^*\mathbf{Z} \tag{20.28}$$

Now, using $\mathbf{\Delta}^*\mathbf{Z} = \mathbf{V}$ we have

$$\mathbf{F}_i = \mathbf{V}_i - s_i \mathbf{V} \tag{20.29}$$

Expression (20.29) is known as the Heckscher-Ohlin-Vanek equation. It shows that the factor content of trade vector of a country is given by the difference between the endowment vector of the country and the world endowment vector, the latter multiplied by the country's share in world GDP. Let v_{ij} be country i's share in world endowment of factor j, i.e., $v_{ij} \equiv V_{ij}/V_j$.

Definition 20.1. A country is abundant in factor j iff

$$v_{ji} > s_i. \tag{20.30}$$

It is immediate from (20.29) and (20.30) that each country is a net exporter of the services of its abundant factors. Indeed, expression (20.29) may be written as

$$\mathbf{F}_i = diag\left(v_{ji} - s_i\right)\mathbf{V} \tag{20.31}$$

which proves the Heckscher-Ohlin-Vanek theorem.

If $M = N = 2$ expression (20.31) becomes

$$\mathbf{F}_i = \begin{bmatrix} (v_{Ki} - s_i)\,\bar{K} \\ (v_{Li} - s_i)\,\bar{L} \end{bmatrix} \tag{20.32}$$

and $s_i = \left(p_K v_{Ki}\bar{K} + p_L v_{Li}\bar{L}\right) / \left(p_K \bar{K} + p_L \bar{L}\right)$. It is clear that

$$\text{if } v_{Li} > v_{Ki}, \quad \text{then } v_{Li} > s_i > v_{Ki}, \tag{20.33}$$

$$\text{if } v_{Li} < v_{Ki}, \quad \text{then } v_{Li} < s_i < v_{Ki}, \tag{20.34}$$

Note that $v_{Li} > v_{Ki}$ ($v_{Li} < v_{Ki}$) implies that country i is *relatively* abundant in factor L (K). Thus, the two-by-two version of the Heckscher-Ohlin-Vanek theorem states that each country is the net exporter of the services of its *relatively* abundant factor, as we have already seen above.

References

Bhagwati, J. N. (1972). The Heckscher-Ohlin theorem in the multi-commodity case.
Brecher, R. A., & Choudhri, E. U. (1982). The factor content of international trade without factor-price equalization.
Chipman, J. S. (1966). A survey of the theory of international trade: Part 3, The Modern Theory.
Deardoff, A. V. (1982). The general validity of the Heckscher-Ohlin theorem.
Deardoff, A. V. (1994). The possibility of factor price equalization, revisited.
Deardoff, A. V. (1979). Weak links in the chain of comparative advantage.
Dixit, A. K., & Woodland, A. (1982). The relationship between factor endowments and commodity trade.
Ethier, W. J. (1984). Higher dimensional trade theory.
Ethier, W. J. (1982). The general role of factor intensity in the theorems of international trade.
Feenstra, R. C. (2004). *Advanced International Trade: Theory and evidence* (chap. 3).
Ford, J. L. (1982). The Ricardian and Heckscher-Ohlin explanations of trade: A general proof of an equivalence theorem and its empirical implications.
Ford, J. L. (1985). Comments on P.J. Lloyd "The Ricardian-Ohlin Explanation of Trade: A Comment on a General Theorem which is not General".
Gale, D., & Nikaidô, H. (1965). The Jacobian matrix and global univalence of mappings.
Gandolfo, G. (1971). *Mathematical methods and models in economic dynamics.*
Harkness, J. P. (1978). Factor abundance and comparative advantage.
Harkness, J. P. (1983). The factor-proportions model with many nations.
Helpman, E. (1984a). The factor content of international trade.
Horiba, Y. (1974). General equilibrium and the Heckscher-Ohlin theory of trade: The multi-country case.
Jones, R. W. (1956). Factor proportions and the Heckscher-Ohlin theorem.
Lloyd, P. J. (1985). The Ricardian and Heckscher-Ohlin explanations of trade: A comment on a general theorem which is not general.
Neary, J. P. (1984). The Heckscher-Ohlin model as an aggregate.
Neary, J. P. (1985a). The observational equivalence of the Ricardian and Heckscher-Ohlin explanations of trade patterns.
Neary, J. P. (1985b). Two-by-two international trade theory with many goods and factors.
Samuelson, P. A. (1953). Prices of factors and goods in general equilibrium.
Svensson, L. E. O. (1984). Factor trade and goods trade.
Vanek, J. (1968). The factor proportions theory.

Chapter 21
Appendix to Chapter 5

21.1 The Factor-Price-Equalization Theorem

A proof of this theorem in its general version can easily be given by using the dual relations due to Jones (1965) and illustrated in Sect. 19.5. If we consider the last two equations of set (19.65) and solve for p_L^*, p_K^*, we get

$$p_L^* = \frac{\theta_{KB} p_A^* - \theta_{KA} p_B^*}{\theta_{LA}\theta_{KB} - \theta_{LB}\theta_{KA}}, \quad p_K^* = \frac{\theta_{LA} p_B^* - \theta_{LB} p_A^*}{\theta_{LA}\theta_{KB} - \theta_{LB}\theta_{KA}}. \tag{21.1}$$

This shows that the prices of factors depend solely on the prices of commodities: as the latter are internationally identical, so also are the former.

Note that when there is complete specialization these relations would not exist: in fact, with complete specialization, either $\theta_{LA} = \theta_{KA} = 0$ (complete specialization in commodity B) or $\theta_{LB} = \theta_{KB} = 0$ (complete specialization in commodity A).

21.2 The Stolper-Samuelson Theorem

The same dual relations given in Eqs. (21.1) allow us to give a simple proof of this theorem (Stolper and Samuelson, 1941) in its general version.

We can assume, without loss of generality, that commodity A is the numéraire, so that $p_A^* = 0$. A positive (negative) value of p_B^* therefore means an increase (decrease) in the relative price p_B/p_A and, likewise, a positive (negative) value of p_L^* means an increase (decrease) in the unit real reward (i.e., in terms of the numéraire) of labour.

Let us now assume, for example, that commodity B is labour-intensive and that the relative price of this commodity increases. Given the definitions of the θ's (see Sect. 19.5), the greater relative labour intensity of B amounts to the inequality $\theta_{LB}/\theta_{KB} > \theta_{LA}/\theta_{KA}$ and, therefore, the denominator of the fractions in (21.1) is

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_21,

negative. As we have assumed $p_B^* > 0$, it follows that $p_L^* > 0$, $p_K^* < 0$. The increase in the unit real reward of the factor used intensively in the industry producing the commodity with the relative price increase is thus proved.

The *magnification effect* is also easily proved. Assuming $p_A^* = 0$, from Eqs. (21.1) we have

$$p_L^* = \frac{\theta_{KA}}{\theta_{LB}\theta_{KA} - \theta_{LA}\theta_{KB}} p_B^*.$$

Now, $p_L^* > p_B^*$ when

$$\frac{\theta_{KA}}{\theta_{LB}\theta_{KA} - \theta_{LA}\theta_{KB}} > 1,$$

which is certainly true. In fact, $\theta_{LB}\theta_{KA} - \theta_{LA}\theta_{KB} < \theta_{KA}$ because this is equivalent to $\theta_{KA}(\theta_{LB} - 1) - \theta_{LA}\theta_{KB} < 0$, which follows from the fact that $\theta_{LB} < 1$.

21.3 The Rybczynski Theorem

A simple proof of the Rybczynski theorem (Rybczynski, 1955) can be given by way of the dual relations illustrated in Sect. 19.5.

From the first two equations in (19.65), we can express A^* and B^* in terms of L^*; since $K^* = 0$ and $p_L^* = p_K^* = 0$ by assumption, we obtain

$$A^* = \frac{\lambda_{KB} L^*}{\lambda_{LA}\lambda_{KB} - \lambda_{KA}\lambda_{LB}}, \quad B^* = \frac{-\lambda_{KA} L^*}{\lambda_{LA}\lambda_{KB} - \lambda_{KA}\lambda_{LB}}. \tag{21.2}$$

If commodity A is labour intensive, the denominator of these expressions is positive and so $A^* > 0$, $B^* < 0$, which proves the theorem.

Besides, given the assumptions, the expression $\lambda_{KB}/(\lambda_{LA}\lambda_{KB} - \lambda_{KA}\lambda_{LB})$ is greater than one, so that $A^* > L^*$ (the *magnification effect*).

This is another example of the fact that the dual approach in various cases enables us to give simple proofs of the fundamental theorems of the pure theory of international trade.

Also note that, if we compare this proof with that of the Stolper-Samuelson theorem given in the previous section, we see that changes in outputs are related to changes in factor endowments through the λ coefficients in the same way as the θ coefficients link factor price changes to commodity price changes. This *duality* between the Rybczynski and Stolper-Samuelson theorems is a basic feature of the general equilibrium model.

To be more precise, the effect of an increase in the endowment of a factor on the output of a commodity (at unchanged prices of factors and goods) is *exactly the same* as the effect of an increase in the price of that commodity (*ceteris paribus*) on that factor's reward. The relations stating the equality of these effects are also called the *reciprocity relations.*

This can be easily checked by using the dual approach. If we compare the results given in Eqs. (21.2) with those of Eqs. (21.1), and substitute the definitions of the λ's and θ's (given in Sect. 19.5) in these expressions, we immediately find the result stated. Alternatively, we could solve the first two equations in (19.61) for $\mathrm{d}A/\mathrm{d}L$, $\mathrm{d}B/\mathrm{d}L$, etc., and the third and fourth for $\mathrm{d}p_L/\mathrm{d}p_A$, $\mathrm{d}p_L/\mathrm{d}p_B$, etc., and find that the resulting expressions are respectively equal.

References

Jones, R. W. (1965). The structure of simple general equilibrium models.
Rybczynski, T. M. (1955). Factor endowment and relative commodity prices.
Stolper, W. F., & Samuelson, P. A. (1941) Protection and real wage.

Chapter 22
Appendix to Chapter 6

22.1 The Specific Factors Model

The specific factors model (Jones, 1971; Samuelson, 1971) can be conveniently examined we follow (Jones, 1971) extending to the present case the treatment already introduced in Sect. 19.5 for the traditional case. Equations (19.59) have to be modified to take account of the presence of specific factors.

Let $a_{ij}, i = K^A, K^B, L; j = A, B$, denote the quantity of factor i required to produce a unit of commodity j. Then we have

$$\begin{aligned} a_{K^A A} A &= K^A, \\ a_{K^B B} B &= K^B, \\ a_{LA} A + a_{LB} B &= L, \end{aligned} \tag{22.1}$$

since $a_{K^A B} \equiv 0, a_{K^B A} \equiv 0$ by the specific factors assumption. These equations emphasize the dual relations between factor endowments and commodity outputs, and derive from full factor employment. We also have the dual relations between commodity prices and factor prices, which derive from competitive equilibrium:

$$\begin{aligned} a_{LA} p_L + a_{K^A A} p_{K^A} &= p_A, \\ a_{LB} p_L + a_{K^B B} p_{K^B} &= p_B. \end{aligned} \tag{22.2}$$

Since, in general, the four input coefficients a_{ij} are variable, the above five equations must be supplemented by four equations to determine these coefficients. Such equations derive from the firm's optimization procedure. As is well known, with constant returns to scale, the input coefficients depend solely upon the factor-price ratio. Therefore

$$a_{ij} = a_{ij} \left(\frac{p_L}{p_{K^j}} \right), \quad i = K^A, K^B, L; \quad j = A, B, \tag{22.3}$$

which are the four equations that we need.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_22,

The nine equations (22.1)–(22.3) describe the production side of the model, and make it possible to determine the nine unknowns a_{ij}, A, B, p_L, p_{K^A}, p_{K^B} given the five parameters L, K^A, K^B, p_A, p_B.

Before going on to examine the equations of change, it is possible to simplify the model by substituting from the first two equations of (22.1) into the third one, thus obtaining

$$\frac{a_{LA}}{a_{K^A A}}K^A + \frac{a_{LB}}{a_{K^B B}}K^B = L. \tag{22.4}$$

Since the a_{ij} depend on factor prices, Eqs. (22.2) and (22.4) provide a set of three equations to determine the three factor prices in terms of the parameters. Let us begin with the total differentials of Eqs. (22.2), that are

$$\begin{aligned} p_L \mathrm{d}a_{LA} + a_{LA}\mathrm{d}p_L + p_{K^A}\mathrm{d}a_{K^A A} + a_{K^A A}\mathrm{d}p_{K^A} &= \mathrm{d}p_A, \\ p_L \mathrm{d}a_{LB} + a_{LB}\mathrm{d}p_L + p_{K^B}\mathrm{d}a_{K^B B} + a_{K^B B}\mathrm{d}p_{K^B} &= \mathrm{d}p_B. \end{aligned} \tag{22.5}$$

If we denote relative changes by an asterisk (namely, $a^*_{LA} = \mathrm{d}a_{LA}/a_{LA}$, etc.) we can rewrite Eqs. (22.5), after simple manipulations (these are the same as shown in the Sect. 19.5) in the form

$$\begin{aligned} \theta_{K^A A} p^*_{K^A} + \theta_{LA} p^*_L &= p^*_A, \\ \theta_{K^B B} p^*_{K^B} + \theta_{LB} p^*_L &= p^*_B, \end{aligned} \tag{22.6}$$

where the θ's denote the factor shares in each sector ($\theta_{K^A A} = a_{K^A A} p_{K^A}/p_A$ etc.), which of course add up to 1, and use has been made of the fact that, as shown in the Eqs. (19.63)

$$\begin{aligned} \theta_{K^A A} a^*_{K^A A} + \theta_{LA} a^*_{LA} &= 0, \\ \theta_{K^B B} a^*_{K^B B} + \theta_{LB} a^*_{LB} &= 0. \end{aligned} \tag{22.7}$$

We next consider the total differential of Eq. (22.4), which is

$$\begin{aligned} &\frac{a_{LA}}{a_{K^A A}}\mathrm{d}K^A + \frac{\mathrm{d}a_{LA}}{a_{K^A A}}K^A - \frac{a_{LA}}{a_{K^A A}}\frac{\mathrm{d}a_{K^A A}}{a_{K^A A}}K^A \\ &+\frac{a_{LB}}{a_{K^B B}}\mathrm{d}K^B + \frac{\mathrm{d}a_{LB}}{a_{K^B B}}K^B - \frac{a_{LB}}{a_{K^B B}}\frac{\mathrm{d}a_{K^B B}}{a_{K^B B}}K^B \\ &= \mathrm{d}L, \end{aligned} \tag{22.8}$$

whence

$$
\begin{aligned}
&\frac{K_A}{L}\left(\frac{a_{LA}}{a_{K^A A}}\mathrm{d}K^A/K^A + \frac{\mathrm{d}a_{LA}}{a_{K^A A}} - \frac{a_{LA}}{a_{K^A A}}\frac{\mathrm{d}a_{K^A A}}{a_{K^A A}}\right)\\
&+\frac{K^B}{L}\left(\frac{a_{LB}}{a_{K^B B}}\mathrm{d}K^B/K^B + \frac{\mathrm{d}a_{LB}}{a_{K^B B}} - \frac{a_{LB}}{a_{K^B B}}\frac{\mathrm{d}a_{K^B B}}{a_{K^B B}}\right)\\
&= \mathrm{d}L/L.
\end{aligned}
$$

Further simple manipulations and use of the definitions of starred variables give

$$
\begin{aligned}
&\lambda_{LA}K^{*A} + \lambda_{LA}a^*_{LA} - \lambda_{LA}a^*_{K^A A}\\
&+\lambda_{LB}K^{*B} + \lambda_{LB}a^*_{LB} - \lambda_{LB}a^*_{K^B B}\\
&= L^*,
\end{aligned}
\tag{22.9}
$$

where $\lambda_{LA} \equiv a_{LA}A/L, \lambda_{LB} \equiv a_{LB}B/L$ denote the fractions of the labour force used in sector A and B respectively. These fractions must of course add up to one, given the full employment condition.

From the definition of elasticity of substitution between factors (see Sect. 19.5) in the two sectors, σ_A, σ_B, we have

$$
\begin{aligned}
a^*_{K^A A} - a^*_{LA} &= \sigma_A(p^*_L - p^*_{K^A}),\\
a^*_{K^B B} - a^*_{LB} &= \sigma_B(p^*_L - p^*_{KB}),
\end{aligned}
\tag{22.10}
$$

which allows us to rewrite (22.9) in the form

$$
\begin{aligned}
&\lambda_{LA}\sigma_A p^*_{K^A} + \lambda_{LB}\sigma_B p^*_{K^B} - (\lambda_{LA}\sigma_A + \lambda_{LB}\sigma_B)\, p^*_L\\
&= \left(L^* - \lambda_{LA}K^{*A} - \lambda_{LB}K^{*B}\right).
\end{aligned}
\tag{22.11}
$$

The system made up of Eqs. (22.6) and (22.11) gives us the equations of change that allow us to determine the effects on factor returns of changes in commodity prices and factor endowments.

Before solving this system, it is as well to observe that it immediately shows why FPE (factor price equalization) does not hold.

In the standard model, two relationships are given to determine two factor prices once commodity prices are known: these are the last two equations in set (19.65). Hence, since commodity prices are internationally identical, with the assumed identical technologies also factor prices are identical across countries. In the present model, the two equations (22.6) are obviously not sufficient to determine the three factor prices in terms of commodity prices only.

Let us now solve our three-equation system. Its determinant is

$$
\begin{aligned}
D &\equiv \begin{vmatrix} \theta_{K^A A} & 0 & \theta_{LA}\\ 0 & \theta_{K^B B} & \theta_{LB}\\ \lambda_{LA}\sigma_A & \lambda_{LB}\sigma_B & -(\lambda_{LA}\sigma_A + \lambda_{LB}\sigma_B)\end{vmatrix}\\
&= -\theta_{K^A A}\{\theta_{K^B B}\,(\lambda_{LA}\sigma_A + \lambda_{LB}\sigma_B) + \theta_{LB}\lambda_{LB}\sigma_B\} - \theta_{LA}\theta_{K^B B}\lambda_{LA}\sigma_A
\end{aligned}
$$

$$= -\lambda_{LA}\sigma_A\theta_{K^B B}\left(\theta_{K^A A} + \theta_{LA}\right) - \lambda_{LB}\sigma_B\theta_{K^A A}\left(\theta_{K^B B} + \theta_{LB}\right)$$
$$= -\theta_{K^A A}\theta_{K^B B}\left(\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}} + \lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}\right), \tag{22.12}$$

where use has been made of the fact that $\theta_{K^A A} + \theta_{LA} = \theta_{K^B B} + \theta_{LB} = 1$.

Simple calculations (for example by Cramer's rule) yield

$$p^*_{K^A} = \frac{1}{\Delta}\left\{\left[\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}} + \frac{1}{\theta_{K^A A}}\lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}\right]p^*_A - \frac{\theta_{LA}}{\theta_{K^A A}}\lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}p^*_B + \frac{\theta_{LA}}{\theta_{K^A A}}\left[L^* - \lambda_{LA}K^{*A} - \lambda_{LB}K^{*B}\right]\right\}, \tag{22.13}$$

$$p^*_{K^B} = \frac{1}{\Delta}\left\{\left[\lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}} + \frac{1}{\theta_{K^B B}}\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}}\right]p^*_B - \frac{\theta_{LB}}{\theta_{K^B B}}\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}}p^*_A + \frac{\theta_{LB}}{\theta_{K^B B}}\left[L^* - \lambda_{LA}K^{*A} - \lambda_{LB}K^{*B}\right]\right\}, \tag{22.14}$$

$$p^*_L = \frac{1}{\Delta}\left\{\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}}p^*_A + \lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}p^*_B + \left[\lambda_{LA}K^{*A} + \lambda_{LB}K^{*B} - L^*\right]\right\}, \tag{22.15}$$

$$p^*_{K^A} - p^*_{K^B} = \frac{1}{\Delta}\left\{\left[\frac{1}{\theta_{K^B B}}\lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}} + \frac{1}{\theta_{K^A A}}\lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}\right]\left(p^*_A - p^*_B\right) + \frac{1}{\theta_{K^A A}\theta_{K^B B}}\left(\theta_{LA} - \theta_{LB}\right)\left[\lambda_{LA}K^{*A} + \lambda_{LB}K^{*B} - L^*\right]\right\}, \tag{22.16}$$

where

$$\Delta \equiv -\frac{D}{\theta_{K^A A}\theta_{K^A A}} = \lambda_{LA}\frac{\sigma_A}{\theta_{K^A A}} + \lambda_{LB}\frac{\sigma_B}{\theta_{K^B B}}. \tag{22.17}$$

Let us observe that the expression $\sigma_j/\theta_{K^j j}$, $j = A, B$, which frequently appears in the above formulae, is the elasticity of the marginal product curve of the mobile factor in sector j. Hence Δ is a weighted average of these elasticities.

It is then an easy matter to show that (a form of) the Stolper-Samuelson theorem holds for the specific factors but not for the mobile factor.

Suppose, for example, that commodity A is the numéraire, so that $p^*_A = 0$. A positive (negative) value of p^*_B then means an increase (decrease) in the relative price p_B/p_A. Consider now an increase in the relative price of commodity B at unchanged factor endowments. From Eqs. (22.13) and (22.14) we see that

$p^*_{K^B} > 0$, $p^*_{K^A} < 0$: the unit real reward of the specific factor used in sector B increases while that of the specific factor used in the other sector decreases. From Eq. (22.15) we see that $p^*_L > 0$, hence the wage rate increases in terms of commodity A.

Let us now take commodity B as the numéraire ($p^*_B = 0$), whereby an increase in the relative price of B means $p^*_A < 0$. From Eqs. (22.13) and (22.14) we again see that $p^*_{K^B} > 0$, $p^*_{K^A} < 0$: the result as regards specific factor rewards is independent of the choice of the numéraire. However, from Eq. (22.15) we see that $p^*_L < 0$, namely the wage rate *decreases* in terms of commodity B. Hence the real wage rate may move in either direction, depending on the composition of the expenditure of wage earners.

The effects of changes in factor endowments on factor prices are unambiguous: a change in any factor endowment causes the return to the mobile factor to change in a direction opposite to the returns to both specific factors. For example, an increase in the labour force has a positive effect on both $p^*_{K^A}$, $p^*_{K^B}$ and a negative effect on p^*_L.

To obtain the comparative statics results concerning outputs, we totally differentiate Eqs. (22.1), thus obtaining

$$\begin{aligned} A\mathrm{d}a_{K^AA} + a_{K^AA}\mathrm{d}A &= \mathrm{d}K^A, \\ B\mathrm{d}a_{K^BB} + a_{K^BB}\mathrm{d}B &= \mathrm{d}K^B, \\ A\mathrm{d}a_{LA} + a_{LA}\mathrm{d}A + B\mathrm{d}a_{LB} + a_{LB}\mathrm{d}B &= \mathrm{d}L. \end{aligned} \tag{22.18}$$

Simple manipulations (see Sect. 19.5) yield

$$\begin{aligned} \lambda_{K^AA}A^* &= K^{*A} - \lambda_{K^AA}a^*_{K^AA}, \\ \lambda_{K^BB}B^* &= K^{*B} - \lambda_{K^BB}a^*_{K^BB}, \\ \lambda_{LA}A^* + \lambda_{LB}B^* &= L^* - (\lambda_{LA}a^*_{LA} + \lambda_{LB}a^*_{LB}). \end{aligned} \tag{22.19}$$

It is possible to express the proportional changes in input coefficients in terms of the proportional changes in factor prices using Eqs. (22.7) and (22.10), whence

$$\left.\begin{aligned} a^*_{Lj} &= -\theta_{K^jj}\sigma_j(p^*_L - p^*_{K^j}), \\ a^*_{K^jj} &= \theta_{Lj}\sigma_j(p^*_L - p^*_{K^j}). \end{aligned}\right. \quad j = A, B. \tag{22.20}$$

Substitution of (22.20) into (22.19) yields

$$\begin{aligned} \lambda_{K^AA}A^* &= K^{*A} - \lambda_{K^AA}\theta_{LA}\sigma_A(p^*_L - p^*_{K^A}), \\ \lambda_{K^BB}B^* &= K^{*B} - \lambda_{K^BB}\theta_{LB}\sigma_B(p^*_L - p^*_{K^B}), \\ \lambda_{LA}A^* + \lambda_{LB}B^* &= L^* + \lambda_{LA}\theta_{K^AA}\sigma_A(p^*_L - p^*_{K^A}) + \lambda_{LB}\theta_{K^BB}\sigma_B(p^*_L - p^*_{K^B}). \end{aligned} \tag{22.21}$$

Differently from the standard 2×2 model, factor prices are not constant in the face of constant commodity prices. In fact, letting $p^*_A = p^*_B = 0$, from Eqs. (22.13) to (22.15) we have

$$\begin{aligned}
p^*_{K^A} &= \frac{1}{\Delta}\frac{\theta_{LA}}{\theta_{K^A A}}[L^* - \lambda_{LA}K^{*A} - \lambda_{LB}K^{*B}],\\
p^*_{K^B} &= \frac{1}{\Delta}\frac{\theta_{LB}}{\theta_{K^B B}}[L^* - \lambda_{LA}K^{*A} - \lambda_{LB}K^{*B}],\\
p^*_L &= \frac{1}{\Delta}[\lambda_{LA}K^{*A} + \lambda_{LB}K^{*B} - L^*],
\end{aligned} \tag{22.22}$$

whence

$$\begin{aligned}
K^{*B} = L^* = 0 \text{ and } K^*_A > 0 &\Longrightarrow p^*_{K^A} < 0, p^*_{K^B} < 0, p^*_L > 0,\\
K^{*A} = L^* = 0 \text{ and } K^*_B > 0 &\Longrightarrow p^*_{K^A} < 0, p^*_{K^B} < 0, p^*_L > 0,\\
K^{*A} = K^*_B = 0 \text{ and } L^* > 0 &\Longrightarrow p^*_{K^A} > 0, p^*_{K^B} > 0, p^*_L < 0.
\end{aligned} \tag{22.23}$$

Equations (22.21) and (22.23) in turn imply

$$\begin{aligned}
K^{*B} = L^* = 0 \text{ and } K^*_A > 0 &\Longrightarrow A^* > 0, B^* < 0,\\
K^{*A} = L^* = 0 \text{ and } K^*_B > 0 &\Longrightarrow A^* < 0, B^* > 0,\\
K^{*A} = K^*_B = 0 \text{ and } L^* > 0 &\Longrightarrow A^* > 0, B^* > 0,
\end{aligned} \tag{22.24}$$

which show that (a form of) the Rybczynski theorem is only valid for specific factors (an increase in a specific factor causes an output increase in the corresponding sector and an output decrease in the other sector) but not for the mobile factor (an increase in the labour force brings about an output increase in both sectors).

22.2 The Cost of Transport

As we pointed out in the text, the rigorous treatment of the cost of transport requires a model which maintains the two-country assumption but with at least four variables present: the two transport services in addition to the two commodities. This takes us at once to a general equilibrium model of the type mentioned in Sect. 3.7. It is of course necessary to add the equations which establish equilibrium between demand and supply on the market for transport services and also the relations stating that exports of a given commodity by a given country occur only if the price in the importing country is equal to that in the exporting country plus the cost of transport. By applying to the resultant model the methods used in mathematical economics to demonstrate the existence of general economic equilibrium in a closed economy, one can see that in effect an equilibrium does exist. The extension of the model of general world equilibrium to a model with more than two countries does not present any further difficulties.

The price to be paid for this generality is, as we have already seen in Sect. 3.7, the loss of the explicative and interpretative power of the model, which does not allow us to establish empirically significant propositions regarding the structure of international trade or the other problems that the pure theory of international trade deals with. For a demonstration of the existence of equilibrium, see Hadley and Kemp (1966) . For further considerations regarding the cost of transport, see Casas (1983) and Casas and Choi (1985b) .

The role of transport cost in bringing about a core-periphery pattern will be examined in Sect. 31.2.

22.3 Intermediate Goods

22.3.1 Final Goods as Inputs

Let us first look at the case in which each product existing in the economy can be used as both an intermediate and final good. For simplicity, we assume that each good enters as an intermediate good only in the production of the other good and let A_B and B_A be respectively the quantity of A used as an intermediate good in the production of B and the quantity of B used in the production of A; with A and B we shall now indicate the *net* quantities of the two goods. We thus have the relationships

$$\begin{aligned} A &= F_A\left(K_A, L_A, B_A\right) - A_{B,} \\ B &= F_B\left(K_B, L_B, A_B\right) - B_A, \end{aligned} \tag{22.25}$$

where F_A and F_B are first-degree homogeneous production functions. Samuelson's theorem states that Eqs. (22.25) can be transformed into the net production functions

$$\begin{aligned} A &= N_A\left(K_A^c, L_A^c\right), \\ B &= N_B\left(K_B^c, L_B^c\right), \end{aligned} \tag{22.26}$$

where K_A^c, L_A^c denote the total quantities of capital and labour (directly and indirectly) required in the production of A as final good, and similarly for K_B^c, L_B^c. On the basis of Eqs. (22.26), each sector may be considered as an *integrated industry*, which produces internally all the intermediate goods (which are not observed from the outside) which are needed to produce the final good. Equations (22.26) are derived from a process of efficient allocation of resources, which consists in maximizing the quantity of the final good that can be obtained with any given combination of total use (direct and indirect) of capital and labour.[1]

Let us consider one of the two integrated industries, for example, that of commodity A (the same argument applies to B). From the point of view of the

[1]Remember that, in general, a production function gives the *maximum* quantity of output for any given combination of inputs. This maximum, in the case of ordinary production functions, such as Eqs. (22.25), is set for us by the state of technology, while in the case we are examining, in which we are trying to cause the intermediate goods to disappear, it is necessary to solve a further problem, that of the efficient allocation of resources.

integrated industry, the other commodity serves solely as an intermediate good, with a production function $B_A = F_B(\ldots)$, so that it is as if we placed $B = 0$ in the second equation of (22.25). The production function of A can therefore be rewritten as

$$A = F_A\left[K_A, L_A, F_B\left(K_A^c - K_A, L_A^c - L_A, A_B\right)\right] - A_{B,} \tag{22.27}$$

since, given the assumptions, made, $K_A^c = K_A + K_B, \quad L_A^c = L_A + L_B$. It is thus a question of maximizing A in (22.27), given K_A^c, L_A^c. The first-order conditions are

$$\begin{aligned} \frac{\partial A}{\partial K_A} &= \frac{\partial F_A}{\partial K_A} - \frac{\partial F_A}{\partial B_A}\frac{\partial F_B}{\partial K_A} &= 0, \\ \frac{\partial A}{\partial L_A} &= \frac{\partial F_A}{\partial L_A} - \frac{\partial F_A}{\partial B_A}\frac{\partial F_B}{\partial L_A} &= 0, \\ \frac{\partial A}{\partial A_B} &= \frac{\partial F_A}{\partial B_A}\frac{\partial F_B}{\partial A_B} - 1 &= 0. \end{aligned} \tag{22.28}$$

The interpretation is very simple: the first two conditions tell us that the marginal productivity (in terms of A) of each primary factor must be the same whether the factor is used directly or indirectly in the production of A (by producing B, which is used as an intermediate good in the production of A). The third condition tells us that the marginal productivity of A in terms of itself (that is, when A is used as an intermediate good to produce B which is used as an intermediate good to produce A) must be equal to one.

The integrated industry is completely described by (22.27) and (22.28). On the basis of the theory of comparative statics, it is possible—provided that the second order conditions for a maximum have been satisfied—to use Eqs. (22.28) to express K_A, L_A, A_B as differentiable functions of the two parameters K_A^c, L_A^c. By substituting these functions in (22.27), we can see that A is ultimately expressed as a function only of K_A^c, L_A^c, that is, $A = N_A(K_A^c, L_A^c)$,[2] which is in fact the first of Eqs. (22.26). The second of Eqs. (22.26) can be obtained in the same way.

22.3.2 Pure Intermediate Goods

Let us now examine the model with a "pure" intermediate good. The first point to be considered is that the classification of goods on the basis of the apparent factor intensity can be different from the classification of goods on the basis of the total factor intensity. If we indicate the pure intermediate good by Z, we get the following

[2]Still making use of the method of comparative statics, it is possible to obtain explicit expressions for the partial derivatives of the N_A function and to show that it is homogeneous of the first degree. See, for example, Chacholiades (1978, pp. 231–232).

equations, which express the full employment of the primary factors and of the intermediate good:

$$
\begin{aligned}
a_{KA}A + a_{KB}B + a_{KZ}Z &= K,\\
a_{LA}A + a_{LB}B + a_{LZ}Z &= L,\\
a_{ZA}A + a_{ZB}B &= Z,
\end{aligned}
\tag{22.29}
$$

where $a_{KA} = K_A/A$ etc., are the apparent technical coefficients. By substituting from the third equation into the previous ones, we get

$$
\begin{aligned}
a^c_{KA}A + a^c_{KB}B &= K,\\
a^c_{LA}A + a^c_{LB}B &= L,
\end{aligned}
\tag{22.30}
$$

where

$$
\begin{aligned}
a^c_{KA} &= a_{KA} + a_{KZ}a_{ZA},\ a^c_{KB} = a_{KB} + a_{KZ}a_{ZB},\\
a^c_{LA} &= a_{LA} + a_{LZ}a_{ZA},\ \ a^c_{LB} = a_{LB} + a_{LZ}a_{ZB},
\end{aligned}
\tag{22.31}
$$

are the total technical coefficients.

Apparent and total factorial intensities are then[3]

$$
\begin{aligned}
&\varrho_A = \frac{a_{KA}}{a_{LA}}, \quad \varrho_B = \frac{a_{KB}}{a_{LB}}, \quad \varrho_Z = \frac{a_{KZ}}{a_{LZ}},\\
&\varrho^c_A = \frac{a^c_{KA}}{a^c_{LA}} = \frac{a_{KA} + a_{KZ}a_{ZA}}{a_{LA} + a_{LZ}a_{ZA}}, \quad \varrho^c_B = \frac{a^c_{KB}}{a^c_{LB}} = \frac{a_{KB} + a_{KZ}a_{ZB}}{a_{LB} + a_{LZ}a_{ZB}}.
\end{aligned}
\tag{22.32}
$$

By introducing the quantities

$$
\gamma_A = \frac{a_{LZ}a_{ZA}}{a_{LA} + a_{LZ}a_{ZA}}, \quad \gamma_B = \frac{a_{LZ}a_{ZB}}{a_{LB} + a_{LZ}a_{ZB}},
\tag{22.33}
$$

it is possible to express the total factor intensities of A and B as weighted averages of the respective apparent intensities and of the factor intensity of Z, that is

$$
\begin{aligned}
\varrho^c_A &= (1-\gamma_A)\,\varrho_A + \gamma_A\varrho_Z,\\
\varrho^c_B &= (1-\gamma_B)\,\varrho_B + \gamma_B\varrho_Z,
\end{aligned}
\tag{22.34}
$$

as can be ascertained by direct substitution. Given the properties of the average, ϱ^c_A will be included between ϱ_A and ϱ_Z, and ϱ^c_B between ϱ_B and ϱ_Z. Thus, if ϱ_Z is included between ϱ_A and ϱ_B, the classification based on total intensities coincides with that based on apparent intensities. In fact, if $\varrho_A > \varrho_Z > \varrho_B$ then, as ϱ^c_A is

[3]As the intermediate good is produced exclusively with primary factors, it shows no distinction between apparent and total coefficients or between apparent and total factor intensities.

included between ϱ_A and ϱ_Z while ϱ_B^c is included between ϱ_Z and ϱ_B, ϱ_A is also greater than ϱ_B^c; likewise if $\varrho_A < \varrho_Z < \varrho_B$.

On the other hand, whenever ϱ_Z is not included between ϱ_A and ϱ_B it is possible for the classification based on total intensity to be different from that based on apparent intensity, giving rise to the problems mentioned in the text.[4]

We now pass to the demonstration of the theorem stated in Sect. 6.4, according to which, if it is assumed that one of the three goods is non-traded and the apparent capital intensity of this good is intermediate between the apparent intensities of the two traded goods, then the Heckscher-Ohlin theorem is valid. For this purpose we use the dual approach (see Sect. 19.5) extended to our case. As the majority of empirical studies regarding intermediate goods take as reference Leontief's input-output model, in which the input coefficients of intermediate goods are assumed to be constant, we too adopt this simplification. The coefficients a_{ZA}, a_{ZB}, are therefore assumed to be constant. The price equations are

$$\begin{aligned} a_{LA}p_L + a_{KA}p_K + a_{ZA}p_Z &= p_A, \\ a_{LB}p_L + a_{KB}p_K + a_{ZB}p_Z &= p_B, \\ a_{LZ}p_L + a_{KZ}p_K \qquad\quad &= p_Z, \end{aligned} \tag{22.35}$$

from which, calculating the total differentials, assuming A as numéraire (whence $\mathrm{d}p_A = 0$) and rearranging terms, we have

$$\begin{aligned} a_{LA}\mathrm{d}p_L + a_{KA}\mathrm{d}p_K + a_{ZA}\mathrm{d}p_Z &= -\left(p_L\mathrm{d}a_{LA} + p_K\mathrm{d}a_{KA}\right), \\ a_{LB}\mathrm{d}p_L + a_{KB}\mathrm{d}p_K + a_{ZB}\mathrm{d}p_Z &= \mathrm{d}p_B - \left(p_L\mathrm{d}a_{LB} + p_K\mathrm{d}a_{KB}\right), \\ a_{LZ}\mathrm{d}p_L + a_{KZ}\mathrm{d}p_K - \mathrm{d}p_Z \qquad &= -\left(p_L\mathrm{d}a_{LZ} + p_K\mathrm{d}a_{KZ}\right). \end{aligned} \tag{22.36}$$

As the minimum cost conditions imply that $p_L\mathrm{d}a_{Li} + p_K\mathrm{d}a_{Ki} = 0, i = A, B, Z$, the terms in brackets on the right-hand side of (22.36) disappear. If we now solve this system, we obtain

$$\begin{aligned} \mathrm{d}p_L &= \frac{a_{KA} + a_{ZA}a_{KZ}}{D}\mathrm{d}p_B, \\ \mathrm{d}p_K &= \frac{-\left(a_{LA} + a_{ZA}a_{LZ}\right)}{D}\mathrm{d}p_B, \\ \mathrm{d}p_Z &= \frac{a_{LA} + a_{LZ}\left(\varrho_A - \varrho_Z\right)}{D}\mathrm{d}p_B, \end{aligned} \tag{22.37}$$

[4] For the two classifications to coincide even in this case, it is necessary for the final commodity, with a capital/labour ratio between the capital/labour ratio of the intermediate good and the capital/labour ratio of the other final good, to have an intensity of use of the intermediate good equal to or greater than that of the other final good. This can be demonstrated by starting from Eqs. (22.32) and afterwards examining the appropriate inequalities.

It is as well at this point to note that, in the model previously examined (A and B are used both as final and intermediate goods) the two classifications necessarily coincide: see Vanek (1963).

where

$$D \equiv a_{LA}a_{LB}\left(\varrho_A-\varrho_B\right)+a_{ZA}a_{LB}a_{LZ}\left(\varrho_Z-\varrho_B\right)+a_{ZB}a_{LA}a_{LZ}\left(\varrho_A-\varrho_Z\right). \tag{22.38}$$

Let us now assume that country 1 is relatively capital abundant in economic terms (see Sect. 4.2), that is, $q_1 > q_2$, where $q = p_L/p_K$. We then begin to consider the case in which the intermediate good is non-traded. As

$$\frac{\mathrm{d}q}{\mathrm{d}p_B}=\left(p_K\frac{\mathrm{d}p_L}{\mathrm{d}p_B}-p_L\frac{\mathrm{d}p_K}{\mathrm{d}p_B}\right)/p_K^2,$$

given Eqs. (22.37) we have

$$\frac{\mathrm{d}q}{\mathrm{d}p_B}=\frac{1}{p_K^2 D},$$

and therefore, assuming that D will be different from zero

$$\frac{\mathrm{d}p_B}{\mathrm{d}q}=p_K^2 D. \tag{22.39}$$

We must remember that, having used A as numéraire ($p_A = 1$), p_B is in effect the relative price of the final goods. Equation (22.39) therefore expresses the relationship between the relative price of the final goods and the relative price of the factors, which must be single-valued for the Heckscher-Ohlin theorem to be valid. In fact, it is necessary that a different relative price of goods in autarky corresponds uniquely to a different relative factor endowment (in economic terms). If, for example, $D > 0$, we have $\mathrm{d}p_B/\mathrm{d}q > 0$ and, with $q_1 > q_2$, this means that $(p_B)_1 > (p_B)_2$ in autarky, so that on opening international trade (which determines a single common price lying between the two autarkic prices) country 2 will export B and country 1 will export A. Does this conform to the Heckscher-Ohlin theorem? The answer is yes, provided that ϱ_Z is included between ϱ_A and ϱ_B. In fact, with $\varrho_A > \varrho_Z > \varrho_B$ we have $D > 0$ and country 1 in fact exports the capital intensive good. Similarly, with $\varrho_B > \varrho_Z > \varrho_A$ we have $D < 0$, and given (22.39), it follows that $(p_B)_1 < (p_B)_2$, so that country 1 will export B, which is now the capital-intensive one. Thus, in any case in which the capital intensity of the intermediate good (which, as we have assumed, is non-traded) is included between those of the two traded goods, it is true that country 1, with a relatively high capital endowment (in economic terms) will export the more capital-intensive good in conformity with the Heckscher-Ohlin theorem.

Let us now examine the case in which the non-traded good is one of the two final goods, for example, A. It is now necessary to find an expression which will give us the derivative of the relative price p_Z/p_B with respect to q and establish the conditions under which it has a unique sign in relation to the factor intensities. Since

$$\frac{\mathrm{d}\left(p_Z/p_B\right)}{\mathrm{d}q} = \frac{1}{p_B^2}\left(p_B\frac{\mathrm{d}p_Z}{\mathrm{d}q} - p_Z\frac{\mathrm{d}p_B}{\mathrm{d}q}\right), \tag{22.40}$$

the procedure consists in calculating $\mathrm{d}p_Z/\mathrm{d}q$, as $\mathrm{d}p_B/\mathrm{d}q$ is already known from (22.39). This calculation can be carried out if one notes that $\mathrm{d}p_Z/\mathrm{d}q = (\mathrm{d}p_Z/\mathrm{d}p_B)(\mathrm{d}p_B/\mathrm{d}q)$ and if one uses Eq. (22.37) to determine $\mathrm{d}p_Z/\mathrm{d}p_B$. We refer the reader to Batra and Casas (1973, p. 307) for the details, and we shall limit our observations to the fact that (22.40) will certainly have a unique sign, if the factor intensity of the non-traded good A is intermediate between that of B and that of Z. More precisely, we have

$$\frac{\mathrm{d}\left(p_Z/p_B\right)}{\mathrm{d}q} > 0 \quad \text{if} \quad \varrho_B > \varrho_A > \varrho_Z, \tag{22.41}$$

and therefore, given that $q_1 > q_2$ we get $(p_Z/p_B)_1 > (p_Z/p_B)_2$, so that country 2 will export Z and country 1 will export B (which is more capital-intensive than Z), in conformity with the Heckscher-Ohlin theorem.

Similarly, it can be seen that, as

$$\frac{\mathrm{d}\left(p_Z/p_B\right)}{\mathrm{d}q} < 0 \quad \text{if} \quad \varrho_B < \varrho_A < \varrho_Z, \tag{22.42}$$

the assumption $q_1 > q_2$ implies $(p_Z/p_B)_1 < (p_Z/p_B)_2$, so that country 1 will export Z (which is now the most capital-intensive).

This completes the demonstration of the theorem given in the text. For other approaches to trade in intermediate goods, see, e.g., Sanyal and Jones (1982) and Sarkar (1985).

22.4 Elastic Supply of Factors

We propose to examine formally the behaviour of the offer curve of a country with the endogenous variation of the supply of labour, with the aim of ascertaining the conditions under which this curve will have an anomalous shape. This means ascertaining the conditions under which the country increases its demand for imports when their price is greater. Let us assume that the country concerned imports B.[5] The demand for imports will be given by the difference between domestic demand and domestic production of the commodity in question:

$$E_B = B^D\left(I_A, p\right) - B\left(p, L\right), \tag{22.43}$$

[5] In the text we assumed that A is the imported commodity, but this has no effect on the conclusions.

where the meaning of the symbols is as given in the Chap. 19: the one thing to note here is that since the supply of labour is variable, the quantity of goods produced is also a function of employment L, in addition to being a function of the relative price $p = p_B/p_A$. We shall now calculate the total derivative of E_B with respect to p, bearing in mind that

$$I_A = A\,(p, L) + pB\,(p, L)\,, \tag{22.44}$$

and that employment L is also a function of p through the labour market. Thus we have

$$\begin{aligned}\frac{\mathrm{d}E_B}{\mathrm{d}p} &= \frac{\partial B^D}{\partial I_A}\frac{\mathrm{d}I_A}{\mathrm{d}p} + \frac{\partial B^D}{\partial p} - \frac{\partial B}{\partial p} - \frac{\partial B}{\partial L}\frac{\mathrm{d}L}{\mathrm{d}p}\\ &= \frac{\partial B^D}{\partial I_A}\left(\frac{\partial I_A}{\partial p} + \frac{\partial I_A}{\partial L}\frac{\mathrm{d}L}{\mathrm{d}p}\right) + \frac{\partial B^D}{\partial p} - \frac{\partial B}{\partial p} - \frac{\partial B}{\partial L}\frac{\mathrm{d}L}{\mathrm{d}p}.\end{aligned} \tag{22.45}$$

We now recall that $\partial I_A/\partial p = B$ (see Eq. (19.22)) and that (see Eq. (27.36)) $\partial I_A/\partial L = g_A - \varrho_A g_A' > 0$; defining the marginal propensity to import A as $\mu \equiv p(\partial B^D/\partial I_A)$, we get

$$\frac{\mathrm{d}E_B}{\mathrm{d}p} = \frac{\partial B^D}{\partial p} + \frac{\mu}{p}\left[B + \left(g_A - \varrho_A g_A'\right)\frac{\mathrm{d}L}{\mathrm{d}p}\right] - \left(\frac{\partial B}{\partial p} + \frac{\partial B}{\partial L}\frac{\mathrm{d}L}{\mathrm{d}p}\right). \tag{22.46}$$

If we note that, on the basis of (27.33), $\partial B/\partial L = -\varrho_A g_B/\left(\varrho_B - \varrho_A\right)$ and rearrange the terms, we get

$$\frac{\mathrm{d}E_B}{\mathrm{d}p} = \left(\frac{\partial B^D}{\partial p} + \frac{\mu}{p}B - \frac{\partial B}{\partial p}\right) + \left[\left(g_A - \varrho_A g_A'\right)\frac{\mu}{p} + \frac{\varrho_A g_B}{\varrho_B - \varrho_A}\right]\frac{\mathrm{d}L}{\mathrm{d}p}. \tag{22.47}$$

As, on the basis of (19.17), $g_A - \varrho_A g_A' = p\left(\varrho_B - \varrho_B g_B'\right)$, we can rewrite the expression in square brackets appropriately and we finally get

$$\begin{aligned}\frac{\mathrm{d}E_B}{\mathrm{d}p} &= \left(\frac{\partial B^D}{\partial p} + \frac{\mu}{p}B - \frac{\partial B}{\partial p}\right)\\ &\quad + \left[\left(g_B - \varrho_B - \varrho_B g_B'\right)\left(\mu - \frac{\varrho_A}{\varrho_A - \varrho_B}\frac{g_B}{g_B - \varrho_B g_B'}\right)\right]\frac{\mathrm{d}L}{\mathrm{d}p}.\end{aligned} \tag{22.48}$$

In the case where the supply of labour is constant, $\mathrm{d}L/\mathrm{d}p = 0$ and the derivative $\mathrm{d}E_B/\mathrm{d}p$ will be given by the first expression in parentheses, which we shall assume to be negative, given the assumption that the basic offer curve is normal.

In the case where the supply of labour is endogenously variable, $\mathrm{d}L/\mathrm{d}p \neq 0$ and the expression in square brackets also comes into play. Let us suppose that B is the labour-intensive commodity, so that $\varrho_A > \varrho_B$. We thus get $\varrho_A/\left(\varrho_A - \varrho_B\right) > 1$; and

$g_B/\left(g_B - \varrho_B g_B'\right)$ is also a magnitude greater than one. Under normal conditions $0 < \mu < 1$, and therefore the expression in square brackets is negative. It can at once be seen that, if $\mathrm{d}L/\mathrm{d}p$ is negative, it is possible that $\mathrm{d}E_B/\mathrm{d}p > 0$, i.e., that the country under consideration demands more imports when their price is higher.

The economic meaning of $\mathrm{d}L/\mathrm{d}p < 0$ has already be clarified in Sect. 6.5: when p increases, the real reward of labour grows (Stolper-Samuelson theorem) and, as long as it lies along the backward-bending branch of the labour supply curve, the supply decreases. The reader can obtain further information, for example, in Kemp (1969b, chap. 5); an alternative approach to the one followed here will be found in Laffer and Miles (1982, chap. 8) .

The fact that $\mathrm{d}L/\mathrm{d}p < 0$ can lead to anomalous (or, as some would say, perverse) quantity-price relations is therefore a theoretically admissible possibility; however, some theorists argue against the probability of this actually occurring (Martin & Neary, 1980) .

22.5 Non-traded Goods

Let us take three goods, A, B, N, of which the first is imported, the second exported and the third non-traded; consequently, excess demand for A is positive, for B negative, and for N zero. The production functions have the usual properties (first-degree homogeneity, etc.), so that

$$A = L_A g_A\left(\varrho_A\right), \quad B = L_B g_B\left(\varrho_B\right), \quad N = L_N g_N\left(\varrho_N\right). \tag{22.49}$$

We also have, in equilibrium—see Eqs. (19.17)—that

$$\begin{aligned} &g_A' = p g_B' = p_n g_N', \\ &g_A - \varrho_A g_A' = p\left(g_B - \varrho_B g_B'\right) = p_n\left(g_N - \varrho_N g_N'\right), \end{aligned} \tag{22.50}$$

where $p = p_B/p_A$ and $p_n = p_N/p_A$ are the relative prices.

Given the existing quantities of factors, the full employment conditions are

$$\sum_i L_i = L, \quad \sum_i K_i = \sum_i \varrho_i L_i = K, \quad i = A, B, N. \tag{22.51}$$

Let us now assume that the prices of the two non-traded goods (or their relative price) are given and let us consider the following equations, which are a sub-set of Eqs. (22.50)

$$\begin{aligned} g_A' - p g_B' \qquad\qquad &= 0, \\ g_A - \varrho_A g_A' - p\left(g_B - \varrho_B g_B'\right) &= 0. \end{aligned} \tag{22.52}$$

These allow us to express ϱ_A, ϱ_B as single-valued functions of p, as already seen in Eqs. (20.20) and (20.21).

If we now consider the sub-system

$$\begin{aligned} g_A' - p_n g_N' &= 0, \\ \left(g_A - \varrho_A g_A'\right) - p_n \left(g_N - \varrho_N g_N'\right) &= 0, \end{aligned} \tag{22.53}$$

we can solve it—provided there is no factor intensity reversal so that its Jacobian is other than zero—obtaining uniquely ϱ_N and p_n as functions of ϱ_A and therefore of p, which demonstrates that, in the context of the traditional model, the price of the non-traded good is uniquely determined by the terms of trade.

22.5.1 The Behaviour of the Offer Curve

Let us now go on to examine the behaviour of the offer curve in order to ascertain the conditions under which the country under consideration has an increased demand for imports when their price is higher. We therefore have to calculate the derivative

$$\frac{\mathrm{d}E_A}{\mathrm{d}p} = \frac{\mathrm{d}A^D}{\mathrm{d}p} - \frac{\mathrm{d}A}{\mathrm{d}p}.$$

We begin with the observation that the production of A is no longer, as in the two-commodity model, an increasing function of p, because of the fact that, following the variations of p, p_n also varies and therefore shifts of resources occur between sector N and sector A. In order to calculate $\mathrm{d}A/\mathrm{d}p$ it is therefore necessary to take account of all these effects.[6] If we start from the production functions (22.49) we get

$$\frac{\mathrm{d}A}{\mathrm{d}p} = \frac{\mathrm{d}}{\mathrm{d}p}\left[L_A g_A\left(\varrho_A\right)\right] = g_A \frac{\mathrm{d}L_A}{\mathrm{d}p} + L_A g_A' \frac{\mathrm{d}\varrho_A}{\mathrm{d}p}. \tag{22.54}$$

Let us now calculate the derivatives $\mathrm{d}L_A/\mathrm{d}p$ and $\mathrm{d}\varrho_A/\mathrm{d}p$. As we shall see, when calculating $\mathrm{d}L_A/\mathrm{d}p$ we shall also calculate $\mathrm{d}\varrho_A/\mathrm{d}p$.

By differentiating Eqs. (22.51) with respect to p, we get

$$\begin{aligned} \frac{\mathrm{d}L_A}{\mathrm{d}p} + \frac{\mathrm{d}L_B}{\mathrm{d}p} + \frac{\mathrm{d}L_N}{\mathrm{d}p} &= 0, \\ \varrho_A \frac{\mathrm{d}L_A}{\mathrm{d}p} + L_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} + \varrho_B \frac{\mathrm{d}L_B}{\mathrm{d}p} + L_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} + \varrho_N \frac{\mathrm{d}L_N}{\mathrm{d}p} + L_N \frac{\mathrm{d}\varrho_N}{\mathrm{d}p} &= 0, \end{aligned} \tag{22.55}$$

---

[6] Since these effects also involve the demand for N—as we shall find from (22.65)—it can be seen at once that it is now no longer possible, as in the traditional model given in Chap. 3 and Appendix, to consider the productive side of the model separately from the demand side.

from which

$$\begin{aligned} \frac{\mathrm{d}L_A}{\mathrm{d}p} + \frac{\mathrm{d}L_B}{\mathrm{d}p} &= -\frac{\mathrm{d}L_N}{\mathrm{d}p}, \\ \varrho_A \frac{\mathrm{d}L_A}{\mathrm{d}p} + \varrho_B \frac{\mathrm{d}L_B}{\mathrm{d}p} &= -\varrho_N \frac{\mathrm{d}L_N}{\mathrm{d}p} - \sum_i L_i \frac{\mathrm{d}\varrho_i}{\mathrm{d}p}, \end{aligned} \tag{22.56}$$

and therefore

$$\frac{\mathrm{d}L_A}{\mathrm{d}p} = -\frac{\varrho_B - \varrho_N}{\varrho_B - \varrho_A} \frac{\mathrm{d}L_N}{\mathrm{d}p} + \frac{1}{\varrho_B - \varrho_A} \sum_i L_i \frac{\mathrm{d}\varrho_i}{\mathrm{d}p}. \tag{22.57}$$

The value of $\mathrm{d}L_A/\mathrm{d}p$ is therefore dependent on $\mathrm{d}L_N/\mathrm{d}p$ and $\mathrm{d}\varrho_i/\mathrm{d}p$, $i = A, B, N$. To get $\mathrm{d}L_N/\mathrm{d}p$ one has merely to start from the condition of equilibrium in the market of the non-traded good, $L_N g_N(\varrho_N) = N^D$ from which, on the basis of the implicit-function rule,

$$\frac{\mathrm{d}L_N}{\mathrm{d}p} = -\frac{1}{g_N} \left(\frac{\mathrm{d}N^D}{\mathrm{d}p} - L_N g'_N \frac{\mathrm{d}\varrho_N}{\mathrm{d}p} \right). \tag{22.58}$$

As for the derivatives $\mathrm{d}\varrho_i/\mathrm{d}p$, one has simply to start out from the conditions of equilibrium given in Eqs. (22.50), and calculate the total derivative thereof with respect to p; solving the consequent system, we get $\mathrm{d}\varrho_i/\mathrm{d}p$. A simpler alternative is to determine $\mathrm{d}\varrho_A/\mathrm{d}p$ and $\mathrm{d}\varrho_B/\mathrm{d}p$ by differentiating system (22.52) and then to calculate $\mathrm{d}\varrho_N/\mathrm{d}p$ by differentiating system (22.53).

The second method gives us

$$\begin{aligned} & g''_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} - g'_B - p g''_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} = 0, \\ & g'_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} - \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} g'_A - \varrho_A g''_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} - \left(g_B - \varrho_B g'_B \right) \\ & \quad - p \left(g'_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} - g'_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} - \varrho_B g''_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} \right) = 0, \end{aligned} \tag{22.59}$$

and, if we simplify and rearrange the terms, we get

$$\begin{aligned} g''_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} - p g''_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} &= g'_B, \\ -\varrho_A g''_A \frac{\mathrm{d}\varrho_A}{\mathrm{d}p} + p \varrho_B g''_B \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} &= g_B - \varrho_B g'_B. \end{aligned} \tag{22.60}$$

By solving, we obtain

$$\begin{aligned}\frac{\mathrm{d}\varrho_A}{\mathrm{d}p} &= \frac{g_B}{g_A''\left(\varrho_B-\varrho_A\right)},\\ \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} &= \frac{g_A}{p^2 g_B''\left(\varrho_B-\varrho_A\right)}.\end{aligned} \tag{22.61}$$

Similarly, if we differentiate system (22.53) with respect to p, simplifying and rearranging the terms gives us

$$\begin{aligned} g_A''\frac{\mathrm{d}\varrho_A}{\mathrm{d}p} - p_n g_N''\frac{\mathrm{d}\varrho_N}{\mathrm{d}p} &= g_N'\frac{\mathrm{d}p_n}{\mathrm{d}p},\\ -\varrho_A g_A''\frac{\mathrm{d}\varrho_A}{\mathrm{d}p} + p_n\varrho_N g_N''\frac{\mathrm{d}\varrho_N}{\mathrm{d}p} &= \left(g_N-\varrho_N g_N'\right)\frac{\mathrm{d}p_n}{\mathrm{d}p},\end{aligned} \tag{22.62}$$

from which, by solving,

$$\begin{aligned}\frac{\mathrm{d}\varrho_A}{\mathrm{d}p} &= \frac{g_N}{g_A''\left(\varrho_N-\varrho_A\right)}\frac{\mathrm{d}p_n}{\mathrm{d}p},\\ \frac{\mathrm{d}\varrho_B}{\mathrm{d}p} &= \frac{g_N}{p_n^2 g_N''\left(\varrho_N-\varrho_A\right)}\frac{\mathrm{d}p_n}{\mathrm{d}p},\end{aligned} \tag{22.63}$$

where the first of Eqs. (22.61) and the first of Eqs. (22.63) must naturally coincide, a fact that enables us to determine the derivative of p_n with respect to p:

$$\frac{\mathrm{d}p_n}{\mathrm{d}p} = \frac{\left(\varrho_N-\varrho_A\right)g_B}{\left(\varrho_B-\varrho_A\right)g_N}. \tag{22.64}$$

We now have all the elements necessary to determine $\mathrm{d}A/\mathrm{d}p$, by substituting in Eq. (22.54) the results obtained by means of Eqs. (22.57), (22.58), (22.61) and (22.63). We thus get

$$\frac{\mathrm{d}A}{\mathrm{d}p} = -\frac{g_A}{g_N}\frac{\varrho_B-\varrho_N}{\varrho_B-\varrho_A}\frac{\mathrm{d}N^D}{\mathrm{d}p} + H, \tag{22.65}$$

where

$$H \equiv \frac{pL_N g_A g_B^2}{p_n^3 g_N g_N''\left(\varrho_B-\varrho_A\right)^2} + \frac{pL_A g_B^2}{g_A''\left(\varrho_B-\varrho_A\right)^2} + \frac{L_B g_A^2}{p^2 g_B^2\left(\varrho_B-\varrho_A\right)^2} < 0. \tag{22.66}$$

Term H tends therefore to make $\mathrm{d}A/\mathrm{d}p$ take on the right sign for the normality of the offer curve. However, we also have to take into account the first terms on the right-hand side of Eq. (22.65), which may very well be positive and of a higher absolute value than H, so that $\mathrm{d}A/\mathrm{d}p > 0$ (the economic meaning of this apparently anomalous sign has been clarified in Sect. 6.6). Even without determining the sign of $\mathrm{d}A^D/\mathrm{d}p$ (which can, in turn, be anomalous: the reader can consult Komiya, 1967; Kemp, 1969b, chap. 6), this is sufficient to establish the possibility that $\mathrm{d}E_A/\mathrm{d}p < 0$

that is, $\mathrm{d}E_A/\mathrm{d}\,(1/p) > 0$; this result means that the demand for imports can rise with the rise in the price of imports $p_A/p_B = 1/p$.

22.6 Specific Factors and De-industrialization

Following Corden and Neary (1982) , we shall analyse the problem by means of the dual approach (see Sect. 19.5), appropriately extended to the case of three goods and modified so as to take into account the presence of specific factors (see also Jones, 1971). Bearing in mind that labour is the only mobile factor between sectors and is fully employed, we get the equation

$$a_{LA}A + a_{LB}B + a_{LN}N = L, \tag{22.67}$$

where a_{Li}, $i = A, B, N$, are the technical coefficients. The demand for the non-traded good N (the market for which is constantly in equilibrium) is a function of real national income y and of the price p_N; for simplicity, we shall neglect the effects of the prices of the other goods, and of income distribution.[7] Using the asterisk to indicate proportional variations, we have

$$N^{D^*} = -\varepsilon_N p_N^* + \eta y^*, \tag{22.68}$$

where ε_N and η are the price elasticity and the income elasticity of demand respectively.

In this model the only source of increase in real income is technical progress in the extractive sector which generates the boom, so that

$$y^* = \theta_A \pi, \tag{22.69}$$

where θ_A is the share of the extractive sector in national income and π is the Hicksian measure of technical progress. By substituting (22.69) in (22.68) we have

$$N^{D^*} = -\varepsilon_N p_N^* + \eta\theta_A\pi. \tag{22.70}$$

If we indicate the specific capital of each sector with K_i, it is necessary to add the full employment conditions of each specific factor, that is

$$a_{KA}A = K_A, \quad a_{KB}B = K_B, \quad a_{KN}N = K_N. \tag{22.71}$$

[7] For the complications introduced by the effects that a changed income distribution at a constant price of N has on spending on N see Corden (1984a, fn. 5 on p. 361).

If we differentiate Eq. (22.67) and transform the result into proportional variations, by following the procedure illustrated in Sect. 19.5 (bearing in mind that now L is constant), we have

$$\lambda_{LA}\left(A^* + a_{LA}^*\right) + \lambda_{LB}\left(B^* + a_{LB}^*\right) + \lambda_{LN}\left(N^* + a_{LN}^*\right) = 0, \tag{22.72}$$

where $\lambda_{LA} = a_{Li}A/L$, etc., denote the fractions of the total labour force employed in the various sectors. Following the usual procedure, from Eqs. (22.71) we get

$$a_{KA}^* + A^* = 0, \quad a_{KB}^* + B^* = 0, \quad a_{KN}^* + N^* = 0, \tag{22.73}$$

and by substituting in Eqs. (22.72), we obtain

$$\lambda_{LA}\left(a_{LA}^* - a_{KA}^*\right) + \lambda_{LB}\left(a_{LB}^* - a_{KB}^*\right) + \lambda_{LN}\left(a_{LN}^* - a_{KN}^*\right) = 0. \tag{22.74}$$

From Eqs. (19.63) we have

$$a_{Li}^* - a_{Ki}^* = -\sigma_i\left(p_L^* - p_{Ki}^*\right), \quad i = A, B, N, \tag{22.75}$$

where σ_i is the elasticity of substitution in sector i. As labour is mobile, p_L^* is equal throughout, while the p_{Ki}^* are specific for each sector. From the equality between price and unit cost—see Eqs. (19.59) and (19.62)—account being taken of the technical progress factor and using B as numéraire, we have

$$\begin{aligned} p_A^* &= \theta_{LA}p_L^* + \theta_{KA}p_{KA}^* - \pi, \\ 0 &= \theta_{LB}p_L^* + \theta_{KB}p_{KB}^*, \\ p_N^* &= \theta_{LN}p_L^* + \theta_{KN}p_{KN}^*, \end{aligned} \tag{22.76}$$

where $\theta_{LA} = a_{LA}p_L/p_A$ etc. is the share of labour in the value of output in sector A and so on. By substituting Eqs. (22.75) and (22.76) in (22.74), assuming that $p_A^* = 0$ as the price of good A is given by the international market and simplifying, we have

$$p_L^* = \xi_A\pi + \xi_N p_N^*, \tag{22.77}$$

where $0 < \xi_i < 1$ is the proportional contribution of sector i to Δ, the elasticity with respect to wages of the aggregate demand for labour:

$$\begin{aligned} \xi_i &\equiv \frac{1}{\Delta}\lambda_{Li}\frac{\sigma_i}{\theta_{Ki}}, i = A, B, N, \\ \Delta &\equiv \lambda_{LA}\frac{\sigma_A}{\theta_{KA}} + \lambda_{LB}\frac{\sigma_B}{\theta_{KB}} + \lambda_{LN}\frac{\sigma_N}{\theta_{KN}}. \end{aligned} \tag{22.78}$$

Turning now to the market for N, supply depends solely on the real wage which entrepreneurs have to meet in this sector. In fact, as K_N is assumed fully employed and immobile, the quantity of N produced will depend on the quantity of labour utilized, which in turn is a function of the real wage,[8] following the optimization principle, according to which the entrepreneur equates the marginal productivity of labour to the real wage. Thus, if, as usual, we consider the proportional variations, we get

$$N^* = \Phi_N \left(p_N^* - p_L^*\right), \tag{22.79}$$

where $\Phi_N \equiv \sigma_N \theta_{LN} / \theta_{KN}$ is the price elasticity of supply.

By equating demand (22.68) and supply (22.79), we obtain

$$\left(\Phi_N + \varepsilon_N\right) p_N^* = \Phi_N p_L^* + \eta \theta_A \pi. \tag{22.80}$$

We can now solve the system made up of Eqs. (22.77) and (22.80) for the unknowns p_N^* and p_L^*, obtaining

$$\begin{aligned} Hp_N^* &= \left(\eta\theta_N + \Phi_N \xi_A\right) \pi > 0, \\ Hp_L^* &= \left[\eta \xi_N \theta_A + \left(\Phi_N + \varepsilon_N\right) \xi_A\right] \pi > 0, \end{aligned} \tag{22.81}$$

where

$$H \equiv \Phi_N \left(1 - \xi_N\right) + \varepsilon_N > 0. \tag{22.82}$$

22.6.1 Effects on Prices, Outputs and Factor Rewards

Relations (22.81) confirm what was said in Sect. 6.7, namely, *that both the relative price of N and the real wage increase*.

In order to see how the production of N varies it is sufficient to substitute p_N^* and p_L^* from (22.81) into (22.79), thus obtaining

$$N^* = \left(\Phi_N / H\right) \left[\eta \theta_A \left(1 - \xi_N\right) - \xi_A \varepsilon_N\right]. \tag{22.83}$$

As can be seen, N can be either positive or negative (i.e., *the production of the non-traded good may either increase or decrease*); with regard to the argument in the text, note that η determines the magnitude of the *spending effect* (which causes the production of N to increase), while ξ_A determines the magnitude of the *resource movement effect* (which causes the production of N to decrease).

[8] It is as well to point out that we use "real wage" in the sense of wage expressed in terms of the product; the real wage expressed in terms of wage-earners' purchasing power will be examined later.

Since, from Eqs. (22.81), the real wage in sector B increases (remember that we have taken B as numéraire, so that p_L^* is expressed in terms of that commodity), employment, and therefore output, in this sector necessarily decrease (*de-industrialization*).

We come now to factor rewards. The real wage, measured in terms of workers' purchasing power, may vary in any direction according to the direction in which p_N varies (remember that p_A and p_B are assumed constant). If we indicate with α_N the share of wages used by workers to buy N, the variation in the real wage from the point of view of the workers will be

$$p_L^* - \alpha_N p_N^* = \frac{1}{H}\left\{\eta\theta_A\left(\xi_N - \alpha_N\right) + \xi_A\left[\Phi_N\left(1-\alpha_N\right) + \varepsilon_N\right]\right\}\pi, \tag{22.84}$$

which may also be negative if $p_N^* > 0$ and if α_N is sufficiently large.

In order to determine the variations in the rewards of the specific factors, all that is needed is to combine Eqs. (22.81) with Eqs. (22.76), by which we obtain

$$\begin{aligned} \theta_{KA}Hp_{KA}^* &= \left[-\eta\xi_N\theta_{LA}\theta_A + \Phi_N\left(1-\theta_{LA}\xi_A-\xi_N\right) + \varepsilon_N\left(1-\theta_{LA}\xi_A\right)\right]\pi, \\ \theta_{KB}Hp_{KB}^* &= -\theta_{LB}\left[\eta\xi_N\theta_A + \xi_A\left(\Phi_N+\varepsilon_N\right)\right]\pi < 0, \\ \theta_{KN}Hp_{KN}^* &= \left[\eta\left(1-\theta_{LN}\xi_N\right)\theta_A + \xi_A\left(\theta_{KN}\Phi_N - \theta_{LN}\varepsilon_N\right)\right]\pi. \end{aligned} \tag{22.85}$$

As can be seen, only the sign of p_{KB}^* is certain, that is, we are able to establish a priori that the reward for specific capital in sector B decreases, but we can say nothing a priori about the direction in which the reward for specific capital will vary in the other two sectors.

22.7 International Factor Mobility

The role of factor mobility in the Heckscher-Ohlin model was examined for the first time by Mundell (1957b: see Mundell, 1968, chap. 6), whose contribution has been set out in the text. Subsequently a line of research was developed (Jones, 1967; Kemp, 1969b, etc.) which dealt with the optimum tax to be imposed on movements of capital and the problem of what the tax should be if at the same time an optimum tariff is also levied on imports (see Sect. 11.1).

A third line of research (Bhagwati, 1973; Markusen & Melvin, 1979, etc.) looked into the effects on the welfare of the host country of a foreign capital inflow, followed by repatriation of profits. This literature aims to throw light on the age-old debate on the question of whether an inflow of capital is indeed a propitious event and thus to be encouraged, or whether it is damaging. It is necessary to note that in this type of analysis a continuous and potentially unlimited inflow is not considered (as in that case Mundell's results are valid), but a once-and-for-all inflow. The ownership of capital remains abroad and profits are repatriated.

The result of this analysis is that the capital inflow may in general have any effect on the welfare of the host country, as the welfare may either increase or decrease. It is fairly easy to demonstrate this result through our previous findings (see in particular Sect. 21.3) and the results of Sect. 27.2 below. In fact, a once-and-for-all capital inflow can be treated—leaving aside for the moment the question of repatriation of profits—as an exogenous increase in the existing stock of capital. The effects of this increase are well known (Rybczynski's theorem) and it is also known that under certain conditions there can be a decrease in welfare (the so-called immiserizing growth case: see Sect. 27.2). Furthermore, account must be taken of the decrease in welfare due to the fact that the profits accruing to foreign capital are to be deducted from national income, because they are repatriated. In other words, the final effect is given by the algebraic sum of two effects:

(a) The loss (or gain) that comes from the increase in capital stock;
(b) The loss that derives from the repatriation of profits on foreign capital.

Effect (a) is the one we shall meet in the analysis in Sect. 27.2, and it is clear that the addition of effect (b), which is certain to be negative, can cause the situation following the capital inflow to worsen in comparison to the initial one, not only when there is immiserizing growth (in which case effect (b) does no more than strengthen effect (a)), but also when (a) would in itself be positive.

By adopting the same criterion of comparison as in Sect. 27.2 (which allows us to avoid the problems inherent in social indifference curves) and taking up Eq. (27.25) below, we see that there will be an improvement or worsening according to whether

$$\frac{\partial I_A}{\partial \gamma} + \frac{\partial E_A/\partial \gamma}{1+\xi_1+\xi_2} \gtrless 0, \tag{22.86}$$

where, for brevity, we have omitted the subscript 1. By substituting the value of $\partial E_A/\partial \gamma$ from (27.19),[9] we have

$$\frac{\partial I_A}{\partial \gamma} + \left(\frac{\partial A^D}{\partial I_A}\frac{\partial I_A}{\partial \gamma} - \frac{\partial A}{\partial \gamma}\right)/1+\xi_1+\xi_2 \gtrless 0, \tag{22.87}$$

that is, by identifying factor γ with capital and rearranging the terms

$$\frac{\partial I_A}{\partial K}\left(1+\frac{\partial A^D/\partial I_A}{1+\xi_1+\xi_2}\right) - \frac{\partial A/\partial K}{1+\xi_1+\xi_2} \gtrless 0. \tag{22.88}$$

It is now necessary to calculate $\partial I_A/\partial K$, taking account of effect (b). We get

[9]Equation (27.19) has been used rather than (27.20), because, as will be seen, $\partial I_A/\partial \gamma = 0$ and thus the passage from the first to the second expression is not valid in this case.

$$\frac{\partial I_A}{\partial K} = \frac{\partial}{\partial K}(A + pB) - \left(g'_A k_A + p g'_B k_B\right), \tag{22.89}$$

where the second expression in parentheses in the right-hand side is the variation in income due to the repatriation of profits: k_A and k_B are the fractions of the capital inflow that are utilized in the two sectors and g'_A, pg'_B are the respective marginal productivities. As—see (19.17)—in equilibrium $g'_A = pg'_B$ and as $k_A + k_B = 1$ by definition, we have

$$\frac{\partial I_A}{\partial K} = \left(\frac{\partial A}{\partial K} + p\frac{\partial B}{\partial K}\right) - g'_A = 0, \tag{22.90}$$

since the expression in parentheses is equal to g'_A, on the basis of Eqs. (27.42) and (27.43). *The repatriation of profits thus entirely absorbs the increase in national income consisting of the additional output made possible by the capital inflow*. This is obvious if one thinks that the increase in output is given by the capital increase (inflow) times its marginal productivity (which in equilibrium is levelled in all sectors); by rewarding foreign capital on the basis of its marginal productivity the balance is zero.

Therefore Eq. (22.88), account being taken of (27.41), becomes

$$-\frac{\partial A/\partial K}{1 + \xi_1 + \xi_2} = \frac{g_A}{(\varrho_B - \varrho_A)(1 + \xi_1 + \xi_2)} \gtrless 0. \tag{22.91}$$

If we assume that A is the imported commodity, and bear in mind that $(1 + \xi_1 + \xi_2) < 0$ for stability, there will be an improvement or a worsening according to whether $\varrho_A \gtrless \varrho_B$ that is, according to whether the imported commodity is more or less capital-intensive than the exported one. This in turn is the same as saying that there will be an improvement or a worsening according to whether the terms of trade are better or worse: in fact, if we consider Eqs. (27.17) and (27.19) and bear in mind that $\partial I_A / \partial K = 0$, we have

$$\frac{\mathrm{d}p}{\mathrm{d}K} = -\frac{\partial A/\partial K}{E_{2B}(1 + \xi_1 + \xi_2)}, \tag{22.92}$$

which—as its denominator is negative—has a sign that coincides with that of (22.91).

This result must not be taken as to be in conflict with that in Sect. 27.2, where it will be demonstrated that the worsening in the terms of trade is only a necessary, not a sufficient, condition for immiserizing growth.

In fact, this result is true when only effect (a) is considered; by introducing effect (b) it can be seen that, *as national income has remained unvaried at the level prior to the foreign capital inflow*, the worsening in the terms of trade is evidently a necessary *and* sufficient condition to produce a worsening in the situation.

On international factor movements in general, see Various Authors (1983), Jones and Dei (1983) , Ruffin (1984), and Wong (1995).

22.7.1 The Theorems of International Trade Theory Under Factor Mobility

A fourth line of research (Ethier & Svensson, 1986; Wong, 1995, chap. 4) has examined the validity of the four core theorems of international trade theory (Heckscher-Ohlin, factor price equalization, Rybczynski, Stolper and Samuelson) in the presence of factor mobility. The general result (Ethier & Svensson, 1986) is that appropriate versions of these theorems still hold provided that the number of commodities and mobile factors is at least as large as the total number of factors. This shows that the theorems are sensitive to the total number of markets (and not to the number of commodities) relative to the number of factors.

We shall illustrate this result by a two-country, two-commodity, three-factor (one which is mobile) model due to Wong (1995, chap. 4, sect. 4.1), on which the following treatment is based.

The basic assumption is that, in addition to capital (K) and labour (L), there is a third primary factor, land (D). Capital is the internationally mobile factor, while labour and land are immobile. The production function in sector $i = A, B$ is

$$Q_i = F_i(K_i, L_i, D_i), \tag{22.93}$$

with the usual properties (first-degree homogeneity, etc.).

The representative firm's optimization problem is to choose the inputs (and hence the output) so as to maximize profit for any given set of prices, namely

$$\max_{K_i, L_i, D_i} \{p_i F_i(K_i, L_i, D_i) - p_K K_i - p_L L_i - p_D D_i\}. \tag{22.94}$$

This maximization can also be carried out in two stages: in the first, the firm maximizes the objective function with respect to K_i taking L_i, D_i as given; in the second stage the result of the first stage is plugged in the objective function, which is maximized with respect to L_i, D_i. Thus we have

$$\max_{L_i, D_i} \left\{ \max_{K_i} [p_i F_i(K_i, L_i, D_i) - p_K K_i] - p_L L_i - p_D D_i \right\}. \tag{22.95}$$

Let us now define for each sector the function

$$H_i(L_i, D_i, r_i) \equiv \max_{K_i} [F_i(K_i, L_i, D_i) - r_i K_i], \tag{22.96}$$

where $r_i \equiv p_K/p_i$ is the real rental rate in terms of commodity i. The solution to this maximization problem is given by

$$\frac{\partial F_i(K_i, L_i, D_i)}{\partial K_i} - r_i = 0. \tag{22.97}$$

Since the conditions of the implicit function theorem are satisfied (we have $\partial^2 F_i/\partial K_i^2 \neq 0$, in particular $\partial^2 F_i/\partial K_i^2 < 0$ by the assumption of decreasing marginal productivity), Eq. (22.97) can be solved for the optimal value of K_i in terms of the parameters, say

$$K_i = G_i(L_i, D_i, r_i).$$

The function G_i is a continuously differentiable function of its arguments by the implicit function theorem.

Since the production function $F_i(K_i, L_i, D_i)$ is first-degree homogeneous, the function $G_i(L_i, D_i, r_i)$ is homogeneous of the first degree with respect to L_i, D_i when given r_i. It follows that the function $H_i(L_i, D_i, r_i)$ is also homogeneous of the first degree with respect to L_i, D_i when given r_i. Besides, the envelope theorem (see, for example, Mas-Colell, Whinston, & Green, 1995, pp. 964–966) shows that the partial derivatives of $H_i(L_i, D_i, r_i)$ with respect to L_i, D_i are equal to the corresponding derivatives of $F_i(K_i, L_i, D_i)$, namely

$$\frac{\partial H_i}{\partial L_i} = \frac{\partial F_i}{\partial L_i}, \quad \frac{\partial H_i}{\partial D_i} = \frac{\partial F_i}{\partial D_i}. \tag{22.98}$$

Finally, the (strict) concavity of F_i implies that H_i is (strictly) concave with respect to L_i, D_i when given r_i. In fact, consider the Hessian matrix of F_i

$$M_{F_i} = \begin{bmatrix} \dfrac{\partial^2 F_i}{\partial K_i^2} & \dfrac{\partial^2 F_i}{\partial K_i \partial L_i} & \dfrac{\partial^2 F_i}{\partial K_i \partial D_i} \\ \dfrac{\partial^2 F_i}{\partial L_i \partial K_i} & \dfrac{\partial^2 F_i}{\partial L_i^2} & \dfrac{\partial^2 F_i}{\partial L_i \partial D_i} \\ \dfrac{\partial^2 F_i}{\partial D_i \partial K_i} & \dfrac{\partial^2 F_i}{\partial D_i \partial L_i} & \dfrac{\partial^2 F_i}{\partial D_i^2} \end{bmatrix},$$

which is negative definite when F_i is (strictly) concave. The Hessian matrix of H_i is

$$M_{H_i} = \begin{bmatrix} \dfrac{\partial^2 H_i}{\partial L_i^2} & \dfrac{\partial^2 H_i}{\partial L_i \partial D_i} \\ \dfrac{\partial^2 H_i}{\partial D_i \partial L_i} & \dfrac{\partial^2 H_i}{\partial D_i^2} \end{bmatrix}.$$

From (22.98) it follows that, in the neighbourhood of the optimum point, M_{H_i} coincides with the south-east leading principal submatrix of M_{F_i} (the matrix obtained by deleting the first row and column of M_{F_i}). Hence if M_{F_i} satisfies the conditions for negative definiteness (the principal minors alternate in sign, beginning with minus), M_{H_i} satisfies them as well.

From all this it follows that $H_i(L_i, D_i, r_i)$ behaves like a production function in the two factors L_i, D_i.

Consider now the firm's optimization problem, that—by Eqs. (22.95) and (22.96)—can be written as

$$\max_{L_i, D_i} \{p_i H_i(L_i, D_i, r_i) - p_L L_i - p_D D_i\}. \tag{22.99}$$

As shown above, we can take $H_i(L_i, D_i, r_i)$ as a production function, so that we can use (22.99) to define a framework similar to the standard two-factor, two-sector framework. This stratagem greatly simplifies the analysis.

22.7.1.1 The Heckscher-Ohlin Theorem

We make the usual assumptions (identical technologies, homothetic preferences etcetera: see Chap. 4). The functions $H_i(L_i, D_i, r_i)$ are internationally equal since the only possible element of difference (r_i) is equalized by the international mobility of goods and capital. Thus we can concentrate on labour and land and their (relative) abundance. In exactly the same manner as in Sect. 4.2, we can show that, at the same commodity-price ratio, a country abundant in a factor has a production bias in favour of the commodity which uses that factor more intensively, and hence that it will export that commodity given the internationally identical and homothetic structure of demand.

22.7.1.2 Factor Price Equalization

Rental rates are equalized by free trade and free capital mobility. Then we can use the traditional arguments (Sect. 4.3) on the functions $H_i(L_i, D_i, r_i)$ to show that with internationally identical commodity prices and rental rates, the prices of labour and land are also equalized.

22.7.1.3 The Rybczynski Theorem

Consider a closed economy, and suppose that p_K, p_A, p_B are given, hence r_i is also given. We know that, given r_i, the functions $H_i(L_i, D_i, r_i)$ behave like ordinary production functions in the arguments (factors) L_i, D_i. Without loss of generality we can assume that commodity A is labour intensive, with a higher labour/land ratio than B. Then we can apply the traditional arguments (see Sect. 5.4) to show that the increase in the quantity of a factor (say, labour) causes an increase in the output of the commodity intensive in that factor (A) and a decrease in the output of the other commodity, at unchanged commodity and factor prices (i.e., given also p_L, p_D).

This proves the Rybczynski theorem.

22.7.1.4 The Stolper-Samuelson Theorem

Let $a_{ij}, i = K, L, D; j = A, B$ the (optimal) input coefficients, namely the amount of factor i required to produce one unit of commodity j when costs are minimized. Then we have

$$\begin{aligned} p_A &= a_{KA} p_K + a_{LA} p_L + a_{DA} p_D, \\ p_B &= a_{KB} p_K + a_{LB} p_L + a_{DB} p_D. \end{aligned} \tag{22.100}$$

If we differentiate both sides (keeping p_K constant) and consider proportional changes (denoted by an asterisk) we obtain, by the same procedure followed in Sect. 19.5,

$$\begin{aligned} p_A^* &= \theta_{LA} p_L^* + \theta_{DA} p_D^* + \theta_{KA} a_{KA}^* + \theta_{LA} a_{LA}^* + \theta_{DA} a_{DA}^*, \\ p_B^* &= \theta_{LB} p_L^* + \theta_{DB} p_D^* + \theta_{KB} a_{KB}^* + \theta_{LB} a_{LB}^* + \theta_{DB} a_{DB}^*, \end{aligned} \tag{22.101}$$

where θ_{ij} is the share of factor i in sector j ($\theta_{LA} = a_{LA} p_L / p_A$, etc.). Cost minimization (see Sect. 19.5) implies

$$\theta_{KA} a_{KA}^* + \theta_{LA} a_{LA}^* + \theta_{DA} a_{DA}^* = 0,$$

$$\theta_{KB} a_{KB}^* + \theta_{LB} a_{LB}^* + \theta_{DB} a_{DB}^* = 0,$$

hence Eqs. (22.101) reduce to

$$\begin{aligned} p_A^* &= \theta_{LA} p_L^* + \theta_{DA} p_D^*, \\ p_B^* &= \theta_{LB} p_L^* + \theta_{DB} p_D^*. \end{aligned} \tag{22.102}$$

These equations can be solved for p_L^*, p_D^* in terms of p_A^*, p_B^*, thus obtaining

$$\begin{aligned} p_L^* &= \frac{\theta_{DB} p_A^* - \theta_{DA} p_B^*}{\theta_{LA}\theta_{DB} - \theta_{LB}\theta_{DA}}, \\ p_D^* &= \frac{\theta_{LA} p_B^* - \theta_{LB} p_A^*}{\theta_{LA}\theta_{DB} - \theta_{LB}\theta_{DA}}. \end{aligned} \tag{22.103}$$

Let us assume, for example, that commodity B is labour intensive. Given the definition of the θ's, this implies $\theta_{LB}/\theta_{DB} > \theta_{LA}/\theta_{DA}$ and hence that the denominator of the fractions in Eqs. (22.103) is negative. Without loss of generality we can assume that commodity A is the numéraire, $p_A = 1$, hence $p_A^* = 0$. Thus a positive (negative) value of p_B^* means an increase (decrease) in the relative price p_B/p_A, and a positive (negative) value of p_L^* means an increase in the real reward of labour.

If we then let $p_B^* > 0$, we see from Eqs. (22.103) that $p_L^* > 0$, $p_D^* < 0$. Thus an increase of the relative price of a commodity causes an increase in the real reward of the factor intensively used in the production of this commodity. This proves the Stolper-Samuelson theorem.

Let us now again differentiate Eqs. (22.100), this time keeping commodity prices constant but letting p_K vary. We obtain, using the cost minimization conditions,

$$\begin{aligned} 0 &= \theta_{KA} p_K^* + \theta_{LA} p_L^* + \theta_{DA} p_D^*, \\ 0 &= \theta_{KB} p_K^* + \theta_{LB} p_L^* + \theta_{DB} p_D^*. \end{aligned} \tag{22.104}$$

These equations show that, if the reward to capital increases, the price of at least one immobile factor must decrease. To obtain more definite results we can solve Eqs. (22.104) for p_L^*, p_D^* in terms of p_K^*, whence

$$\begin{aligned} p_L^* &= \frac{\theta_{KB}\theta_{DA} - \theta_{KA}\theta_{DB}}{\theta_{LA}\theta_{DB} - \theta_{LB}\theta_{DA}} p_K^*, \\ p_D^* &= \frac{\theta_{KA}\theta_{LB} - \theta_{KB}\theta_{LA}}{\theta_{LA}\theta_{DB} - \theta_{LB}\theta_{DA}} p_K^*. \end{aligned} \tag{22.105}$$

Let us keep for the moment to the assumption that commodity B is labour intensive (hence A is land intensive), which means that $D_A/L_A > D_B/L_B$, or $D_A/D_B > L_A/L_B$. The denominator of the fractions in Eqs. (22.105) is negative. Then as a result of an increase in p_K the price of labour increases when $\theta_{KB}\theta_{DA} - \theta_{KA}\theta_{DB} < 0$, or, using the definitions of the θ's and a's, when $K_A/K_B > D_A/D_B$. Since we have assumed $D_A/D_B > L_A/L_B$, the condition for p_L^* to be positive when p_K^* is positive becomes

$$K_A/K_B > D_A/D_B > L_A/L_B. \tag{22.106}$$

When commodity B is land intensive, the denominator of the fractions in Eqs. (22.105) is positive, and the condition becomes

$$L_A/L_B > D_A/D_B > K_A/K_B. \tag{22.107}$$

In both cases D_A/D_B is included between L_A/L_B and K_A/K_B, and land is called a middle factor by Wong (1995, p. 143). Obviously, when $p_L^* > 0$, then $p_D^* < 0$.

Similarly it can be shown than, when labour is the middle factor, then $p_K^* > 0$ gives rise to $p_D^* > 0$, $p_L^* < 0$.

Finally, if capital is the middle factor, then Eqs. (22.105) imply that an increase in p_K causes a decrease in both p_L and p_D.

All these results can be summarized by saying that, when capital is not the middle factor, and increase in its reward causes a decrease in the reward of the middle factor and an increase in the reward of the other immobile factor. When capital is the middle factor, an increase in its reward causes a decrease in the rewards of both immobile factors.

22.7.2 Factor Mobility in the Specific Factors Model

The effects on factor rewards and outputs of an inflow of labour or of a specific capital have already been determined in the treatment of this model in Sect. 22.1, so that we only reproduce them here:

(I) Effects on factor rewards:

$$\begin{aligned} K^{*B} &= L^* = 0 \text{ and } \ K_A^* > 0 \Longrightarrow p_{K^A}^* < 0, p_{K^B}^* < 0, p_L^* > 0, \\ K^{*A} &= L^* = 0 \text{ and } \ K_B^* > 0 \Longrightarrow p_{K^A}^* < 0, p_{K^B}^* < 0, p_L^* > 0, \\ K^{*A} &= K_B^* = 0 \text{ and } \ L^* > 0 \Longrightarrow p_{K^A}^* > 0, p_{K^B}^* > 0, p_L^* < 0. \end{aligned}$$

(II) Effects on outputs:

$$\begin{aligned} K^{*B} &= L^* = 0 \text{ and } \ K_A^* > 0 \Longrightarrow A^* > 0, B^* < 0, \\ K^{*A} &= L^* = 0 \text{ and } \ K_B^* > 0 \Longrightarrow A^* < 0, B^* > 0, \\ K^{*A} &= K_B^* = 0 \text{ and } \ L^* > 0 \Longrightarrow A^* > 0, B^* > 0. \end{aligned}$$

22.8 Uncertainty and International Trade

Here, following Dumas (1980), we shall examine the case of generalized uncertainty, in which the production function of a generic good Y takes the form

$$Y_s = F_s(K, L), \tag{22.108}$$

where the subscript s indicates the states of nature. Thanks to first-degree homogeneity, we can write

$$y_s = g_s(\varrho), \quad y_s \equiv Y_s/L, \quad \varrho \equiv K/L. \tag{22.109}$$

Let us assume that the factors are rewarded at the beginning of the period and let us introduce Arrow-Debreu uncertainty, where we indicate with Φ_s the price of elementary or pure securities. As it would not be possible to show here the basis of these theories of uncertainty, the reader is referred to Arrow (1964), Debreu (1959), and Hirshleifer (1970). We only recall that an "elementary security" of index s is a security with a price equal to one, if the state of nature s occurs, equal to zero otherwise. As $p_s g_s'(\varrho)$ is the value of the marginal product of capital if the state of nature s occurs, and as only one of these states of nature will occur, then $\sum_s \Phi_s p_s g_s'(\varrho)$, given the definition of Φ_s, is the value of the marginal product of capital which is actually found.

Competitive equilibrium implies

$$\begin{aligned} p_K &= \sum_s \Phi_s p_s g_s'(\varrho), \\ p_L &= \sum_s \Phi_s p_s g_s(\varrho) - \varrho \sum_s \Phi_s p_s g_s'(\varrho), \end{aligned} \tag{22.110}$$

where p_s is the price of the commodity in each state of nature.

Let us assume that two commodities, A and B, are produced and let us consider the *present market value* of each product in each sector

$$\begin{aligned} V_A &= \sum_s \Phi_s p_{sA} L_A g_{sA}(\varrho_A) = L \sum_s \Phi_s p_{sA} l_A g_{sA}(\varrho_A), \\ V_B &= \sum_s \Phi_s p_{sB} L_B g_{sB}(\varrho_B) = L \sum_s \Phi_s p_{sB} l_B g_{sB}(\varrho_B), \end{aligned} \tag{22.111}$$

where $l_A = L_A/L$, $l_B = L_B/L$ are the fractions of the total labour force employed in the two sectors. As there is full employment of labour, $l_A + l_B = 1$: then, considering the condition of full employment of the capital stock and denoting the given total capital/labour ratio with $\bar{\varrho}$, we get

$$l_A \varrho_A + l_B \varrho_B = \bar{\varrho}, \tag{22.112}$$

which, together with the condition of full employment of labour, makes it possible to obtain

$$l_A = \frac{\bar{\varrho} - \varrho_B}{\varrho_A - \varrho_B}, l_B = \frac{\varrho_A - \bar{\varrho}}{\varrho_A - \varrho_B}. \tag{22.113}$$

If we consider the ratio between the present market values of the future outputs, $v = V_A/V_B$, given Eqs. (22.111) and (22.113), we have

$$v = \frac{\sum_s \Phi_s p_{sA} g_{sA}(\varrho_A)}{\sum_s \Phi_s p_{sB} g_{sB}(\varrho_B)} \frac{\bar{\varrho} - \varrho_B}{\varrho_A - \bar{\varrho}}. \tag{22.114}$$

Let us now assume that there are two countries and that commodities are freely traded in all states of nature in both countries, so that p_{sA} and p_{sB} are the same everywhere. Let us also assume that the pure security markets are unified at world level, so that the Φ_s are equal in the two countries. The production functions are internationally identical and there is no factor-intensity reversal.

Without any loss of generality we can assume that A is the capital intensive commodity, so that $\varrho_A > \bar{\varrho} > \varrho_B$. It can then be seen at once from (22.114) that the country in which $\bar{\varrho}$ is higher will have a higher v, that is, a relatively greater V_A. This shows that *the capital-abundant country produces a relatively greater present market value of the capital-intensive commodity*, and vice versa for the labour-abundant country. Obviously, this proposition is the extension to the case of uncertainty (with present market value in the place of certain quantity) of the proposition at the basis of the Heckscher-Ohlin theorem (see Sect. 4.2).

If we now assume, as in the Heckscher-Ohlin theory, identical demand structures in the two countries (no element of uncertainty being introduced on the demand side), it immediately follows that *each country has a positive present value of exports of the commodity which makes relatively intensive use of the relatively plentiful factor.* This extends the Heckscher-Ohlin theorem to the case of uncertainty.[10]

Assuming absence of complete specialization, it is possible to demonstrate the factor-price equalization: given Eqs. (22.110), inside each country we shall have

$$\begin{aligned} p_K &= \sum_s \Phi_s p_{sA} g'_{sA}(\varrho_A) = \sum_s \Phi_s p_{sB} g'_{sB}(\varrho_B), \\ p_L &= \sum_s \Phi_s p_{sA} g_{sA}(\varrho_A) - \varrho_A \sum_s \Phi_s p_{sA} g'_{sA}(\varrho_A) \\ &= \sum_s \Phi_s p_{sB} g_{sB}(\varrho_B) - \varrho_B \sum_s \Phi_s p_{sB} g'_{sB}(\varrho_B), \end{aligned} \tag{22.115}$$

from which

$$\begin{aligned} &\sum_s \Phi_s p_{sA} g'_{sA}(\varrho_A) - \sum_s \Phi_s p_{sB} g'_{sB}(\varrho_B) = 0, \\ &\left[\sum_s \Phi_s p_{sA} g_{sA}(\varrho_A) - \varrho_A \sum_s \Phi_s p_{sA} g'_{sA}(\varrho_A)\right] \\ &-\left[\sum_s \Phi_s p_{sB} g_{sB}(\varrho_B) - \varrho_B \sum_s \Phi_s p_{sB} g'_{sB}(\varrho_B)\right] = 0, \end{aligned} \tag{22.116}$$

which is a system of two implicit functions. On the basis of the implicit-function theorem, if the Jacobian with respect to ϱ_A, ϱ_B is different from zero at the equilibrium point, it is possible to express ϱ_A and ϱ_B as single-valued differentiable functions of the other $3s$ variables (Φ_s, p_{sA}, p_{sB}).

The Jacobian is

$$\begin{vmatrix} \sum_s \Phi_s p_{sA} g''_{sA}(\varrho_A) & -\sum_s \Phi_s p_{sB} g''_{sB}(\varrho_B) \\ -\varrho_A \sum_s \Phi_s p_{sA} g''_{sA}(\varrho_A) & -\varrho_B \sum_s \Phi_s p_{sB} g''_{sB}(\varrho_B) \end{vmatrix}$$

$$= \left[\sum_s \Phi_s p_{sA} g''_{sA}(\varrho_A)\right]\left[\sum_s \Phi_s p_{sB} g''_{sB}(\varrho_B)\right](\varrho_B - \varrho_A), \tag{22.117}$$

which is different from zero because, given the assumption of absence of factor-intensity reversals, there will always be $\varrho_A > \varrho_B$ or $\varrho_B > \varrho_A$,

[10] It is as well to observe that the extension of this theorem from the deterministic case to one with uncertainty is valid only if the *physical* definition of relative abundance is used, whereas if the definition in terms of relative factor prices is used, then such an extension is no longer valid.

As we have assumed that the production functions are internationally identical and the variables Φ_s, p_{sA}, p_{sB} likewise, the values of ϱ_A and ϱ_B derived from Eqs. (22.116) will be identical in both countries so that, by substituting in Eqs. (22.115), we get the same factor prices in both countries.

For a demonstration of the validity of the other traditional theorems (Stolper-Samuelson, Rybczynski) we refer the reader to Dumas (1980). See also Helpman and Razin (1978) , Eaton (1979), Pomery (1979, 1984), Anderson (1981), Grossman and Razin (1985), and Grinols (1985).

22.9 Smuggling

Let us take as example the case in which the real costs of smuggling are made up exclusively of a loss of part of the commodity smuggled. We start from the following model (Bhagwati & Srinivasan, 1974)

$$\begin{aligned} C_A &= A + m_{Ag} + m_{As}, \\ C_B &= f(A) - pm_{Ag} - p_s m_{As}, \\ U_A &= p_h U_B, \\ -f'(A) &= p_h. \end{aligned} \tag{22.118}$$

The first equation defines the domestic consumption of the imported commodity (we assume that it is A), given by domestic output plus imports, distinguished in legal imports m_{Ag} and illegal ones m_{As}. The second equation defines the domestic consumption of commodity B, equal to domestic production less exports. Domestic production of B is connected to that of A by way of the transformation curve $B = f(A)$. Exports of commodity B are equal, in equilibrium, to the values of the corresponding imports of A in the two branches of trade (legal trade and smuggling), where p and p_s are the international relative price of A[11] for legal trade and the relative price of the same commodity illegally traded ($p_s > p$).

Given a social welfare function $U = U(A, B)$, with positive partial derivatives U_A, U_B, the optimum condition is given by the equality between the marginal rate of substitution (U_A/U_B) and the domestic relative price p_h, hence the third equation. The fourth and last expresses the fact that, on the basis of the efficiency conditions (see Sect. 19.1), the marginal rate of transformation is equal to the domestic relative price p_h.

Given that the domestic (relative) price charged by the smugglers (henceforth "domestic illegal price" for brevity) is less than the legal (relative) domestic price (which is equal to the international price plus tariff), legal trade will disappear, so that $p_h = p_s, m_{Ag} = 0$. We propose to calculate the direction in which social

[11]To symplify analysis we use the relative price of commodity A instead of that of B as we did in Sect. 6.10.

welfare moves with the variation in the price of the domestic illegal price $p_s = p_h$, in the interval $p \leq p_s \leq p(1+d)$, where d is the tariff rate, assuming that p is constant. From the social welfare function, we get

$$\begin{aligned} \frac{\mathrm{d}U}{\mathrm{dp}_h} &= U_A \frac{\mathrm{d}C_A}{\mathrm{d}p_h} + U_B \frac{\mathrm{d}C_B}{\mathrm{d}p_h} = U_B \left(\frac{U_A}{U_B} \frac{\mathrm{d}C_A}{\mathrm{d}p_h} + \frac{\mathrm{d}C_B}{\mathrm{d}p_h} \right) \\ &= U_B \left(p_h \frac{\mathrm{d}C_A}{\mathrm{d}p_h} + \frac{\mathrm{d}C_B}{\mathrm{d}p_h} \right), \end{aligned} \tag{22.119}$$

given the third equation of (22.118). The last expression in parentheses is formally identical to the following

$$\frac{\mathrm{d}}{\mathrm{d}p_h} \left(p_h C_A + C_B \right) - C_A.$$

Remembering that $m_{Ag} = 0$, $p_h = p_s$, it follows from the first two equations in (22.118) that

$$p_h C_A + C_B = p_h A + f\left(A\right),$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}p_h} \left(p_h C_A + C_B \right) = A + p_h \frac{\mathrm{d}A}{\mathrm{d}p_h} + f'\left(A\right) \frac{\mathrm{d}A}{\mathrm{d}p_h} = A,$$

given the fourth of Eqs. (22.118). So, by substituting in (22.119), we have

$$\frac{\mathrm{d}U}{\mathrm{d}p_h} = U_B \left(A - C_A \right) = -U_B m_{As} < 0, \tag{22.120}$$

given the first of (22.118) and the fact that $m_{Ag} = 0$. It follows from (22.120) that social welfare is a monotonically decreasing function of the domestic illegal price. There will obviously be maximum welfare at the lower bound of the interval, that is, when $p_s = p$ (the free trade price), while there will be minimum welfare at the upper bound of the interval, that is when $p_s = p(1+d)$. Now, as we have seen in the text, this minimum is inferior to that which the society would have if there were no smuggling and the legal domestic price were equal to $p(1+d)$. We can therefore establish that

$$U^f > U^d > U^s_{\min},$$

where U^f = welfare in the case of free trade, U^d = welfare in the case of a tariff and legal trade, U^s_{min} = welfare in the case of smuggling with a relative price equal to that of legal trade with tariff. The U^s welfare that the society enjoys in the case of smuggling will therefore be included between U^f and U^s_{min} and, given the monotonic relationship between welfare and the domestic illegal price, it is demonstrated that U^s can be less or greater than U^d, according to the value assumed by $p_s = p_h$.

The economic theory of smuggling can be put in the general framework of the theory of DUP (Directly UnProductive) activities, for which see Bhagwati and Srinivasan (1983, chap. 30, and references therein). For a crime-theoretical approach see Martin and Panagariya (1984).

References

Anderson, J. E. (1981). The Heckscher-Ohlin and Travis-Vanek theorems under uncertainty.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk bearing.

Batra, R. N., & Casas, F. R. (1973). Intermediate products and the pure theory of international trade: A Neo-Heckscher-Ohlin framework.

Bhagwati, J. N., & Srinivasan, T. N. (1974). An alternative proof of the Bhagwati-Hansen results on smuggling and welfare.

Bhagwati, J. N., & Srinivasan, T. N. (1983). *Lectures on international trade*.

Casas, F. R. (1983). International trade with produced transport services.

Casas, F. R., & Choi, E. K. (1985b). Some paradoxes of transport costs in international trade.

Chacholiades, M. (1978). *International trade theory and policy* (chap. 8).

Corden, W. M. (1984a). Booming sector and Dutch disease economic: Survey and consolidation.

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialization in a small open economy.

Debreu, G. (1959). *Theory of value* (chap. 7).

Dumas, B. (1980). The theorems of international trade under generalized uncertainty.

Eaton, J. (1979). The allocation of resources in an open economy with uncertain terms of trade.

Ethier, W. J., & Svensson, L. E. O. (1986). The theorems of international trade with factor mobility.

Grinols, E. L. (1985). International trade with incomplete markets.

Grossman, G. M., & Razin, A. (1985). The pattern of trade in a Ricardian model with country-specific uncertainty.

Hadley, G., & Kemp, M. C. (1966). Equilibrium and efficiency in international trade.

Helpman, E., & Razin, A. (1978). *A theory of international trade under uncertainty*.

Hirshleifer, J. (1970). *Investment, interest, and capital*.

Jones, R. W. (1967). International capital movements and the theory of tariffs and trade.

Jones, R. W. (1971). A three-factor model in theory, trade, and history.

Jones, R. W., & Dei, F. (1983). International trade and foreign investment: A simple model.

Kemp, M. C. (1969b). *The pure theory of international trade and investment* (chaps. 5–7, 9).

Komiya, R. (1967). Non-traded goods and the pure theory of international trade.

Laffer, A. B., & Miles, M. A. (1982). *International economics in an integrated world* (chap. 8).

Markusen, J. R., & Melvin, J. R. (1979). Tariffs, capital mobility, and foreign ownership.

Martin, J. P., & Neary, J. P. (1980). Variable labour supply and the pure theory of international trade: An empirical note.

Martin, L., & Panagariya, A. (1984). Smuggling, trade and price disparity: A crime-theoretic approach.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.*

Mundell, R. A. (1968). *International economics* (chaps. 5 and 6).

Pomery, J. (1979). Uncertainty and international trade.

Pomery, J. (1984). Uncertainty in trade models.

Ruffin, R. J. (1984). International factor movements.

Sanyal, K. K., & Jones, R. W. (1982). The theory of trade in middle products.
Samuelson, P. A. (1971). Ohlin was right.
Sarkar, A. (1985). A model of trade in intermediate goods.
Vanek, J. (1963). Variable factor proportions and interindustry flows in the theory of international trade.
Various Authors, (1983). Paradoxes in capital theory: A symposium.
Wong, K.-Y. (1995). *International trade in goods and factor mobility.*

Chapter 23
Appendix to Chapter 9

23.1 A Neo-Heckscher-Ohlin Model

We examine Falvey's model (Falvey, 1981). For the reader's convenience we report here Eqs. (9.1) from the text:

$$\begin{aligned} p_1(\alpha) &= W_1 + \alpha R_1, \\ p_2(\alpha) &= W_2 + \alpha R_2, \end{aligned}$$

where α is a continuous index over the interval $\underline{\alpha}$, $\overline{\alpha}$; the units are chosen such that the production of one unit of α requires the input of α units of capital and one unit of labour.

The solution for α_0, the marginal quality such that $p_1(\alpha) = p_2(\alpha)$, is

$$\alpha_0 = \frac{W_1 - W_2}{R_2 - R_1}, \tag{23.1}$$

which is clearly positive, since we have assumed that $W_1 > W_2$ and $R_1 < R_2$. For any other quality we have $p_1 \neq p_2$, and precisely

$$p_1(\alpha) - p_2(\alpha) = (W_1 - W_2) + \alpha(R_1 - R_2),$$

from which, using the fact that Eq. (23.1) yields $R_2 - R_1 = (W_1 - W_2)/\alpha_0$,

$$p_1(\alpha) - p_2(\alpha) = (W_1 - W_2)(\alpha_0 - \alpha)/\alpha_0. \tag{23.2}$$

It can readily be seen from (23.2) that $p_1(\alpha) \lesseqgtr p_2(\alpha)$ according as $\alpha \gtreqless \alpha_0$; this means that the home country produces the qualities higher than the marginal quality α_0 at lower unit costs than the rest of the world and vice versa. From this result, one can anticipate that under free trade and with no transport costs the home country will

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_23,

export the qualities higher than α_0 and import the qualities lower than α_0. This intra-industry trade will follow the lines of the Heckscher-Ohlin proposition, as shown in the text, Sect. 9.1.

Let us now explicitly consider the demand side. The demand for each quality is assumed to depend only on the relative prices of qualities; since we are in a partial equilibrium context, consumers' income as well as the prices of the products of other industries can be taken as given and hence can be ignored. Since perfect competition obtains in the industry, prices will equal unit production costs and so, as the wage rate is given, will depend only on the rate of profit. Thus we can write the demands for quality α as

$$\begin{aligned} D_1 &= D_1(R_1, R_2; \alpha), \\ D_2 &= D_2(R_1, R_2; \alpha). \end{aligned}$$

We must now determine the equilibrium rates of return on capital, R_{1E} and R_{2E}, which are the rates that bring into equality the demand for capital and the (given) supply of it. Let $\underline{\alpha}$ and $\overline{\alpha}$ be the indices of the lowest and highest quality respectively, and K_1, K_2 the industry's stock of capital in the two countries. Since α also measures the capital input, and given the results on the pattern of trade, we have

$$D_{1K}(R_{1E}, R_{2E}) \equiv \int_{\alpha_0}^{\overline{\alpha}} \alpha[D_1(R_{1E}, R_{2E}; \alpha) + D_2(R_{1E}, R_{2E}; \alpha)]\mathrm{d}\alpha = K_1, \tag{23.3}$$

because all of the world demand (domestic plus foreign) for the qualities higher than α_0 will be met by the home country's output. Similarly,

$$D_{2K}(R_{1E}, R_{2E}) \equiv \int_{\underline{\alpha}}^{\alpha_0} \alpha[D_1(R_{1E}, R_{2E}; \alpha) + D_2(R_{1E}, R_{2E}; \alpha)]\mathrm{d}\alpha = K_2, \tag{23.4}$$

since all of the world demand for the qualities below α_0 will be met by the rest-of-the-world's output. Note that in (23.3) and (23.4), α_0 is a function of $(R_{1E} - R_{2E})$ through (23.1).

We observe that in (23.3) an increase in R_1 reduces the home country's excess demand for capital for two reasons. First, this increase raises the prices of domestically produced qualities relative to foreign-produced ones and so—assuming that demand functions are normal—induces a substitution of the latter for the former. Second, the increase reduces the range of qualities where the home country has a cost advantage over the rest of the world.

Conversely, an increase in R_2 causes the excess demand for capital in the home country to increase. Therefore, if we denote this excess demand by $E^1(R_1, R_2) = D_{1K} - K_1$, the partial derivatives will be $E^1_{R_1} < 0$, $E^1_{R_2} > 0$.
Similar considerations applied to $E^2(R_1, R_2)$ give $E^2_{R_1} > 0$, $E^2_{R_2} < 0$.

Let us now consider the stability of equilibrium. This requires that any change which raises (reduces) the price of the qualities produced in a country, with other prices constant, brings about a decrease (increase) in the overall demand for capital. This implies that $E^1_{R_1} + E^2_{R_1} < 0$, $E^1_{R_2} + E^2_{R_2} < 0$; these inequalities will be used in the following comparative statics analysis (this use is an application of Samuelson's correspondence principle).

We now examine the effects of an increase in the home country's wage rate on the free trade equilibrium values of R_1 and R_2. Since the wage rate is given, it can be introduced as a shift parameter in the excess demand functions for capital defined above. We can then calculate the total differentials of these excess demand functions and obtain the system

$$E^1_{R_1}\mathrm{d}R_1 + E^1_{R_2}\mathrm{d}R_2 + E^1_{W_1}\mathrm{d}W_1 = 0, \tag{23.5}$$

$$E^2_{R_1}\mathrm{d}R_1 + E^2_{R_2}\mathrm{d}R_2 + E^2_{W_1}\mathrm{d}W_1 = 0, \tag{23.6}$$

which has the solution

$$\mathrm{d}R_1 = -\frac{E^1_{W_1}E^2_{R_2} - E^2_{W_1}E^1_{R_2}}{\Delta}\mathrm{d}W_1, \tag{23.7}$$

$$\mathrm{d}R_2 = -\frac{E^2_{W_1}E^1_{R_1} - E^1_{W_1}E^2_{R_1}}{\Delta}\mathrm{d}W_1, \tag{23.8}$$

where $\Delta \equiv E^1_{R_1}E^2_{R_2} - E^2_{R_1}E^1_{R_2}$ is positive given the stability condition discussed above. If we extend stability considerations to the effects of a change in wages, we can assume that $E^1_{W_1} + E^2_{W_1} < 0$, with $E^1_{W_1} < 0$ and $E^2_{W_1} > 0$.

From all these stability considerations it follows that $\left|E^2_{R_2}\right| > \left|E^1_{R_2}\right|$ and $\left|E^1_{W_1}\right| > \left|E^2_{W_1}\right|$, so that from (23.7) we have $\mathrm{d}R_1/\mathrm{d}W_1 < 0$; but the sign of $\mathrm{d}R_2/\mathrm{d}W_1$ remains ambiguous since $\left|E^1_{R_1}\right| > \left|E^2_{R_1}\right|$.

The economic interpretation of these results is the following. At the initial rates of return to capital, the increase in W_1 causes an increase in the domestically produced qualities and so a decrease in the range of qualities in which the home country has a cost advantage (as can be seen from (23.1), an increase in W_1 raises α_0 at unchanged R_1, R_2, W_2). Since foreign prices are unchanged, in world demand there will be a substitution in favour of foreign-produced qualities, and so an excess supply of capital in the home country industry. This excess reduces the rate of return to the domestic industry's capital, which tends to offset the initial effect of the higher wage on costs. In the new equilibrium, R_1 will therefore be lower, while the final position of R_2 is ambiguous (since it increases initially, because of the excess demand for it due to the excess demand for foreign-produced qualities, and then decreases).

23.2 A Model of Monopolistic Competition

23.2.1 Love for Variety and Demand

This model is based on the contributions by Krugman and Helpman (Helpman, 1990; Helpman & Krugman, 1985, 1989; Krugman, 1979, 1980, 1990). It starts from the S-D-S (Spence-Dixit-Stiglitz) approach to consumer preferences and demand, according to which consumers love variety and so their utility increases as the number of goods consumed increases, other things being equal (Dixit & Stiglitz, 1977; Spence, 1976). This means that the consumer will be better off by consuming a greater number of goods at the given prices and income. A simple way of modelling this (Dixit & Stiglitz, 1977) is to assume that the representative consumer has a utility function of the type

$$u = \left(\sum_{i=1}^{n} D_i^{\alpha} \right)^{1/\alpha}, \quad 0 < \alpha < 1, \tag{23.9}$$

where D_i is the quantity consumed of good i, and n the number of goods. This functional form, also used in production theory, is of the well-known constant elasticity of substitution (CES) type, which is homogeneous of degree one in the quantities and has the convenient property that the elasticity of substitution between any two goods is constant, $\sigma = 1/(1-\alpha) > 1$.

The consumer maximises u subject to the budget constraint $I = \sum_{i=1}^{n} D_i p_i$, where I is the consumer's money income. It is a well-known result (Varian, 1992, p. 112) that in the case of a CES utility function the demand functions deriving from the consumer's maximization process have the form

$$D_i = \frac{p_i^{-\sigma}}{\sum_i^n p_i^{1-\sigma}} I. \tag{23.10}$$

To show that utility increases with the number of goods consumed, let us assume that all goods have the same price p. From (23.10) it follows that the optimal quantities of each good will be equal, hence the consumer's income will be divided equally among all available commodities, $D_i = I/np$. By substituting these into the utility function (23.9) we obtain the optimal utility $\overline{u}$

$$\overline{u} = n^{(1-\alpha)/\alpha} \left(\frac{I}{p} \right),$$

which clearly increases as n increases.

The n goods can be taken as the n varieties of a horizontally differentiated product. In the case of m differentiated products, each of which has several varieties, say v_k, the situation is much more complicated. The number of goods will be $n = \sum_{k=1}^{m} v_k$. A convenient way of simplifying the problem is to assume that the overall utility function has the separability property, namely that the subutility deriving from the consumption of the different varieties of a product is independent of the quantities of the varieties of other products being consumed. It follows that the overall utility function can be written as

$$U = U[u_1(.), u_2(.), \ldots, u_m(.)], \tag{23.11}$$

where $u_1(.)$ is the subutility function whose arguments are the different varieties of product 1, and so on. Note that the presence of homogeneous products is easily accommodated, for in the case of a homogeneous product there will only be one variety of it, hence the subutility function relating to it will have one argument only. Thus, if k is a generic good, we have $u_k = u_k(D_k)$ when k is homogeneous, while if k is a differentiated good we have $u_k = u_k(D_{k1}, D_{k2}, \ldots., D_{kv_{k,}})$, where $D_{k\omega}$ (for $\omega = 1, 2, \ldots, v_k$) is the quantity of variety ω that is being consumed and v_k is the number of varieties of good k. The subutility function of any product is assumed to be of the CES type discussed above, which clearly reduces to $u_k = D_k$ in the case of a homogeneous product.

It is well known (Green, 1964, chap. 4) that in the case of homogeneous functional separability the solution of the consumer's maximization problem can be carried out as a two-stage maximization procedure. For each group, namely for each subset consisting of all the varieties of each differentiated good, we define a price index $P_k = f_k(p_{k1}, p_{k2}, \ldots, p_{kn_k})$ as a function of the prices of members of the group, and a quantity index $D_k = g_k(D_{k1}, D_{k2}, \ldots., D_{kn_{k,}})$ as a function of their quantities; both functions must be homogeneous of degree one in their respective arguments for homogeneous separability to obtain. Then the two-stage budgeting procedure is carried out as follows.

First, the optimal distribution of the consumer's given income among the groups is determined by reference to the price and quantity indices alone, namely the subutility functions u_k in U are replaced with the quantity indices D_k, and the utility function $U = U(D_1, D_2, \ldots., D_m)$ is maximised with respect to the D_k's subject to $\sum_{k=1}^{m} P_k D_k = I$. This determines the expenditure $I_k = P_k D_k$ on each group.

Second, the expenditure allocated to the various groups is distributed among the members of the group on the basis of their individual prices, namely by carrying out the maximization of each subutility function taking I_k, the expenditure allocated to the group, as given. It is clear that the second stage can also be carried out first, considering I_k as a parameter to be determined in the subsequent optimal-expenditure-allocation stage taken as second. This is the approach chosen by Dixit and Stiglitz (1977) and followed by Helpman and Krugman, but we prefer to follow the traditional sequence for clarity of exposition.

Let us then consider a model in which there are two commodities, one homogeneous and the other differentiated with n varieties. Let Y be the consumption of the homogeneous good, and D_i the consumption of variety i of the differentiated commodity. The subutility functions are of the CES type, hence turn out to be Y for the homogeneous commodity and $\left(\sum_{i=1}^{n} D_i^{\alpha}\right)^{1/\alpha}$ for the differentiated commodity. The overall utility function is assumed to be

$$U = Y + A\theta^{-1}\left[\left(\sum_{i=1}^{n} D_i^{\alpha}\right)^{1/\alpha}\right]^{\theta}, \qquad 0 < \theta < 1, \tag{23.12}$$

where A is a constant.

To carry out the first stage of the maximization process we must preliminarily define a quantity index and a price index for the differentiated commodity. These are

$$D = \left(\sum_{i=1}^{n} D_i^{\alpha}\right)^{1/\alpha}, \qquad P = \left(\sum_{i=1}^{n} p_i^{\alpha/(\alpha-1)}\right)^{(\alpha-1)/\alpha}, \tag{23.13}$$

which clearly satisfy the condition of being homogeneous of degree one. The first stage consists in maximising $U = Y + A\theta^{-1}D^{\theta}$ with respect to Y and D subject to the budget constraint $I = Y + PD$, where the prices are expressed in terms of the homogeneous good. From the first-order conditions we get

$$AD^{\theta-1} = P,$$

hence

$$D = (A^{-1}P)^{-1/(1-\theta)} = BP^{-\epsilon}, \qquad \epsilon = \frac{1}{1-\theta}, \tag{23.14}$$

is the aggregate demand function, with constant price-elasticity ϵ. Having thus determined PD, the budget allocated to the differentiated good, we can go on to the second stage, where we maximise the subutility function $\left(\sum_{i=1}^{n} D_i^{\alpha}\right)^{1/\alpha}$ subject to the budget constraint $PD = \sum_{i=1}^{n} D_i p_i$. The solution is of type (23.10), namely

$$D_i = \frac{p_i^{-\sigma}}{\sum_i^n p_i^{1-\sigma}} PD = \frac{p_i^{-\sigma}}{P^{-1}\sum_i^n p_i^{1-\sigma}} D. \tag{23.15}$$

If we use the definition of P and the fact that $1-\sigma = -\alpha/(1-\alpha) = \alpha/(\alpha-1)$, we can manipulate the denominator of the last fraction as follows:

$$P^{-1}\sum_{i=1}^{n} p_i^{1-\sigma} = \left(\sum_{i=1}^{n} p_i^{\alpha/(\alpha-1)}\right)^{\frac{-(\alpha-1)}{\alpha}} \left(\sum_{i=1}^{n} p_i^{\alpha/(\alpha-1)}\right) = \left(\sum_{i=1}^{n} p_i^{\alpha/(\alpha-1)}\right)^{\frac{-(\alpha-1)}{\alpha}+1}$$
$$= \left(\sum_{i=1}^{n} p_i^{\alpha/(\alpha-1)}\right)^{1/\alpha}.$$

The last term is clearly $P^{1/(\alpha-1)} = P^{-\sigma}$. Hence the demand for quality i turns out to be

$$D_i = \left(\frac{p_i}{P}\right)^{-\sigma} D, \tag{23.16}$$

which can also be written as

$$D_i = Bp_i^{-\sigma} P^{\sigma-\epsilon} \tag{23.17}$$

since $D = BP^{-\epsilon}$ as shown in (23.14).

This result has an important implication: if a single firm produces good i, and if this firm is small enough with respect to the economy so that it considers itself as unable to influence D and P, it will perceive itself as facing a downward sloping demand curve with constant elasticity σ. This will indeed be the case in a monopolistically competitive market, with imperfect competition due to economies of scale in the production of the several varieties of the differentiated good: given the large number of symmetric potential products, there is no reason for two firms trying to produce the same good. More precisely, if a firm chose a variety that is already produced by another firm, it would have to share the market for this variety: given the equality of the demand curves for the various goods (varieties) when D and P are taken as given, the profits to be gained are clearly lower than the profits that the incumbent firm could make by choosing another variety as yet unproduced. Hence each good will be produced by a different firm.

23.2.2 The Production Side

Let us now turn to the production side. The homogeneous commodity is produced under constant returns to scale in a perfectly competitive market, while the n varieties of the differentiated good are produced under increasing returns to scale in a monopolistically competitive market. Hence the pricing rule of the representative firm producing the homogeneous commodity is *price* = *marginal cost*, which in turn equals average cost at the equilibrium point, given the no-profit condition. The representative monopolistically competitive firm will apply the *marginal revenue* = *marginal cost* pricing rule, with the usual mark-up over price. However, if we assume absent any restriction on entry and exit, monopolistic competition will also reduce profits to zero, hence a selling price equal to average cost in this market as well.

As regards the structure of production, namely the factor inputs, one could consider the traditional two-factor setting (Helpman & Krugman, 1985, chap. 7) , but the essentials of the monopolistic competition approach to international trade can be brought out in a much simpler way if we use the one-factor setting (Helpman & Krugman, 1989; Krugman, 1979, 1980, 1990) .

Let us then assume that there is only one factor, labour (for this reason the model has also been called a "Chamberlinian-Ricardian" model). We first consider the simpler case in which only the differentiated good exists. If $g(x_i)$ is the labour input of the firm producing the quantity x of variety i, we have $g'(x_i) > 0$, but $\mathrm{d}[g(x_i)/x_i]/\mathrm{d}x_i < 0$ due to increasing returns. Marginal cost is $wg'(x_i)$, where w is the given wage rate. From the demand function (23.16) we get the inverse demand function $p_i = (PD^{1/\sigma})D_i^{-1/\sigma}$, hence marginal revenue is $\mathrm{d}(p_i D_i)/\mathrm{d}D_i = \mathrm{d}[(PD^{1/\sigma})D_i^{(\sigma-1)/\sigma}]/\mathrm{d}D_i = [(\sigma-1)/\sigma]p_i$. Thus the pricing rule of the monopolistically competitive firm gives

$$wg'(x_i) = [(\sigma-1)/\sigma]p_i,$$

from which

$$\frac{p_i}{w} = \frac{g'(x_i)\sigma}{\sigma-1}. \tag{23.18}$$

If we assume free entry and exit, we additionally have

$$\frac{p_i}{w} = \frac{g(x_i)}{x_i}. \tag{23.19}$$

These two equations together determine the output and price of the representative firm. Since the demand functions are identical across varieties and the cost functions have also been assumed identical, output per firm and price (relative to the wage rate) turn out to be the same for all varieties produced. It remains to determine the number of varieties produced. This can easily be obtained from the full employment condition and the fact that output per firm is the same and labour input also. Hence $ng(x) = L$, from which

$$n = \frac{L}{g(x)}, \tag{23.20}$$

where x is taken from the previous solution and L is the labour force. We do not know which n goods are produced, but this is unimportant, since all goods are symmetric.

23.2.3 International Trade

If we now consider a world consisting of two such economies, and assume identical tastes (technology needs not be identical, but for simplicity's sake we shall assume that it is), it is easy to see the determinants of international trade. Country 1 will produce n_1 goods and country 2 will produce n_2 different goods. Given the love for variety, each will consume some of the other products, and consumers will be better off since the number of goods increases. Thus there will be mutually beneficial intra-industry trade.

Let us now introduce the homogeneous good into the picture. If we denote by a_{LY} the constant labour-input coefficient in the production of Y, we have $p_Y = wa_{LY}$. Now let us assume that in equilibrium both countries produce some of this good, and that trade in Y can occur costlessly (no transport costs, no tariffs, etc.). Then p_Y must be the same in both countries, and this ties down relative wage rates in the two countries:

$$\frac{w_1}{w_2} = \frac{a_{2LY}}{a_{1LY}}. \tag{23.21}$$

We already know from Eqs. (23.18) and (23.19) the producer price and output of differentiated products in terms of labour and thus also in terms of the homogeneous good. If we denote by x and p the output and the selling price of a generic firm producing a variety of the differentiated commodity, and assume identical technology, x and p will be the same in both countries, and will also be the same across varieties. Let us then consider the varieties which are internationally traded. Clearing of the product market requires output to equal the sum of the two countries' demands, $x = D^1 + D^2$, where D^1, D^2 are given by Eq. (23.17), namely $D^1 = B_1 p^{-\sigma} P_1^{\sigma-\epsilon}$, $D^2 = B_2 p^{-\sigma} P_2^{\sigma-\epsilon}$, where we have omitted the country subscript from p since it is equal in both countries, as we have seen above. We now introduce transport costs of the usual iceberg type, so that for every unit shipped, only $1/(1+\phi)$ units reach the foreign market, where $\phi > 0$. Hence the price to domestic consumers of one unit of an imported good will be $(1+\phi)p$. Taking transport costs into account and letting x_{12} be the quantity produced by country 1 to serve country 2's market we can write the usual *supply* = *demand* condition

$$(1+\phi)^{-1} x_{12} = B_2[p(1+\phi)]^{-\sigma} P_2^{\sigma-\epsilon},$$

whence

$$x_{12} = (1+\phi)^{1-\sigma} B_2 p^{-\sigma} P_2^{\sigma-\epsilon}. \tag{23.22}$$

As regards the domestic market, we have

$$x_{11} = B_1 p^{-\sigma} P_1^{\sigma-\epsilon}, \tag{23.23}$$

where x_{11} is the quantity produced by country 1's firm to serve the domestic market.

It follows that the overall market-clearing condition for country 1's firm can be written

$$x_1 = B_1 p^{-\sigma} P_1^{\sigma-\epsilon} + (1+\phi)^{1-\sigma} B_2 p^{-\sigma} P_2^{\sigma-\epsilon}. \tag{23.24}$$

We similarly find that the market clearing condition for country 2's firm is

$$x_2 = (1+\phi)^{1-\sigma} B_1 p^{-\sigma} P_1^{\sigma-\epsilon} + B_2 p^{-\sigma} P_2^{\sigma-\epsilon}. \tag{23.25}$$

Since $x_1 = x_2 = x$ as seen above, and p is also given, the system consisting of Eqs. (23.24) and (23.25) determines the price indices P_1, P_2 or, to simplify the solution, $P_1^{\sigma-\epsilon}$, $P_2^{\sigma-\epsilon}$, that is to say, *transformations* of the consumer price indices for differentiated products in each country. Note that the fact that producer prices and quantities of each variety are given implies that any change in the price indices is brought about by a change in the number of firms active in each country, as can be immediately seen from the definition of the price index given in Eqs. (23.13).

23.2.3.1 A Simple Gravity Equation

The monopolistic competition model gives rise to the gravity equation in a very simple and direct way. A number of slightly different specification of the gravity equation exist in the literature. Here we derive the *odds and friction*[1] specification since it obtains directly from the model described above. In this specification the dependent variable is the ratio of foreign to domestic trade (purchase from abroad divided by purchase from home). Consider for instance country 1 and let n_i be the number of varieties produced in country i. Recall from (23.24) and (23.25) that x_{21} and x_{11} are, respectively, country 1's imports and domestic sales of a any single variety. Therefore country 1's total import of the differentiated good is n_2 times the imports of a single variety, $n_2 x_{21} = n_2 (1+\phi)^{1-\sigma} B_1 p^{-\sigma} P_1^{\sigma-\varepsilon}$, and the value of domestic trade is $n_1 x_{11} = n_1 B_1 p^{-\sigma} P_1^{\sigma-\varepsilon}$.

The ratio of imports to domestic trade, denoted χ_{12}, is equal to $n_2 x_{21}/n_1 x_{11}$ which gives

$$\chi_{12} = (n_2/n_1)\left(\frac{1}{1+\phi}\right)^{\sigma-1}. \tag{23.26}$$

The term (n_2/n_1) represents the *odds* and the term $(1+\phi)^{1-\sigma}$ represents the *friction* due to trade costs. Recalling that $\sigma > 1$ it is clear that any increase in trade costs reduces the ratio of imports to domestic trade. Equation (23.26) is not suitable for empirical estimation because the number of varieties is rarely available in the data and when it is available is typically subject to large measurement errors. To get around this problem let v_i denote the value of sectorial GDP, $v_i \equiv p x n_i$, where we recall from (23.18) and (23.19) that x is the firm's total output. Now, noting that $n_2/n_1 = p x n_2 / p x n_1$ we can rewrite Eq. (23.26) as

$$\chi_{12} = \frac{v_2}{v_1}(1+\phi)^{1-\sigma}. \tag{23.27}$$

[1]This convenient term is used in Combes, Lafourcade, and Mayer (2005).

which is the simplest *odds and friction* specification. The equation written in this way is more suitable for empirical studies because the value of sectorial GDP is more easily measurable than the number of varieties.

23.3 Homogeneous Goods, Oligopoly, and Trade

23.3.1 A Cournot-Type Model

Brander (1981) and Brander and Krugman (1983) model increasing returns in a very simple way, assuming a cost function (equal in both countries) of the type

$$C(q) = F + cq, \tag{23.28}$$

where F is fixed cost and c the (constant) marginal cost. Transport costs are modelled according to the iceberg assumption, so that if a quantity x is exported from country 1 to country 2, the quantity gx arrives in country 2, where $0 \leq g \leq 1$ is the same for both countries. The higher g, the lower transport costs. The markets are located in the two countries and are really segmented as explained in the text. The two firms, one located in country 1 and the other in country 2, compete in the two markets (for the case in which they compete in a third market only, see Brander & Spencer, 1984, 1985, and below, Sect. 24.4.3.2) and behave as Cournot duopolists.

The demand functions are identical in the two countries, and for simplicity's sake we assume them to be normal (downward sloping) and linear, so that

$$p_1 = a - b(q_{11} + q_{21}), \tag{23.29}$$

$$p_2 = a - b(q_{12} + q_{22}), \tag{23.30}$$

where q_{ij} is the quantity offered by firm i in market j, and $a > 0, b > 0$. We can now specify the profit functions. For firm 1 we have

$$\pi_1 = \{[a - b(q_{11} + q_{21})]q_{11} + [a - b(q_{12} + q_{22})]q_{12}\} - [F + c(q_{11} + \frac{1}{g}q_{12})], \tag{23.31}$$

where we observe that, if the quantity *offered* in market 2 is q_{12}, the corresponding quantity *produced* must be $(1/g)q_{12}$, given transport costs. Similarly for firm 2 we have

$$\pi_2 = \{[a - b(q_{11} + q_{21})]q_{21} + [a - b(q_{12} + q_{22})]q_{22}\} - [F + c(q_{22} + \frac{1}{g}q_{21})]. \tag{23.32}$$

23.3.2 The Equilibrium Solution

Cournot behaviour implies that each firm maximises profit taking as given the quantities offered by the other firm. The first-order conditions for a maximum are

$$\begin{aligned}
\frac{\partial \pi_1}{\partial q_{11}} &= [-2bq_{11} - bq_{21} + a] - c = 0, \\
\frac{\partial \pi_1}{\partial q_{12}} &= [-2bq_{12} - bq_{22} + a] - c/g = 0, \\
\frac{\partial \pi_2}{\partial q_{21}} &= [-2bq_{21} - bq_{11} + a] - c/g = 0, \\
\frac{\partial \pi_2}{\partial q_{22}} &= [-2bq_{22} - bq_{12} + a] - c = 0,
\end{aligned} \tag{23.33}$$

whose solution will yield the optimal quantities q_{ij} provided that the second-order conditions are satisfied. The Hessian of firm 1's profit function is

$$\begin{bmatrix} -2b & 0 \\ 0 & -2b \end{bmatrix},$$

whose leading principal minors alternate in sign, starting from minus. Hence the second-order conditions are satisfied. The same holds for firm 2.

The four first-order conditions can be interpreted, as usual, as the equality between marginal revenue and marginal cost for each firm in each market. Note that marginal cost (and hence marginal revenue) for delivering an export unit (c/g) is higher than for a unit of domestic sales (c) owing to transport costs.

Equations (23.33) also define the reaction functions implicitly. For example, if we solve the first equation for q_{11}, we get the optimal quantity offered by firm 1 in market 1 in terms of the quantity offered by firm 2 in the same market. This reaction curve is

$$q_{11} = -\frac{1}{2}q_{21} + \frac{c-a}{2b}. \tag{23.34}$$

It can now be observed that the system of the four first-order conditions is separable: the first and third equation, in fact, only contain the two unknowns q_{11}, q_{21} and can be solved independently of the two other. Similarly, the second and fourth equations independently determine the unknowns q_{22}, q_{12}. This separability property depends on the constant marginal cost assumption, for if marginal cost were a function of output, q_{12} would enter the first equation, q_{11}would enter the second equation and so on; the four equations would all be linked. We also observe that the two subsystems are perfectly symmetric, so that the set of solutions to the first is also the set of solutions to the second, with $q_{11} = q_{22}$ and $q_{12} = q_{21}$. Hence we need consider only one subsystem, for example the first. This is a simple linear system, whose solution is

$$q_{11}^E = \frac{a + c/g - 2c}{3b}, \tag{23.35}$$

$$q_{21}^E = \frac{a + c - 2c/g}{3b}. \tag{23.36}$$

We are interested in a positive solution for q_{21}^E, the amount of "invasion" of country 2's firm into market 1, because $q_{21} = 0$ (and hence $q_{12} = 0$ as well, given the symmetry of the two subsystems) would mean no international trade. It is easy to see that for q_{21}^E to be positive we must have

$$g > \frac{2c}{a + c}, \tag{23.37}$$

which means that transport costs must be below a certain critical level before invasion will occur (recall that transport costs are inversely related to g). When transport costs tend to zero ($g \to 1$), the solution will tend to the Cournot solution

$$q_{11}^E = q_{21}^E = \frac{a - c}{3b}, \tag{23.38}$$

while for positive transport costs $q_{21}^E < q_{11}^E$, as can easily be determined from (23.35) and (23.36), namely the domestic firm has a higher share of the domestic market than the foreign firm. It is also easy to see that q_{11}^E decreases as g increases (i.e., as transport costs decrease), and that q_{21}^E increases as g increases. Hence the foreign firm's share of the domestic market increases, and that of the domestic firm decreases, as transport costs decrease, both approaching $1/2$. The opposite obviously holds when g decreases.

Since each firm has a smaller share of the foreign market than of the domestic market, marginal revenue is higher in the foreign market than in the domestic market, which we already knew from the first-order conditions. But there is more to it than that. Given the symmetry conditions, the overall quantity supplied to each market will be the same in both markets, hence the price also will be the same in both markets. If we now recall that a firm's mark-up over cost is defined as $(p - MC)/p$, where p is the selling price and MC the marginal cost, it follows that each firm's mark-up over cost is lower in its export market than in its domestic market. In fact, $(p - c/g)/p < (p - c)/p$ due to transport costs. Since the selling price is the same in both markets, and transport costs are borne by the exporting firm, the f.o.b. price of exports is below the domestic price, and—as Brander and Krugman (1983) note—there is *reciprocal* dumping.

23.3.3 Stability

Let us now come to stability. The usual way of modelling the dynamic process underlying the reactions is to introduce a lag. Given the quantity offered in period t by firm 2 in the market under consideration, firm 1 will use its own reaction curve to determine the quantity that it will offer in the next period. Firm 2 will act similarly. This amounts to considering the system of difference equations

$$q_{11,t+1} = -\frac{1}{2}q_{21,t} + \frac{c-a}{2b}, \tag{23.39}$$

$$q_{21,t+1} = -\frac{1}{2}q_{11,t} + \frac{c/g-a}{2b}. \tag{23.40}$$

The roots of the characteristic equation of this system are 1/2, −1/2 (for the procedure see Gandolfo, 2009, chaps. 9 and 10, sect. 10.1). Since they are both less than unity in absolute value, the equilibrium is dynamically stable.

23.4 Vertically Differentiated Goods, Oligopoly, and Trade

The model that we present is based on the works of Gabszewicz, Shaked, Sutton, and Thisse (1981) and Shaked and Sutton (1982, 1983, 1984).

23.4.1 Consumers

There is a continuum of consumers who are assumed to have identical tastes, but different incomes, which are uniformly distributed over some interval $0 < a \leq I \leq b$. There are n vertically differentiated goods which are ranked according to quality in the same way by all consumers, say

$$0 < u_1 < \ldots < u_n, \tag{23.41}$$

where $u_k, k = 1, \ldots, n$, is the universally accepted measure of the quality of good k. Given that n may be large, for the moment we are in a context of monopolistic competition rather than of oligopoly, but the model will end up in an oligopolistic situation, as we shall see.

Given the difference in income, richer consumers are willing to pay more for a higher quality product. Each consumer makes indivisible and mutually exclusive purchases from among the n substitute goods, in the sense that any consumer either buys exactly one unit of the chosen good or buys nothing. The utility function of the representative consumer is denoted by $U(I, k)$, which indicates the utility achieved by consuming one unit of good k, and I units of income on "other things" (the latter are referred to by Gabszewicz et al. (1981) and by Shaked and Sutton (1983, 1984), as a Hicksian composite commodity). The utility obtained from consuming I units of income only is indicated by $U(I,0)$.

These properties can be captured by a simple utility function of the form

$$U(I,k) = u_k \cdot I \quad \text{for } k = 1,2,\ldots,n; \text{ and } U(I,0) = u_0 \cdot I, \tag{23.42}$$

where $u_0 > 0$ is conventionally taken to be smaller than u_1. If we denote by p_k the price of good k in terms of I, with the assumption that p_k increases as the quality increases, it is easy to see that a consumer having a given income $\overline{I}$ can obtain a utility

$$U(\overline{I} - p_k, k) = u_k \cdot (\overline{I} - p_k) \tag{23.43}$$

by devoting p_k units of income to the purchase of one unit of good k and $(\overline{I} - p_k)$ to "other things".

We can now define an income level I_k such that a consumer endowed with this income will be indifferent between good k at price p_k and good $k-1$ at price p_{k-1}. Using (23.43) and taking account of the second definition in (23.42), we have

$$u_k \cdot (I_k - p_k) = u_{k-1} \cdot (I_k - p_{k-1}), \tag{23.44}$$

$$u_1 \cdot (I_1 - p_1) = u_0 I_1, \tag{23.45}$$

respectively for $k > 1$ and $k = 1$.

If we define

$$r_{k-1,k} = \frac{u_k}{u_k - u_{k-1}}, \tag{23.46}$$

which is clearly greater than one, from (23.44) we get

$$\begin{aligned} I_k &= u_k p_k/(u_k - u_{k-1}) - u_{k-1} p_{k-1}/(u_k - u_{k-1}) \\ &= r_{k-1,k} p_k + (1 - r_{k-1,k}) p_{k-1} \\ &= p_{k-1} + (p_k - p_{k-1}) r_{k-1,k}, \end{aligned} \tag{23.47}$$

for all $k > 1$, and from (23.45)

$$I_1 = p_1 r_{0,1}$$

for the case of indifference between consuming no differentiated good and consuming the lowest quality of it.

It can easily be shown that a consumer with income above I_k will prefer the higher-quality good k at price p_k to the lower-quality good $k-1$ at price p_{k-1}, while a consumer with income below I_k will do exactly the opposite. Let us consider a consumer with income $I_k + \mathrm{d}I$, where $\mathrm{d}I \gtrless 0$, and I_k is as defined in Eqs. (23.44) and (23.45). Then the consumers' utility deriving from the consumption of good k or $k-1$ is respectively

$$\begin{aligned} U_k &= u_k \cdot (I_k + \mathrm{d}I - p_k) = u_k \cdot (I_k - p_k) + u_k \cdot \mathrm{d}I, \\ U_{k-1} &= u_{k-1} \cdot (I_k + \mathrm{d}I - p_{k-1}) = u_{k-1} \cdot (I_k - p_{k-1}) + u_{k-1} \cdot \mathrm{d}I, \end{aligned}$$

from which, given Eqs. (23.44) and (23.45), we immediately obtain $U_k \gtrless U_{k-1}$ according as $(u_k - u_{k-1})\mathrm{d}I \gtrless 0$, i.e. according as $\mathrm{d}I \gtrless 0$, since $u_k - u_{k-1} > 0$ given (23.41). This result is of course a consequence of the fact that the utility function has been designed just to have the property that higher-income consumers are willing to spend more to get a higher-quality good.

23.4.2 Firms, and Market Equilibrium

The behaviour of firms is based on a three-stage non-cooperative game. In the first stage firms decide whether or not to enter the industry. In the second stage each firm chooses the quality of its product (each firm is assumed to produce only one good). In the third stage each firm chooses its price, and only variable costs enter the pricing decision, given the assumption that all fixed costs have been incurred in the previous stages and are sunk costs. This three-stage process, as Shaked and Sutton (1982) observe, is meant to capture what happens in reality: the price can be varied easily, but a change in the specification of a product involves modification in the appropriate production facilities, and entry into the industry requires construction of a plant.

The solution that the authors seek is a perfect equilibrium, namely an n-tuple of strategies such that, after any stage, that part of the strategies pertaining to the game consisting of the remaining stages form a Nash equilibrium in that game. This allows Shaked and Sutton (1982, 1983, 1984) to study the three-stage game by first examining price competition in the third stage, taking qualities as given. This amounts to considering the short run. In the long run all stages of the game have to be considered, and the qualities are endogenously determined.

From our previous treatment of consumer's choice it follows that consumers are partitioned into segments or income brackets corresponding to the successive market shares of rival firms. More precisely, if we assume that each firm only produces one good, firm k will sell to consumers with income I_k to I_{k+1} for $k < n$ (with income I_k to b for firm n), where I_k, I_{k+1} are given by (23.47). Since each consumer buys one unit of the good, and there is a continuum of consumers, the number of units sold by firm k will be $(I_{k+1} - I_k)$. It is important to observe that a firm may be "just" excluded from the market in the sense that $I_k - I_{k-1} = 0$, so that this firm has a market share of zero, but a slight (infinitesimal) decrease in its price or a slight increase in the price set by any of its two neighbouring firms will make its market share positive.

Unit variable cost is assumed to be an increasing function of the quality but independent of the level of output, hence we denote it by c_k. Therefore the profit of any firm k becomes, for $k = 1, 2, \ldots, k-1$ and for $k = n$ respectively

$$\pi_k = (p_k - c_k)(I_{k+1} - I_k). \qquad \text{and} \qquad \pi_n = (p_n - c_n)(b - I_n). \tag{23.48}$$

If $p_k < c_k$, the firm will undergo losses and hence that quality will not be produced. Also note that for p_k sufficiently high the sales of the firm will be zero. Hence, we consider only the range in which $\pi_k > 0$.

The next question is whether an equilibrium exists. This will be a non-cooperative equilibrium (Nash equilibrium), namely a price vector such that, for any firm k, given the prices set by the other firms, the price fixed by firm k is its profit maximising price. To show that such equilibrium exists, Shaked and Sutton (1983, p. 1475) begin by proving the following

> Lemma I: For any given products $u_1, u_2, \ldots, u_n$ and corresponding prices $p_1, p_2, \ldots, p_n$, for all k, the profit of the kth firm is a single peaked function of its price.

The market share of any firm k is included between that of two neighbouring firms, $k+1$ and $k-1$. As p_k falls, it may happen that one (or both) neighbouring firm, say firm $k-1$, is driven out of the market, so that firm k will acquire firm $k-2$ as a new neighbour. We first consider the case in which the neighbours are firm $k-1$ and $k+1$. If we examine the profit function (23.48) we find that any turning point of π_k is a maximum, so that π_k is single peaked. In fact, we have

$$\pi_k' = (I_{k+1} - I_k) + (p_k - c_k)\frac{\mathrm{d}}{\mathrm{d}p_k}(I_{k+1} - I_k),$$

and from (23.47) we get

$$I_{k+1} - I_k = [p_k + (p_{k+1} - p_k)r_{k,k+1}] - [p_{k-1} + (p_k - p_{k-1})r_{k-1,k}],$$

whose derivative with respect to p_k is $(1 - r_{k,k+1} - r_{k-1,k})$. Thus we have

$$\begin{aligned} \pi_k' &= (p_k - c_k)(1 - r_{k,k+1} - r_{k-1,k}) + (I_{k+1} - I_k) \\ \pi_k'' &= 2(1 - r_{k,k+1} - r_{k-1,k}) < 0. \end{aligned} \tag{23.49}$$

Since the second derivative is always negative, any turning point of π_k is a global maximum, hence π_k is single peaked. It can easily be checked that such property also holds for $k = n$.

We must now consider the case in which p_k falls sufficiently for driving the neighbouring firm $k-1$ out of the market, so that the new neighbours are firms $k-2$ and $k+1$. By the same procedure used above, it can be seen that with these new neighbours firm k's profit (say, $\hat{\pi}_k = (p_k - c_k)(I_{k+1} - I_{k-1})$) remains a single peaked function of p_k. We know that a zero market share of firm $k-1$ means $I_k - I_{k-1} = 0$, and we show that at the price at which this happens we also have

$$\hat{\pi}'_k > \pi'_k, \tag{23.50}$$

so that $\hat{\pi}_k$ is a fortiori increasing at this point if π_k is increasing there. By the same procedure used for computing π'_k we get

$$\hat{\pi}'_k = (p_k - c_k)(1 - r_{k,k+1} - r_{k-2,k}) + (I_{k+1} - I_{k-1}). \tag{23.51}$$

By using $I_k = I_{k-1}$ and the fact that the definition (23.46) implies $r_{k-2,k} < r_{k-1,k}$, we can easily see from (23.49) and (23.51) that (23.50) does indeed hold. This completes the demonstration of the lemma.

The lemma implies that each firm's profit function is quasi-concave, hence (Friedman, 1977, p. 157) a noncooperative price equilibrium $p_1, p_2, \ldots, p_n$ exists for any set of products $1, 2, \ldots, n$.

The next step of the analysis is to prove that, under normal conditions (namely when all consumers strictly rank the goods in the same way, as assumed at the beginning) the market has the finiteness property. This means that, at any Nash equilibrium involving a number of products drawn from the existing interval of qualities, there is an upper bound B to the number of firms which enjoy positive market shares and prices exceeding unit variable cost. This can be shown quite simply for the particular case studied by Gabszewicz et al. (1981) and Shaked and Sutton (1982) , in which variable cost is assumed to be zero and the distribution of income is not much dispersed, namely $b < 4a$. In this case at most two products (the top two) have a positive market share at equilibrium. The first order condition for a maximum implies $\pi'_k = 0$, hence from (23.49) we have, letting $c_k = 0$,

$$\begin{aligned} p_k(1 - r_{k,k+1}) - p_k r_{k-1,k} + I_{k+1} - I_k &= 0, \\ -p_n r_{n-1,n} + b - I_n &= 0. \end{aligned} \tag{23.52}$$

From the definition of I_k given in (23.47) we get $p_k r_{k-1,k} = I_k + (r_{k-1,k} - 1)p_{k-1}$, and by substituting this into (23.52) we get

$$\begin{aligned} p_k(1 - r_{k,k+1}) - (r_{k-1,k} - 1)p_{k-1} + I_{k+1} - 2I_k &= 0, \\ -p_{n-1}(r_{n-1,n} - 1) + b - 2I_n &= 0, \end{aligned}$$

whence

$$\begin{aligned} I_{k+1} &= 2I_k + p_k(r_{k,k+1} - 1) + (r_{k-1,k} - 1)p_{k-1} > 2I_k, \\ b &= 2I_n + p_{n-1}(r_{n-1,n} - 1) \qquad\qquad\qquad\quad > 2I_n, \end{aligned} \tag{23.53}$$

where the inequalities follow from the fact that the r's, as defined in (23.46), are greater than one. From (23.53) we get, by letting $k + 1 = n$ in the first equation,

$$4I_{n-1} < b. \tag{23.54}$$

Now by assumption $b < 4a$, which in conjunction with (23.54) implies

$$I_{n-1} < a. \tag{23.55}$$

This inequality means that I_{n-1} is already lower than the lower bound to the distribution of income. Since I_{n-1} is the income at which a consumer endowed with this income is indifferent between good $n-1$ at price p_{n-1} and good $n-2$ at price p_{n-2} (see Sect. 23.4.1), it follows that no such income exists, and hence the two top firms (n and $n-1$) cover the market. For further reference note that this result can be strengthened to the case in which $2a < b < 4a$ (Shaked & Sutton, 1982). For the general case of the finiteness property see Shaked and Sutton (1983).

This completes the study of the third stage of the game. In the second stage (choice of quality) a Nash equilibrium exists that involves two distinct qualities produced by two firms that both earn positive profits. The entry of further firms would lead to a configuration in which all firms would earn zero profits. It can also be shown that the top quality firm enjoys a greater revenue than its rival, and that the revenues of both firms increase as the quality of the better product improves. Finally, the examination of the first stage of the game, which involves the decision whether to enter or not the market (it is at this stage that the fixed costs are assumed to be incurred), allows to conclude that a perfect equilibrium exists in which two firms enter, produce distinct products, and have positive profits. No perfect equilibrium exists in which more than two firms enter. The proofs of the results concerning the second and first stage are rather lengthy, hence we refer the reader to Shaked and Sutton (1982).

23.4.3 International Trade

The extension of this model to international trade is straightforward. If we start from two identical economies, then in autarky each will support the same B goods by the finiteness property. When free trade (no transport costs are assumed) is opened, the combined world economy will have the same properties of the two identical autarkic economies (same income distribution, etc.). Hence it will support the same B goods, and international trade will be generated by the fact that consumers in both countries demand the same B goods as before, which will be produced partly in one country and partly in the other. Consumers will be better off as shown in the text.

When the two countries are not identical, then the combined world economy can support more than the number of goods supported by each in isolation. Gabszewicz et al. (1981) study the simplified case (see above) in which each of the two autarkic economies supports two goods, which means that $2a < b < 4a$ in both countries; all four goods are assumed to be different. The income distributions in the two countries are different, but not too much: more precisely, there exists an overlapping interval a_1, a_2 such that

$$\frac{a_1}{2} < a_2 < a_1 < \frac{b_1}{2} < b_2 < b_1,$$

where (a_1, b_1) and (a_2, b_2) are the intervals over which the income distributions in the two countries are defined. Now, the market share of the highest-quality good n will extend below $b_1/2$, that of good $(n-1)$ will extend below a_1, and that of the third good $(n-2)$ will extend below a_2 (see Gabszewicz et al., 1981). Hence, these three goods will cover the market.

For further developments of this approach to international trade see Motta (1992).

23.5 Horizontal Differentiation, Oligopoly, and Trade

The Eaton and Kierzkowski (1984) model assumes that there are only two basic commodities, one homogeneous (A) and the other horizontally differentiated (B). The homogeneous commodity plays a secondary role: it only serves to allow consumers to spend income when they do not purchase the differentiated commodity. Thus, the analysis can be concentrated on the latter.

23.5.1 Demand for Characteristics

The demand for the differentiated commodity follows Lancaster's approach, with the simplifying assumption that such commodity only contains one characteristic Z (which can be measured by a real number). Each consumer has an ideal model of good B represented by a value of Z, say θ_i, where i denotes the consumer. The consumer's utility declines as the model actually consumed becomes more distant from the ideal. Hence consumer i will buy an alternative model only if the price of this non-ideal alternative is sufficiently lower than the price of the ideal model. Finally, if the price for all available varieties of the commodity exceeds a certain upper limit, the consumer will not buy this commodity, and concentrate expenditure on the homogeneous good A.

The formalisation corresponding to these assumptions is the utility function (suggested by Salop, 1979)

$$U(Y, p_i, \theta_i, Z_i, \overline{p}) = \max[Y - (p_i + |\theta_i - Z_i|), Y - \overline{p}], \tag{23.56}$$

where Z_i is the model consumed by individual i, p_i the price paid for it, Y income, and $\overline{p}$ the maximum price. This utility function has the following characteristics: At most one unit of the differentiated commodity will be purchased. The maximum price that the individual is willing to pay is $\overline{p}$, provided that the model corresponds exactly to the ideal, namely $Z_i = \theta_i$. When there is no such correspondence, the individual will be willing to purchase the non-ideal model at a price not higher than $\overline{p} - |\theta_i - Z_i|$; clearly, this price is the lower, the greater the distance from the ideal. In general, the consumer will choose the model for which $p_i + |\theta_i - Z_i|$ is at a minimum, if this amount is less than or equal to $\overline{p}$. Hence if there is no model for which this is true, namely if $p_i + |\theta_i - Z_i| > \overline{p}$, i.e. $p_i > \overline{p} - |\theta_i - Z_i|$ for all existing models, the consumer will not purchase the differentiated commodity.

23.5.2 The Production Side

Let us now consider the production side. There are increasing returns to scale, and the total cost of producing an amount x of a particular model is assumed to be $K + cx$, where the marginal cost c is constant and the fixed cost K is a sunk cost, namely it must be incurred by the firm at the moment of the choice of the model to produce, before the levels of output and price are determined. Hence when the firm sets these levels, the cost K is sunk and the model is already determined. Finally, a single firm can produce only a single model of the commodity.

In spite of its apparent simplicity, this model gives rise to a rich taxonomy, according to the number of firms (one, hence monopoly, or two, hence duopoly) and to the categories of consumers (the types of consumers are distinguished according to the type of ideal model). Here we consider only one case, referring the reader to Eaton and Kierzkowski (1984) for the others. It is the case of two types of consumers and one firm.

There are n_1 consumers of type 1, all having θ_1 as the ideal model, and n_2 of type 2, with ideal model θ_2. The single firm can produce a single model (say Z_1), which can be assumed to be closer to the ideal θ_1 without loss of generality. Assuming that price discrimination is not possible, the firm must decide whether:

(a) Not to produce at all; or
(b) To produce and sell to just the type of consumers whose ideal is closer to Z_1. This means charging the limit price for type 1 consumers, $\overline{p} - |\theta_1 - Z_1|$. Since θ_1 is closer to Z_1 than θ_2, this lim it price will certainly be higher than $\overline{p} - |\theta_2 - Z_1|$, the limit price for type 2 consumers, who will not buy the commodity. Hence the firm's current profit will be

$$\pi_1 = (\overline{p} - |\theta_1 - Z_1| - c)n_1 - K; \tag{23.57}$$

or

(c) To sell to both types of consumers, charging the limit price for type 2 consumers, $\overline{p} - |\theta_2 - Z_1|$. Since this is lower than the limit price for type 1 consumers, these will also buy the commodity, and the firm's current profit will be

$$\pi_{1,2} = (\overline{p} - |\theta_2 - Z_1| - c)(n_1 + n_2) - K. \tag{23.58}$$

If $\overline{p} - |\theta_1 - Z_1| < c$, the firm will not produce, since the highest selling price, which is the limit price for type 1 consumers, is below marginal cost. Excluding this case, alternative (c) or (b) will be selected according as $\pi_{1,2} \gtrless \pi_1$, from which, after simple manipulations,

$$(\overline{p} - c)(1 - \lambda) \gtrless |\theta_2 - Z_1| - |\theta_1 - Z_1|\,\lambda, \tag{23.59}$$

where λ is defined as the proportion of type 1 consumers in the overall market, namely $n_1/(n_2 + n_2)$. It is easy to see that selling to the broader set of consumers is the superior alternative when $\overline{p} - c$ is high, λ is low, and the two ideal qualities are not substantially different.

23.5.3 International Trade

The extension of this model to international trade is straightforward, if we assume that in the home country, where the producing firm is located, there are only type 1 consumers, say n_1 (with ideal model θ_1), while in the foreign country (denoted by an asterisk) there are only type 2 consumers, say n_2^*, whose ideal model is θ_2. With no loss of generality we can assume $\theta_2 > \theta_1$. In autarky, the firm produces exactly model θ_1 and charges the maximum price $\overline{p}$. When trade is opened, the firm will consider exporting to the foreign market. If we assume no transport costs and no segmentation, the firm can sell to both types of markets at the (lower) price that type 2 consumers are willing to pay, namely, since $Z_1 = \theta_1$, at the price $\overline{p} - |\theta_2 - \theta_1| = \overline{p} - (\theta_2 - \theta_1)$. If we compare the firm's profits in the two situations, we find that the firm will begin to export to the foreign market if

$$[\overline{p} - (\theta_2 - \theta_1) - c](n_1 + n_2^*) > (\overline{p} - c)n_1. \tag{23.60}$$

We see that the more similar are the demand patterns and the larger is the foreign country, the more likely is that trade will take place. Similarity in demand patterns is again, contrary to conventional theory, a cause of trade. Another important difference with respect to the traditional theory is that trade is indifferent to the foreign country, that will receive no gain. This occurs because the sole producer of the differentiated product will be able to fix the price at a level that will leave foreign consumers indifferent between consuming only the homogeneous good (as before trade) and both the differentiated and homogeneous good. The domestic consumers will benefit from lower prices and the domestic producer's profit will be larger. Hence, the domestic country's welfare improves while the foreign country receives no benefit from trade.

References

Brander, J. A. (1981). Intra-industry trade in identical commodities.

Brander, J. A., & Krugman, P. R. (1983). A reciprocal dumping model of international trade.

Brander, J. A., & Spencer, B. J. (1984). Tariff protection and imperfect competition.

Brander, J. A., & Spencer, B. J. (1985). Export subsidies and international market share rivalry.

Combes, P.-P., Lafourcade, M., & Mayer, T. (2005). The trade-creating effects of business and social networks: Evidence from France.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity.

Eaton, J., & Kierzkowski, H. (1984). Oligopolistic competition, product variety and international trade.

Falvey, R. E. (1981). Commercial policy and intra-industry trade.

Friedman, J. W. (1977). *Oligopoly and the theory of games*.

Gabszewicz, J. J., Shaked, A., Sutton, J., & Thisse, J.-F. (1981). International trade in differentiated products.

Gandolfo, G. (2009). *Economic dynamics*.

Green, H. A. J. (1964). *Aggregation in economic analysis*.

Helpman, E. (1990). Monopolistic competition in trade theory.

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade*.

Helpman, E., & Krugman, P. R. (1989). *Trade policy and market structure*.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade.

Krugman, P. R. (1980). Scale economies, product differentiation, and the pattern of trade.

Krugman, P. R. (1990). *Rethinking international trade*.

Motta, M. (1992). Sunk costs and trade liberalization.

Salop, S. C. (1979). Monopolistic competition with outside goods.

Shaked, A., & Sutton, J. (1982). Relaxing price competition through product differentiation.

Shaked, A., & Sutton, J. (1983). Natural oligopolies.

Shaked, A., & Sutton, J. (1984). Natural oligopolies and international trade.

Spence, A. M. (1976). Product selection, fixed costs, and monopolistic competition.

Varian, H. R. (1992). *Microeconomic analysis*.

Chapter 24
Appendix to Chapter 10

24.1 Tariffs, Terms of Trade, Domestic Relative Price

If we assume that country 1 imports commodity A and exports commodity B whilst the opposite holds for country 2, international equilibrium is determined in accordance with Eq. (19.27), which we rewrite here

$$E_{2B}(p) + E_{1B}(p) = 0, \tag{24.1}$$

or

$$E_{2B}(p) = -E_{1B}(p), \tag{24.2}$$

that is, the excess demand for commodity B by country 2 (country 2's demand for imports) is equal in absolute value to the excess supply of this commodity by country 1 (country 1's supply of exports).

In the case that a country, say country 2, levies a duty, the domestic relative price of that country—to which its economic agents respond—is no longer p, but $p_d = p(1+d)$. Therefore E_{2B} will be a function of p_d instead of p. Besides, we must introduce the spending of the revenue by the government, which in real terms is dE_{2B}. We assume that the government spends a fraction $0 < \varphi < 1$ of this revenue to purchase commodity B and the remaining fraction $(1-\varphi)$ to purchase commodity A; consequently country 2's total (private + public) demand for imports will be $(1+\varphi d)E_{2B}$.

Thus have the relations

$$\begin{aligned} (1+\varphi d)\, E_{2B}(p_d) + E_{1B}(p) &= 0, \\ p_d - p(1+d) &= 0. \end{aligned} \tag{24.3}$$

Equation (24.3) constitute a set of two implicit functions in three variables (p_d, p, d). Therefore, provided that the Jacobian of these functions with respect

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_24,

to p_d and p is different from zero at the equilibrium point, by the implicit function theorem we can express p_d and p as differentiable functions of d in a neighbourhood of the equilibrium point and perform exercises in comparative statics. In particular, we are interested in the effects on p and p_d of the introduction of a tariff and in determining the conditions for the Metzler and Lerner cases to occur.

The Jacobian of (24.3) is

$$J = \begin{vmatrix} (1+\varphi d)\, E'_{2B} & E'_{1B} \\ 1 & -(1+d) \end{vmatrix} = -(1+\varphi d)(1+d)\, E'_{2B} - E'_{1B}, \tag{24.4}$$

which, evaluated at the initial (free trade) equilibrium point (hence $d = 0$), becomes

$$J = -\left(E'_{2B} + E'_{1B}\right). \tag{24.5}$$

If we multiply and divide by E_{2B}/p we get

$$J = -\frac{E_{2B}}{p}\left(E'_{2B}\frac{p}{E_{2B}} + E'_{1B}\frac{p}{E_{2B}}\right), \tag{24.6}$$

and, since $E_{2B} = -E_{1B}$ in the initial equilibrium situation (in which (24.1) holds), we have

$$J = -\frac{E_{2B}}{p}\left(E'_{2B}\frac{p}{E_{2B}} - E'_{1B}\frac{p}{E_{1B}}\right),$$

that is, by using the definitions of the elasticities given in (19.44) and (19.37),

$$J = -\frac{E_{2B}}{p}\left(\xi_2 - \varepsilon_1\right). \tag{24.7}$$

By using the relation (see (19.42)) $\varepsilon_1 = -(1+\xi_1)$, we finally get

$$J = -\frac{E_{2B}}{p}\left(1 + \xi_1 + \xi_2\right). \tag{24.8}$$

If we now apply Samuelson's correspondence principle and assume that the equilibrium is stable on the basis of the dynamic process of adjustment to excess demand, we can use condition (19.49), that is

$$(1 + \xi_1 + \xi_2) < 0, \tag{24.9}$$

and so

$$J > 0.$$

Let us now calculate p_d' and p', the derivatives of p_d and p with respect to d. By totally differentiating system (24.3) with respect to d we get

$$\begin{aligned} \varphi E_{2B} + (1+\varphi d)\, E_{2B}' p_d' + E_{1B}' p' &= 0, \\ p_d' - p'(1+d) - p &= 0, \end{aligned} \tag{24.10}$$

that is, by using the fact that the derivatives are computed at the initial free-trade equilibrium situation ($d = 0$), and rearranging terms,

$$\begin{aligned} E_{2B}' p_d' + E_{1B}' p' &= -\varphi E_{2B}, \\ p_d' - p' &= p. \end{aligned} \tag{24.11}$$

If we solve for p_d' and p' we get

$$\begin{aligned} p_d' &= \frac{\varphi E_{2B} - E_{1B}' p}{J}, \\ p' &= \frac{\varphi E_{2B} + E_{2B}' p}{J}. \end{aligned} \tag{24.12}$$

By replacing J with expression (24.8) we get

$$\begin{aligned} p_d' &= \frac{p}{E_{2B}} \frac{\varphi E_{2B} - E_{1B}' p}{-(1+\xi_1+\xi_2)} = p \frac{\varphi - E_{1B}' \frac{p}{E_{2B}}}{-(1+\xi_1+\xi_2)} \\ &= p \frac{\varphi + E_{1B}' \frac{p}{E_{1B}}}{-(1+\xi_1+\xi_2)} = p \frac{\varphi + \varepsilon_1}{-(1+\xi_1+\xi_2)}. \end{aligned} \tag{24.13}$$

Similarly we obtain

$$p' = p \frac{\varphi + \xi_2}{-(1+\xi_1+\xi_2)}. \tag{24.14}$$

Given condition (24.9), the sign of p_d' and p' depends only on the numerator of the relevant fraction.

It should be remembered that *Metzler's case* (Metzler, 1949) occurs when, as a consequence of the imposition of a tariff by country 2 on its imports of B, this country's domestic relative price (p_B/p_A) decreases, instead of increasing, with respect to that (equal to the terms of trade) existing in the initial free trade situation. Formally, this amounts to $p_d' < 0$, that is, $\varphi + \varepsilon_1 < 0$. Since $\varepsilon_1 = -(1+\xi_1)$ from (19.42), we have

$$\varphi - \xi_1 - 1 < 0, \tag{24.15}$$

that is,

$$\varphi - \xi_1 < 1. \tag{24.16}$$

In the normal case (non-inferior goods etc.) the elasticity $\xi_1$is negative, so that the condition for Metzler's case to occur is that the sum of the fraction φ and the absolute value of the elasticity of the rest-of-the-world's demand for imports should be smaller than one. This is equivalent to saying that the rest-of-the-world's import demand must be sufficiently rigid. If, on the contrary, we have an abnormal case (for example, commodity A is an inferior good for country 1), the elasticity ξ_1 is positive and (24.16) is satisfied for any non negative φ. This is the case illustrated graphically in Fig. 10.5.

As regards *Lerner's case* (Lerner, 1936), this occurs when, after the imposition of tariff, the terms of trade are higher, instead of being lower, than in the initial free trade situation. In formal terms this means $p' > 0$, that is, given (24.14), $\varphi + \xi_2 > 0$ or

$$-\xi_2 < \varphi. \tag{24.17}$$

As before, two cases must be distinguished. In the normal case the elasticity ξ_2 is negative, so that the condition for Lerner's case to occur is that the tariff-imposing country's demand for imports is sufficiently rigid, with an elasticity in absolute value smaller than the fraction φ. On the contrary, in abnormal cases (for example, when commodity B is an inferior good for country 2), the elasticity ξ_2 is positive and (24.17) is verified for any non-negative φ. This is the case illustrated graphically in Fig. 10.6.

24.2 Cartels

Let q_i be the quantity produced by the i–th country participating in the cartel, and $C_i(q_i)$ the corresponding total cost. The whole output $q = \sum_{i=1}^{n} q_i$, is sold by the cartel as a monopolist. If we denote total revenue by $R = p \cdot q$, where p is related to q through the demand function, the problem is to maximize the profit function

$$\pi = R(q) - \sum_{i=1}^{n} C_i(q_i) = R\left(\sum_{i=1}^{n} q_i\right) - \sum_{i=1}^{n} C_i(q_i). \tag{24.18}$$

The first order conditions are

$$\frac{\partial \pi}{\partial q_i} = R' - C_i' = 0, \tag{24.19}$$

that is,

$$R' - C_1' = C_2' = \ldots = C_n'. \tag{24.20}$$

Marginal cost in each country must equal the marginal revenue of the output as a whole.

The second order conditions require that the leading principal minors of the Hessian

$$\begin{bmatrix} R'' - C_1'' & R'' & R'' & \ldots & R'' \\ R'' & R'' - C_2'' & R'' & \ldots & R'' \\ \ldots.. & \ldots.. & \ldots.. & \ldots.. & \ldots.. \\ R'' & R'' & R'' & \ldots & R'' - C_n'' \end{bmatrix} \tag{24.21}$$

alternate in sign, beginning with minus. In the normal case, $R'' < 0$ and $C_i'' > 0$, so that the second order conditions are satisfied.

In the case of a quasi-monopolistic cartel, the demand for the cartel's output is, by definition, equal to the difference between total world demand for the commodity, D, and the supply of independent producers, S, that is, for any given price p,

$$D_c(p) = D(p) - S(p), \tag{24.22}$$

so that

$$\frac{\mathrm{d}D_c}{\mathrm{d}p} = \frac{\mathrm{d}D}{\mathrm{d}p} - \frac{\mathrm{d}S}{\mathrm{d}p}. \tag{24.23}$$

By simple manipulations, we get

$$-\frac{D_c}{p}\left(-\frac{p}{D_c}\frac{\mathrm{d}D_c}{\mathrm{d}p}\right) = -\frac{D}{p}\left(-\frac{p}{D}\frac{\mathrm{d}D}{\mathrm{d}p}\right) - \frac{S}{p}\left(\frac{p}{S}\frac{\mathrm{d}S}{\mathrm{d}p}\right). \tag{24.24}$$

We now define the various elasticities

$$\eta_c \equiv -\frac{p}{D_c}\frac{\mathrm{d}D_c}{\mathrm{d}p}, \quad \eta_w \equiv -\frac{p}{D}\frac{\mathrm{d}D}{\mathrm{d}p}, \quad \eta_s \equiv \frac{p}{S}\frac{\mathrm{d}S}{\mathrm{d}p}, \tag{24.25}$$

so that (24.24) becomes

$$-\frac{D_c}{p}\eta_c = -\frac{D}{p}\eta_w - \frac{S}{p}\eta_s, \tag{24.26}$$

whence

$$\eta_c = \frac{D\eta_w + S\eta_s}{D_c} = \frac{\eta_w + (S/D)\,\eta_s}{D_c/D}. \tag{24.27}$$

The fraction D_c/D is the cartel's share in total world consumption, that we denote by k; given Eq. (24.22) we have $S/D = 1-k$. Therefore the final formula is

$$\eta_c = \frac{\eta_w + (1-k)\,\eta_s}{k}. \tag{24.28}$$

24.3 The Effective Rate of Protection

In the general case of n intermediate goods, the pre-tariff value added in sector j is

$$v_j = p_j - \sum_{i=1}^{n} p_i q_{ij} = p_j \left(1 - \sum_{i=1}^{n} a_{ij}\right), \tag{24.29}$$

where $a_{ij} = p_i q_{ij}/p_j$.

After the introduction of a tariff schedule we have

$$v_j' = \left(1 + d_j\right) p_j - \sum_{i=1}^{n} (1 + d_i)\, p_i q_{ij} = p_j \left[\left(1 + d_j\right) - \sum_{i=1}^{n} (1 + d_i)\, a_{ij}\right], \tag{24.30}$$

so that the effective rate of protection turns out to be

$$g_j = \frac{v_j' - v_j}{v_j} = \frac{d_j - \sum\limits_{i=1}^{n} a_{ij} d_i}{1 - \sum\limits_{i=1}^{n} a_{ij}} = d_j + \frac{\left(d_j - \overline{d}_i \sum\limits_{i=1}^{n} a_{ij}\right)}{1 - \sum\limits_{i=1}^{n} a_{ij}}, \tag{24.31}$$

where $\overline{d}_i = \sum\limits_{i=1}^{n} a_{ij} d_i / \sum\limits_{i=1}^{n} a_{ij}$ is a weighted average of the nominal tariff rates. It immediately follows from (24.31) that the same conclusions reached in the text in the case of a single intermediate good hold in the general case as well, if we consider the average rate $\overline{d}_i$ in the place of d_i.

This analysis, it should be noted, is based on the simplifying assumptions of fixed input coefficients of intermediate goods which are all traded. For a more general analysis which relaxes these assumptions see, for example, Various Authors (1973), Yabuuchi and Tanaka (1981, and references therein).

A second observation concerns the definition itself of effective rate of protection. The one used in the text and here is that originally suggested by Corden (1966), who subsequently (Corden, 1969) suggested an alternative definition, namely the proportionate change (due to the tariff structure) in the "price of value added". In general the two definitions give different results, but in the case of separable production functions with fixed input coefficients of intermediate goods they coincide (see, for example, Bhagwati and Srinivasan in Various Authors, 1973).

24.4 Imperfect Competition and Trade Policy

24.4.1 A Tariff Under Vertical Product Differentiation

We consider the effects on the returns to capital and on the range of intra-industry trade of a tariff imposed by country 1 in the context of the model examined in Sect. 23.1.

As regards the returns to capital, the increase in the tariff-inclusive prices of the qualities imported by country 1 will give rise to a range of qualities that country 1 can now produce at a lower cost than the cum-tariff import price instead of importing them as before. Country 1's consumers switch from imports to domestic production of these qualities, hence the demand for domestic capital grows and that for foreign capital decreases. The impact effect is a tendency for the domestic return to capital to increase and for the foreign return to capital to decrease, but the final effect is less clear-cut. Formally, if we introduce the tariff rate d in the demands for capital D_{1K}, D_{2K} as a shift parameter and differentiate the excess demand functions totally, we get

$$E_{R_1}^1 \mathrm{d}R_1 + E_{R_2}^1 \mathrm{d}R_2 + E_d^1 \mathrm{d}d = 0, \tag{24.32}$$

$$E_{R_1}^2 \mathrm{d}R_1 + E_{R_2}^2 \mathrm{d}R_2 + E_d^2 \mathrm{d}d = 0, \tag{24.33}$$

where $E_d^1 > 0$, $E_d^2 < 0$ according to the impact effect. Also note that $E_d^1 + E_d^2 < 0$, because at the world level there is a net decrease in the demand for capital since overall prices are higher. Hence by solving this system for $\mathrm{d}R_1$, $\mathrm{d}R_2$ we obtain the final effect

$$\mathrm{d}R_1 = -\frac{E_d^1 E_{R_2}^2 - E_d^2 E_{R_2}^1}{\Delta} \mathrm{d}d, \tag{24.34}$$

$$\mathrm{d}R_2 = -\frac{E_d^2 E_{R_1}^1 - E_d^1 E_{R_1}^2}{\Delta} \mathrm{d}d. \tag{24.35}$$

Given our assumptions we have $\left|E_{R_1}^1\right| > \left|E_{R_1}^2\right|$, $\left|E_{R_2}^2\right| > \left|E_{R_2}^1\right|$ and $\left|E_d^1\right| < \left|E_d^2\right|$, so that from (24.35) we have $\mathrm{d}R_2/\mathrm{d}d < 0$; but the sign of $\mathrm{d}R_1/\mathrm{d}d$ remains ambiguous.

24.4.1.1 Tariffs and Intra-industry Trade

Let us now consider the effects on intra-industry trade. We have stated above that the tariff imposed by country 1 will give rise to a range of qualities that country 1 can now produce at a lower cost than the cum-tariff import price instead of importing them as it did in the pre-tariff situation. Country 2 will of course go on producing these qualities for its internal consumption. More precisely, we must now distinguish two marginal qualities, $\alpha_2^d < \alpha_1^d$, with country 2 being the sole producer in the range $(\underline{\alpha}, \alpha_2^d)$, country 1 only producing in the range $(\alpha_1^d, \overline{\alpha})$,and both countries producing (but neither trading) in the range (α_2^d, α_1^d). To determine these marginal qualities we first observe that country 1, account being taken of the tariff, will import a quality α, be indifferent between importing it or producing it domestically, produce it domestically, according as $p_1(\alpha) \gtreqless (1+d)p_2(\alpha)$; similarly country 2 will import a quality α, be indifferent between importing it or producing it domestically, produce it domestically, according as $p_2(\alpha) \gtreqless p_1(\alpha)$. Hence the two marginal qualities are defined by

$$p_1(\alpha_2^d) = (1+d)p_2(\alpha_2^d), \tag{24.36}$$

$$p_1(\alpha_1^d) = p_2(\alpha_1^d). \tag{24.37}$$

If we take account of Eq. (9.1), from (24.36) we get

$$W_1 + \alpha_2^d R_1 = (1+d)(W_2 + \alpha_2^d R_2),$$

whence

$$\alpha_2^d = \frac{W_1 - (1+d)W_2}{(1+d)R_2 - R_1}, \tag{24.38}$$

and from (24.37) we get

$$\alpha_1^d = \frac{W_1 - W_2}{R_2 - R_1}, \tag{24.39}$$

where R_1, R_2 are the cum-tariff rental rates.

It is easy to check that $p_1(\alpha) \gtreqless (1+d)p_2(\alpha)$ according as $\alpha \lesseqgtr \alpha_2^d$, and that $p_2(\alpha) \gtreqless p_1(\alpha)$ according as $\alpha \gtreqless \alpha_1^d$. Hence country 1 will import the qualities lower than α_2^d, and country 2 will import the qualities higher than α_1^d. When $d = 0$, it is clear that $\alpha_2^d = \alpha_1^d = \alpha_0$, and we are back in the initial free trade situation. To complete our demonstration we must show that $\alpha_2^d < \alpha_1^d$. This follows from the fact that the fraction in (24.38) has both a greater denominator and a smaller numerator than the fraction in (24.39). Hence in the range (α_2^d, α_1^d) both countries will produce but neither will trade. It is also easy to see that α_2^d is a decreasing function of d, hence the range of non-traded qualities is an increasing function of the tariff rate.

24.4.2 Monopolistic Competition and Welfare-Improving Tariff

Let us examine commercial policy in the context of the model studied in Sect. 23.2, in particular the effects of the imposition of a tariff. A surprising result (Helpman & Krugman, 1989; Venables, 1987) is that the imposition of a tariff seems to cause a decrease in the consumer price index of differentiated goods in the tariff-imposing country, which will then be unambiguously better off.

To show this, let us assume that country 1 imposes a tariff at the rate d on imports of the differentiated good, but not on imports of the homogeneous good. The domestic price of the imported goods will rise to $(1+d)p$, hence the market-clearing condition for country 2's firm becomes

$$x_2 = (1+\phi)^{1-\sigma}(1+d)^{-\sigma}B_1 p^{-\sigma}P_1^{\sigma-\epsilon} + B_2 p^{-\sigma}P_2^{\sigma-\epsilon}, \tag{24.40}$$

where $x_2 = x$ as before. To ascertain the effects of the imposition of a tariff on the transformed consumer-price indices for the differentiated goods we compute the differentials of Eqs. (23.24) and (24.40) with respect to d. These are

$$\begin{aligned} &B_1 p^{-\sigma}\mathrm{d}(P_1^{\sigma-\epsilon}) + (1+\phi)^{1-\sigma}B_2 p^{-\sigma}\mathrm{d}(P_2^{\sigma-\epsilon}) = 0, \\ &[(1+\phi)^{1-\sigma}(1+d)^{-\sigma}B_1 p^{-\sigma}]\mathrm{d}(P_1^{\sigma-\epsilon}) + B_2 p^{-\sigma}\mathrm{d}(P_2^{\sigma-\epsilon}) \\ &= [(1+\phi)^{1-\sigma}B_1 p^{-\sigma}P_1^{\sigma-\epsilon}]\sigma(1+d)^{-\sigma-1}\mathrm{d}d, \end{aligned} \tag{24.41}$$

from which

$$\begin{aligned} \frac{\mathrm{d}P_1^{\sigma-\epsilon}}{\mathrm{d}d} &= -\{[(1+\phi)^{1-\sigma}B_1 p^{-\sigma}P_1^{\sigma-\epsilon}]\sigma(1+d)^{-\sigma-1}\}(1+\phi)^{1-\sigma}B_2 p^{-\sigma}/\Delta, \\ \frac{\mathrm{d}P_2^{\sigma-\epsilon}}{\mathrm{d}d} &= \{B_1 p^{-\sigma}[(1+\phi)^{1-\sigma}B_1 p^{-\sigma}P_1^{\sigma-\epsilon}]\sigma(1+d)^{-\sigma-1}\}/\Delta, \end{aligned} \tag{24.42}$$

where

$$\begin{aligned} \Delta &\equiv B_1 p^{-\sigma}B_2 p^{-\sigma} - [(1+\phi)^{1-\sigma}(1+d)^{-\sigma}B_1 p^{-\sigma}][(1+\phi)^{1-\sigma}B_2 p^{-\sigma}] \\ &= B_1 B_2 p^{-2\sigma}[1-(1+\phi)^{2(1-\sigma)}(1+d)^{-\sigma}] \end{aligned}$$

is positive, because both $(1+\phi)^{2(1-\sigma)}$ and $(1+d)^{-\sigma}$ are smaller than one, given the definition of σ. The numerator of $\mathrm{d}(P_1^{\sigma-\epsilon})/\mathrm{d}d$ is clearly negative while the numerator of $\mathrm{d}(P_2^{\sigma-\epsilon})/\mathrm{d}d$ is positive. These signs remain valid when the derivatives are evaluated at the pre-tariff point ($d = 0$).

Hence the imposition of a tariff causes a decrease in the (transformed) price index of differentiated goods in the tariff-imposing country and an increase in the other country's index. Is this enough to say (as in Flam & Helpman, 1987; Helpman &

Krugman, 1989; Venables, 1987, chap. 7) that a tariff is beneficial? Not at all. What we have shown is that the transformed price indices vary in the directions indicated. What we need to know is how the price indices themselves vary. This depends on the sign of $\sigma - \epsilon$. If this is positive, then the price indices will vary in the same direction as the transformed indices, and the result of the welfare-improving effect of a tariff holds. But this is no longer true when $\sigma - \epsilon$ is negative: in this case the actual price index of the tariff-imposing country will increase, leading to the standard result of a welfare loss. Hence, as noted by Helpman (1990, chap. 4), all depends on the magnitude of the elasticity of substitution in the consumer's subutility function relative to the magnitude of the price-elasticity of aggregate demand.

Be it as it may, the economic reason behind the result that a tariff may improve welfare is the "home market effect": the protected home market becomes a preferential place where to produce to supply goods also to the foreign market. The gains, when the domestic price index falls, derive from the fact that domestic consumers obtain a greater number of cheaper domestic goods and a smaller number of more expensive foreign goods at an overall cost, as measured by the price index, which is lower than in the pre-tariff situation.

24.4.3 Strategic Trade Policy Under Oligopoly with Homogeneous Good

24.4.3.1 Tariffs

In the context of the model treated in Sect. 23.3.1 there is a particularly convenient way of dealing with tariffs, namely to assume that the tariff is levied in terms of the commodity being exported to the tariff-imposing country. This means that, if a quantity x is being exported from, say, country 2 to country 1, a quantity $(1-d)x$ will actually reach country 1's market, where d is the tariff rate imposed by country 1. Hence $(1-d)$ can be treated exactly like g, the transport-cost parameter. An increase (decrease) in g can now be taken as a decrease (increase) in the tariff rate. We already know that a decrease in g causes a decrease in q_{21}^E and an increase in q_{11}^E. The overall quantity is

$$q_{21}^E + q_{11}^E = \frac{2a - c - c/g}{3b}, \tag{24.43}$$

which clearly varies in the same direction as g. Hence, the size of the market decreases as the tariff rate increases. Given the market's downward-sloping demand curve, the price will increase.

As regards the generalisation to the free entry case, we must carefully distinguish two cases. The first is when the number of firms in each economy is arbitrarily fixed or, more precisely, taken as exogenously given and unchanged by trade. This case is not very interesting; besides, the results are ambiguous. The interesting case arises when the number of firms is endogenously determined. Venables (1985) studied the effects of tariffs in such a case, and proved that the imposition of a tariff unambiguously raises welfare in the tariff-imposing country and reduces welfare in the other country. For details see Venables (1985, sect. 6) and Helpman and Krugman (1989, sect. 7.5).

24.4.3.2 Subsidies

For this purpose we consider the case in which the two firms only produce for export and compete in a third market (Brander & Spencer, 1985) . Let

$$p = p(q), \quad q = q_1 + q_2, \quad p' \equiv \mathrm{d}p/\mathrm{d}q < 0 \tag{24.44}$$

be the third market's demand function, where q_i is the quantity offered by firm i. Without loss of generality we can take country 1 as the home country, that subsidizes the domestic firm.

The domestic firm maximizes profit π_1 given by

$$\pi_1 = q_1 p(q) - c_1(q_1) + sq_1, \tag{24.45}$$

where $c_1(q_1)$ is the cost function and s the per-unit subsidy. Since the firm behaves like a Cournot duopolist, the first- and second-order conditions are

$$\begin{aligned} \frac{\partial \pi_1}{\partial q_1} &\equiv \pi_1' = q_1 p' + p - c_1' + s = 0, \\ \frac{\partial^2 \pi_1}{\partial q_1^2} &\equiv \pi_1'' = 2p' + q_1 p'' - c_1'' < 0. \end{aligned} \tag{24.46}$$

For country 2's firm (that receives no subsidy) we have the profit function and the optimum conditions:

$$\begin{aligned} \pi_2 &= q_2 p(q) - c_2(q_2), \\ \frac{\partial \pi_2}{\partial q_2} &\equiv \pi_2' = q_2 p' + p - c_2', \\ \frac{\partial^2 \pi_2}{\partial q_2^2} &\equiv \pi_2'' = 2p' + q_2 p'' - c_2'' < 0. \end{aligned} \tag{24.47}$$

The first-order conditions, as usual in Cournot models, define the reaction functions implicitly. Brander and Spencer also introduce the additional conditions

$$\frac{\partial^2 \pi_1}{\partial q_1 \partial q_2} \equiv \frac{\partial \pi_1'}{\partial q_2} = p' + q_1 p'' < 0; \frac{\partial^2 \pi_2}{\partial q_2 \partial q_1} \equiv \frac{\partial \pi_2'}{\partial q_1} = p' + q_2 p'' < 0, \tag{24.48}$$

$$\frac{\partial^2 \pi_1}{\partial q_1^2} < \frac{\partial^2 \pi_1}{\partial q_1 \partial q_2}; \quad \frac{\partial^2 \pi_2}{\partial q_2^2} < \frac{\partial^2 \pi_2}{\partial q_2 \partial q_1}. \tag{24.49}$$

Conditions (24.48) state that each firm's marginal revenue declines with an increase in the quantity produced by the other firm. Conditions (24.49) mean that the own effects of output on marginal profit are greater (in absolute value) than the cross effects. Note that conditions (24.49)—given conditions (24.48) and the second-order conditions—are always satisfied if marginal cost is nondecreasing.

Given the second-order conditions, inequalities (24.48) imply that the reaction functions are downward sloping. Consider, for example, the domestic firm, whose reaction function is implicitly given by the first-order optimum condition $\partial\pi_1/\partial q_1 = 0$. By the implicit function theorem we can calculate the slope of the reaction function R_1 as

$$\left(\frac{\mathrm{d}q_1}{\mathrm{d}q_2}\right)_{R_1} = -\frac{\partial^2 \pi_1/\partial q_1 \partial q_2}{\partial^2 \pi_1/\partial q_1^2}, \tag{24.50}$$

which is negative, given $\partial^2\pi_1/\partial q_1^2 < 0$, when $\partial^2\pi_1/\partial q_1 q_2 < 0$. Similarly we obtain the slope of the reaction function R_2

$$\left(\frac{\mathrm{d}q_1}{\mathrm{d}q_2}\right)_{R_2} = -\frac{\partial^2 \pi_2/\partial q_2^2}{\partial^2 \pi_2/\partial q_2 \partial q_1} < 0. \tag{24.51}$$

Together, conditions (24.48) and (24.49) imply

$$D \equiv \frac{\partial^2 \pi_1}{\partial q_1^2}\frac{\partial^2 \pi_2}{\partial q_2^2} - \frac{\partial^2 \pi_1}{\partial q_1 \partial q_2}\frac{\partial^2 \pi_2}{\partial q_2 \partial q_1} > 0. \tag{24.52}$$

Let us now go on to comparative statics. We first calculate the effects of the subsidy on outputs and price. Since, at the equilibrium point, the Jacobian of the system formed by the two first-order conditions is different from zero (this Jacobian is simply D), by the implicit function theorem we can express q_1, q_2 as differentiable functions of the parameter s. Then we can compute the derivatives $\mathrm{d}q_1/\mathrm{d}s$, $\mathrm{d}q_2/\mathrm{d}s$ by differentiating the first order-conditions with respect to s. This gives

$$\frac{\partial^2 \pi_1}{\partial q_1^2}\frac{\mathrm{d}q_1}{\mathrm{d}s} + \frac{\partial^2 \pi_1}{\partial q_1 \partial q_2}\frac{\mathrm{d}q_2}{\mathrm{d}s} = -1, \tag{24.53}$$

$$\frac{\partial^2 \pi_2}{\partial q_2 \partial q_1}\frac{\mathrm{d}q_1}{\mathrm{d}s} + \frac{\partial^2 \pi_2}{\partial q_2^2}\frac{\mathrm{d}q_2}{\mathrm{d}s} = 0, \tag{24.54}$$

whence

$$\frac{\mathrm{d}q_1}{\mathrm{d}s} = -\frac{\partial^2 \pi_2}{\partial q_2^2}/D > 0, \tag{24.55}$$

$$\frac{\mathrm{d}q_2}{\mathrm{d}s} = \frac{\partial^2 \pi_2}{\partial q_2 \partial q_1}/D < 0, \tag{24.56}$$

$$\frac{\mathrm{d}q_1}{\mathrm{d}s} + \frac{\mathrm{d}q_2}{\mathrm{d}s} = \left(\frac{\partial^2 \pi_2}{\partial q_2 \partial q_1} - \frac{\partial^2 \pi_2}{\partial q_2^2}\right)/D > 0, \tag{24.57}$$

where the signs derive from (24.46), (24.48), (24.49), (24.52). This shows that an increase in the subsidy causes a decrease in the foreign firm's output and an increase in the output of the domestic firm, a fairly intuitive result (shown in the text, Fig. 10.12). It also shows that total output $q_1 + q_2$ increases, and hence that price decreases, given the downward-sloping demand function (24.44).

Let us now examine the effects of the subsidy on profits. For the domestic firm we have

$$\frac{\mathrm{d}\pi_1}{\mathrm{d}s} = \frac{\partial \pi_1}{\partial q_1}\frac{\mathrm{d}q_1}{\mathrm{d}s} + \frac{\partial \pi_1}{\partial q_2}\frac{\mathrm{d}q_2}{\mathrm{d}s} + q_1,$$

hence, since $\frac{\partial \pi_1}{\partial q_1} = 0$ by the first-order conditions,

$$\frac{\mathrm{d}\pi_1}{\mathrm{d}s} = \frac{\partial \pi_1}{\partial q_2}\frac{\mathrm{d}q_2}{\mathrm{d}s} + q_1 = q_1 p' \frac{\mathrm{d}q_2}{\mathrm{d}s} + q_1 > 0. \tag{24.58}$$

For the foreign firm we have

$$\frac{\mathrm{d}\pi_2}{\mathrm{d}s} = \frac{\partial \pi_2}{\partial q_1}\frac{\mathrm{d}q_1}{\mathrm{d}s} + \frac{\partial \pi_2}{\partial q_2}\frac{\mathrm{d}q_2}{\mathrm{d}s} = q_2 p' \frac{\mathrm{d}q_1}{\mathrm{d}s} < 0. \tag{24.59}$$

These results show that a subsidy increases domestic profit and lowers foreign profit.

The additional (and less intuitive) effect of the subsidy is to increase domestic surplus (net of the subsidy). Domestic surplus $G(s)$ is defined as the domestic firm's profit (deriving from exports) minus the cost of the subsidy:

$$G(s) = \pi_1 - sq_1 \tag{24.60}$$

hence

$$\begin{aligned} \frac{\mathrm{d}G}{\mathrm{d}s} &= \frac{\mathrm{d}\pi_1}{\mathrm{d}s} - q_1 - s\frac{\mathrm{d}q_1}{\mathrm{d}s} \\ &= q_1 p' \frac{\mathrm{d}q_2}{\mathrm{d}s} - s\frac{\mathrm{d}q_1}{\mathrm{d}s}, \end{aligned} \tag{24.61}$$

where we have used (24.58) to substitute for $\mathrm{d}\pi_1/\mathrm{d}s$. At $s = 0$, $\mathrm{d}G/\mathrm{d}s$ is clearly positive since we have shown above that $\mathrm{d}q_2/\mathrm{d}s < 0$. This shows that a marginal increase in the subsidy (from a zero-subsidy situation) increases domestic welfare.

It can also be shown that the optimal subsidy, namely the subsidy that maximizes domestic surplus, is positive. In fact, setting $\mathrm{d}\, G/\mathrm{d}s = 0$ we get

$$s = q_1 p' \frac{\mathrm{d}q_2}{\mathrm{d}s} / \frac{\mathrm{d}q_1}{\mathrm{d}s} > 0. \tag{24.62}$$

Actually, the optimal domestic subsidy moves the domestic firm from a Cournot equilibrium to a Stackelberg equilibrium with the domestic firm as leader. To show this, let us consider what would, in the absence of the subsidy, be the Stackelberg equilibrium with the domestic firm as leader. Without the subsidy, the domestic firm's profit is

$$\pi_1 = q_1 p(q_1 + q_2) - c_1(q_1).$$

The Stackelberg leader (see, for example, Varian, 1992, chap. 16) chooses its optimal quantity taking into account that the follower will react along its Cournot reaction curve. In other words, firm 1 does not take q_2 as given, but knows that $q_2 = f(q_1)$ along firm 2's reaction curve R_2. Firm 2, the follower, continues to behave like a Cournot duopolist.

Thus firm 1's optimum condition is

$$\begin{aligned}\pi_1' &= q_1 p' + q_1 p' \left(\frac{\mathrm{d}q_2}{\mathrm{d}q_1}\right)_{R_2} + p - c_1' \\ &= q_1 p' - q_1 p' \frac{\partial^2 \pi_2/\partial q_2 \partial q_1}{\partial^2 \pi_2/\partial q_2^2} + p - c_1' \\ &= 0,\end{aligned} \tag{24.63}$$

where we have used (24.51).

If we now consider the first-order optimum condition for firm 1 when it behaves like a Cournot duopolist with subsidy—see Eq. (24.46)—and substitute the optimum subsidy as given by (24.62) we get

$$\begin{aligned}\pi_1' &= q_1 p' + p - c_1' + q_1 p' \frac{\mathrm{d}q_2}{\mathrm{d}s} / \frac{\mathrm{d}q_1}{\mathrm{d}s} \\ &= q_1 p' + p - c_1' - q_1 p' \frac{\partial^2 \pi_2/\partial q_2 \partial q_1}{\partial^2 \pi_2/\partial q_2^2} \\ &= 0,\end{aligned} \tag{24.64}$$

where we have used (24.55) and (24.56) to substitute for $\mathrm{d}q_1/\mathrm{d}\ s$ and $\mathrm{d}q_2/\mathrm{d}s$. Conditions (24.63) and (24.64) are identical, which proves the statement.

For further analysis of strategic trade policy in the context of the Brander-Spencer model see, for example, Brainard and Martimort (1997) and Bandyopadhyay (1997). For the case in which firms behave like Bertrand duopolists (i.e., their decisional variable is price rather than quantity) see Neary (1991).

24.4.4 Strategic Trade Policy Under Oligopoly with Differentiated Good

It is easy to see that in the model treated in Sect. 23.5 free trade is not the first-best policy for the foreign country, which can improve its welfare by imposing a tariff on imports of the differentiated commodity. Let us consider a specific tariff d^*. Given the assumptions, the selling price will have to remain the same in both countries as in the free trade situation. Hence, the consumers in the tariff-imposing country will suffer no loss, and the country will have a gain which coincides with $d^*n_2^*$, the fiscal revenue from the tariff. This of course will happen provided that the producing firm finds that after the tariff the alternative of exporting in addition to serving the domestic market remains more profitable than the alternative of serving the domestic market only.

To analyse this point let us observe that the specific tariff can be considered as an additional cost to the producing firm as regards the part of its output exported. Hence its profit will become

$$[\overline{p} - (\theta_2 - \theta_1) - c]n_1 + [\overline{p} - (\theta_2 - \theta_1) - (c + d^*)]n_2^* - K, \tag{24.65}$$

which has to be compared with the profit of serving only the domestic market, $(\overline{p} - c)n_1 - K$. The domestic firm will be indifferent when these two expressions are equal. Thus the optimal specific tariff d_E^*, that is to say the specific tariff that taxes away from the producing firm all profits in excess of profits it makes by selling only to the domestic market, is easily computed by equating the two expressions, from which

$$d_E^* = [\overline{p} - (\theta_2 - \theta_1) - c]\frac{n_1 + n_2^*}{n_2^*} - (\overline{p} - c)\frac{n_1}{n_2^*}. \tag{24.66}$$

If there was trade in the pre-tariff situation, condition (23.60) above had to be satisfied, hence d_E^* is clearly positive.

References

Bandyopadhyay, S. (1997). Demand elasticities, asymmetry and strategic trade policy.
Brainard, S. L., & Martimort, D. (1997). Strategic trade policy with incompletely informed policymakers.
Brander J. A., & Spencer, B. J. (1985). Export subsidies and international market share rivalry.
Corden, W. M. (1966). The structure of a tariff system and the effective protective rate.
Corden, W. M. (1969). Effective protective rates in the general equilibrium model: A geometric note.
Flam, H., & Helpman, E. (1987). Industrial trade policy under monopolistic competition.
Helpman, E. (1990). *Monopolistic competition in trade theory*.
Helpman, E., & Krugman, P. R. (1989). *Trade policy and market structure.*
Lerner, A. P. (1936). *The symmetry between import and export taxes*.
Metzler, L. A. (1949). *Tariffs, the terms of trade, and the distribution of national income*.
Neary, J. P. (1991). *Export subsidies and price competition*.
Varian, H. R. (1992). *Microeconomic analysis.*
Various Authors. (1973). *The Theory of effective protection in general equilibrium: A symposium.*
Venables, A. J. (1985). Trade and trade policy with imperfect competition: The case of identical products and free entry.
Venables, A. J. (1987). Trade and trade policy with differentiated products: A Chamberlinian-Ricardian model.
Yabuuchi, S., & Tanaka, K. (1981). Non-traded inputs, interindustry flows, resource allocation and the ERP theory.

Chapter 25
Appendix to Chapter 11

25.1 The Optimum Tariff

If we denote by v the social welfare function having the quantities demanded (consumed) of the two commodities as arguments, we have, for country 2,

$$v = v\left(A_2^D, B_2^D\right) = v\left(A_2 + E_{2A}, B_2 + E_{2B}\right), \tag{25.1}$$

as $E_{2A} = A_2^D - A_2$ etc. (see Sect. 19.3). We have to maximize (25.1) under the constraints of country 1's offer curve and of the relations linking the variables of the model of general international equilibrium. Instead of using Lagrange multipliers, it is simpler here to introduce the constraints directly into the maximand. For this purpose, it should be remembered that $A_2 = \psi(B_2)$ through country 2's transformation curve, that $E_{2B} = -E_{1B}$, that $E_{2A} = -E_{1A} = pE_{1B}$ (see in particular Eqs. (19.25) and (19.27)). We thus have to maximize

$$v = \left[\psi(B_2) + pE_{1B}(p), B_2 - E_{1B}(p)\right], \tag{25.2}$$

with respect to its arguments, which are now B_2 and p. We obtain the first-order conditions (for brevity, we ignore the second order ones)

$$\begin{aligned} \frac{\partial v}{\partial B_2} &= v_A\psi' + v_B = 0, \\ \frac{\partial v}{\partial p} &= v_A\left(E_{1B} + E'_{1B}p\right) - v_B E'_{1B} = 0. \end{aligned} \tag{25.3}$$

From the first, we get

$$v_B/v_A = -\psi', \tag{25.4}$$

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_25,

and from the second, with simple manipulations,

$$E_{1B}\left[v_A\left(1+E'_{1B}\frac{p}{E_{1B}}\right)-v_B\frac{1}{p}E'_{1B}\frac{p}{E_{1B}}\right]=0, \tag{25.5}$$

whence, given the definition of ε_1 in (19.37) and rearranging terms,

$$\frac{v_B}{v_A}=p\frac{1+\varepsilon_1}{\varepsilon_1}. \tag{25.6}$$

From (25.4) and (25.6) we obtain

$$-\psi'=p\frac{1+\varepsilon_1}{\varepsilon_1}. \tag{25.7}$$

Since (see Sect. 19.1) in equilibrium the marginal rate of transformation equals country 2's domestic relative price, which in turn equals the terms of trade plus tariff, we have

$$p\,(1+d)=p\frac{1+\varepsilon_1}{\varepsilon_1}, \tag{25.8}$$

whence

$$d=\frac{1}{\varepsilon_1}. \tag{25.9}$$

Equation (25.9) states that the optimum tariff for country 2 equals the reciprocal of the elasticity of *country* 1*'s* supply of exports. By using relation (19.41), we can also write

$$d=\frac{1-e_1}{e_1}=\frac{1}{e_1}-1, \tag{25.10}$$

that is, the optimum tariff for country 2 equals the reciprocal of the elasticity of *country* 1*'s* offer curve reduced by one.

In some treatments (see, for example, Johnson, 1950, p. 58 of the 1958 reprint) one finds the following formula for country 2's optimum tariff:

d = elasticity of country 1's offer curve reduced by one

but this depends on the different definition of the elasticity of an offer curve.

25.2 The Theory of Second Best

A Pareto-optimum can always be considered as the solution of a constrained maximum problem. Following Lipsey and Lancaster (1956) consider the following problem

$$\max F(x_1, x_2, \ldots, x_n),$$

$$\text{subject to } \psi(x_1, x_2, \ldots, x_n) = 0, \tag{25.11}$$

where, for simplicity, the constraint has been written as an equality. The solution, if we assume that it is found at an interior point, will be characterized by the conditions obtained by maximizing the Lagrangian

$$L = F(x_1, x_2, \ldots, x_n) - \lambda\psi(x_1, x_2, \ldots, x_n),$$

where λ is a Lagrange multiplier. The Paretian conditions are given by the n first-order conditions

$$F_i - \lambda\psi_i = 0, \quad i = 1, 2, \ldots, n, \tag{25.12}$$

where the subscript i denotes the partial derivative with respect to the i–th variables. These conditions can also be written as

$$\frac{F_1}{\psi_1} = \frac{F_2}{\psi_2} = \ldots = \frac{F_n}{\psi_n}. \tag{25.13}$$

Let us now assume that an additional constraint (a distortion) prevents the fulfilment of one of these conditions, for example the first one, so that

$$\frac{F_1}{\psi_1} \neq \frac{F_2}{\psi_2},$$

that is

$$\frac{F_1}{\psi_1} = k\frac{F_2}{\psi_2}, \quad k \neq 1, \tag{25.14}$$

whence

$$\frac{F_1}{F_2} = k\frac{\psi_1}{\psi_2}. \tag{25.15}$$

It is not necessary for k to be constant, but, for simplicity, we shall assume it is. The presence of the additional constraint (25.15) requires the reformulation of the optimum problem in the form

$$\max F(x_1, x_2, \ldots, x_n)$$

subject to

$$\begin{aligned} \psi\left(x_1, x_2, \ldots, x_n\right) &= 0, \\ \frac{F_1}{F_2} = k\frac{\psi_1}{\psi_2} \qquad &= 0. \end{aligned} \tag{25.16}$$

If we maximize the Lagrangian

$$L' = F\left(x_1, x_2, \ldots, x_n\right) - \lambda'\psi\left(x_1, x_2, \ldots, x_n\right) - \mu\left(\frac{F_1}{F_2} - k\frac{\psi_1}{\psi_2}\right),$$

we obtain the new optimum conditions

$$F_i - \lambda'\psi - \mu\left(\frac{F_2F_{1i} - F_1F_{2i}}{F_2^2} - k\frac{\psi_2\psi_{1i} - \psi_1\psi_{2i}}{\psi_2^2}\right) = 0, i = 1, 2, \ldots, n. \tag{25.17}$$

We can now ask whether the conditions for the second best optimum, namely (25.17), are the same as those for the first best Pareto optimum for $i = 2, \ldots, n$, that is, whether in a situation in which one of the Pareto-optimum conditions cannot be fulfilled, the second best solution is obtained by fulfilling the remaining Pareto-optimum conditions. By comparing (25.17) with (25.12), we see that the answer is affirmative if, and only if,

(a) $\mu = 0$, or
(b) $\mu \neq 0$, but the expression in parentheses in (25.17) is zero for all i.

Case (a) Must be excluded, as it can be seen from (25.17) that for $i = 1, 2$ this would imply $F_1/\psi_1 = F_2/\psi_2$, which is excluded by (25.14).

We are left with case (b), which cannot be excluded *a priori*, but nothing can be said about the expression under consideration, which in general may be positive, nil, or negative and, besides, may take on different values for different i's. It follows that, in general, the conditions for the second best optimum, given the additional constraint (25.14), will be different from the corresponding conditions for the Pareto-optimum. This implies that, in the presence of such an additional constraint, the application of those of the Paretian conditions which can still be fulfilled will not, in general, bring about the (second) best solution in the assumed circumstances. Naturally we cannot exclude the possibility that in certain cases (for example in the case of separable functions) this application may bring about the second best solution, but it should be stressed that this is not a generally valid prescription.

References

Johnson, H. G. (1950). Optimum welfare and maximum revenue tariffs.
Lipsey, R. G., & Lancaster, K. (1956). The general theory of second best.

Chapter 26
Appendix to Chapter 12

26.1 Lobbies, Political Parties, and Endogenous Determination of Protection

We examine here a model by Brock, Magee, and Young (1989, Appendix to chap. 3), which considers two lobbies and two parties. Lobby 1 is pro-export (i.e., it favours an export subsidy that is to say a negative tariff). Lobby 2 is protectionist, namely in favour of a tariff on imports. Party 1 is pro-export, while party 2 is protectionist. The bulk of the voters are in favour of free trade but are imperfectly informed and behave in a nonstrategic manner.

The parties maximise their probabilities of election and the lobbies maximise the expected incomes of their membership. The income of the protectionist lobby 2 will obviously be higher under the protectionist party 2 than under the pro-export party 1, and vice versa for lobby 1. In what follows we use the same notation as Brock et al. (1989), where the primed values of a variable denote the pro-export lobby 1.

pro-export lobby 1

$$\max_{C^{1'},C^{2'}} R^{'} = (1-p)r^{1'} + pr^{2'} - C^{1'} - C^{2'}. \tag{26.1}$$

In this equation, p is the probability of election of party 2 and $(1 - p)$ the probability of election of party 1. When the pro-export party 1 is elected the revenue of the pro-export lobby 1 is $r^{1'}$, greater than $r^{2'}$, the revenue of lobby 1 when the protectionist party 2 is elected. If we multiply these incomes by the relevant probabilities we obtain the total *expected* revenue of lobby 1. The expected income is obtained deducting the lobby's costs, that for simplicity's sake are assumed to consist only of the campaign contributions to the two parties, $C^{1'}$ and $C^{2'}$. The strategy of the lobby is to maximise expected income by an appropriate choice of the contributions.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_26,

In a similar way we obtain the expected income of lobby 2:
protectionist lobby 2

$$\max_{C^1, C^2} R = (1-p)r^1 + pr^2 - C^1 - C^2, \tag{26.2}$$

where $r^2 > r^1$, since the income of the protectionist lobby is higher when the protectionist party is elected.

Let us now come to the parties, whose strategy is to maximise their probability of election. This depends on the contributions received and on the level of tariffs and export subsidies. Letting $q = 1 - p$ we have
pro-export party 1

$$\max_{s} q = q[\underbrace{(C^{1'} + C^1)}_{+}, \underbrace{(C^{2'} + C^2)}_{-}, \underset{-}{s}, \underset{+}{t}], \tag{26.3}$$

protectionist party 2

$$\max_{t} p = p[\underbrace{(C^{1'} + C^1)}_{-}, \underbrace{(C^{2'} + C^2)}_{+}, \underset{+}{s}, \underset{-}{t}], \tag{26.4}$$

where $s \geq 0$ is the export subsidy (favoured by party 1) and $t \geq 0$ is the tariff (favoured by party 2). The signs under the variables represent the signs of the partial derivatives of q and p with respect to the variables. Obviously, an increase in the contributions received by a party causes an increase in the party's probability of election (since a dollar is a dollar, it is irrelevant which lobby the contribution comes from), and a decrease in the other party's probability. Given the general attitude of the voters in favour of free trade, an increase in the tariff has an unfavourable effect on the probability of election of the protariff party 2 and hence a favourable effect on the other party's probability. Similarly, an increase in the export subsidy (which is also an impediment to free trade) has an unfavourable effect on the probability of election of the prosusbsidy party 1 and hence a favourable effect on the other party's probability. Also note that the contributions to the parties from the lobbies are themselves functions of s and t.

It is a common-sense observation that it would be irrational for a lobby to contribute to the party which is favourable to the other lobby. This can easily be proved by observing that the derivative of a lobby's income with respect to the contribution given to the party which is favourable to the other lobby is always negative. Consider for example $\mathrm{d}R'/\mathrm{d}C^{2'} = \mathrm{d}p/\mathrm{d}C^{2'}(r^{2'} - r^{1'}) - 1$. Since $\mathrm{d}p/\mathrm{d}C^{2'} > 0$, and $r^{2'} < r^{1'}$, it follows that $\mathrm{d}R'/\mathrm{d}C^{2'} < 0$. Similarly it can be shown that $\mathrm{d}R/\mathrm{d}C^1 < 0$. Thus we know that $C^{2'} = C^1 = 0$.

Given this result, that Brock and Magee call the "campaign-contribution specialization theorem", the model can be simplified by eliminating $C^{2'}$ and C^1.

The first-order conditions for a maximum yield the following four equations (that form the basis of the Brock and Magee analysis):

$$\begin{aligned}
\mathrm{d}R'/\mathrm{d}C^{1'} &= \mathrm{d}p/\mathrm{d}C^{1'}(r^{2'} - r^{1'}) - 1 = 0,\\
\mathrm{d}R/\mathrm{d}C^{2} &= \mathrm{d}p/\mathrm{d}C^{2}(r^{2} - r^{1}) - 1 = 0,\\
\mathrm{d}q/\mathrm{d}C^{1'}(\mathrm{d}C^{1'}/\mathrm{d}s) + \mathrm{d}q/\mathrm{d}s &= 0,\\
\mathrm{d}p/\mathrm{d}C^{2}(\mathrm{d}C^{2}/\mathrm{d}t) + \mathrm{d}p/\mathrm{d}t &= 0,
\end{aligned} \tag{26.5}$$

for the determination of the equilibrium values of $C^{1'}, C^2, s, t$.

Thus the tariff and the export subsidy are endogenously determined together with the lobbies' contributions to the parties. By substituting back into the model we can then determine the two lobbies' expected incomes as well as the two parties' probabilities of election. Numerous alternative mathematical models are contained in Brock et al. (1989).

26.2 Dumping

Let us consider *persistent dumping*, based on the theory of the discriminating monopolist. Let q_j be the quantity sold on the j–th market and $R_j = p_j q_j$ the corresponding revenue, where p_j is linked to q_j through the j–th market's demand curve. As we assume that all production is carried out in a single plant, total cost $C(q)$ is a function of the overall quantity produced to serve all markets, $q = \sum_{j=1}^{m} q_j$. We must now maximize the profit function

$$\pi = \sum_{j=1}^{m} R_j\left(q_j\right) - C\left(\sum_{j=1}^{m} q_j\right). \tag{26.6}$$

If we assume that there are only two markets, the domestic and the foreign, we get the first-order conditions

$$\begin{aligned}
R_1'\left(q_1\right) &= C'\left(q_1 + q_2\right),\\
R_2'\left(q_2\right) &= C'\left(q_1 + q_2\right),
\end{aligned} \tag{26.7}$$

whence

$$R_1'\left(q_1\right) = R_2'\left(q_2\right) = C'\left(q_1 + q_2\right), \tag{26.8}$$

that is, the marginal revenue in each market must equal the marginal cost of the output as a whole.

The second-order conditions require the leading principal minors of the Hessian

$$\begin{bmatrix} R_1'' - C'' & -C'' \\ -C'' & R_2'' - C'' \end{bmatrix}$$

to alternate in sign, beginning with minus. In the normal case, $R_j'' < 0$ and $C'' > 0$, and so these conditions are satisfied.

The layman's concept of dumping, i.e. a sale below cost in foreign markets (such as sporadic dumping), is formally modelled for example by Davies and McGuinness (1982) and Bernhardt (1984). For the case in which each firm dumps into other firms' home markets due to oligopolistic rivalry, see Brander and Krugman (1983).

References

Bernhardt, D. (1984). Dumping costs and uncertainty.

Brander, J. A., & Krugman, P. R. (1983). A "reciprocal dumping" model of international trade.

Brock, W. A., Magee, S. P., & Young, L. (1989). *Black hole tariffs and endogenous policy theory: Political economy in general equilibrium*.

Davies, S. W., & McGuinness, A. J. (1982). Dumping at less than marginal cost.

Chapter 27
Appendix to Chapter 13

27.1 Classification of the Effects of Growth

Let us first consider the consumption effects, which amounts to finding the conditions for the ratio A^D/Y to increase, remain unchanged, decrease, where A is the importable. The derivative of this ratio is

$$\frac{\mathrm{d}\left(A^D/Y\right)}{\mathrm{d}Y} = \frac{Y\left(\mathrm{d}A^D/\mathrm{d}Y\right) - A^D}{Y^2} = \frac{1}{Y}\left(\frac{\mathrm{d}A^D}{\mathrm{d}Y} - \frac{A^D}{Y}\right), \tag{27.1}$$

which can be written as

$$\frac{\mathrm{d}\left(A^D/Y\right)}{\mathrm{d}Y} = \frac{A^D}{Y^2}\left(\eta_{dY} - 1\right), \tag{27.2}$$

where

$$\eta_{dY} \equiv \frac{\mathrm{d}A^D/A^D}{\mathrm{d}Y/Y} = \frac{\mathrm{d}A^D/\mathrm{d}Y}{A^D/Y} = \frac{\mu_{dY}}{\alpha_{dY}} \tag{27.3}$$

is the income elasticity of A^D, and μ_{dY}, α_{dY}, are, respectively, the marginal and average propensity to consume commodity A.

We thus have the following conditions for the consumption effects of growth to be pro-trade-biased (P), neutral (N), anti-trade-biased (A):

$$\frac{\mathrm{d}\left(A^D/Y\right)}{\mathrm{d}Y} \gtreqless 0 \text{ according as } \eta_{dY} \gtreqless 1 \text{ i.e. according as } \mu_{dY} \gtreqless \alpha_{dY}. \tag{27.4}$$

Growth has ultra-pro-trade biased (UP) or ultra-anti-trade biased (UA) consumption effects when $\mu_{dY} > 1$ or $\mu_{dY} < 0$, respectively.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_27,

As regards the production effects, these involve the derivative of the ratio A^S/Y, which is

$$\frac{\mathrm{d}\left(A^S/Y\right)}{\mathrm{d}Y} = \frac{1}{Y}\left(\frac{\mathrm{d}A^S}{\mathrm{d}Y} - \frac{A^S}{Y}\right) = \frac{A^S}{Y}\left(\eta_{sY} - 1\right), \tag{27.5}$$

where

$$\eta_{sY} \equiv \frac{\mathrm{d}A^S/A^S}{\mathrm{d}Y/Y} = \frac{\mu_{sY}}{\alpha_{sY}}. \tag{27.6}$$

Thus the conditions for the production effects of growth to be pro-trade-biased (P), neutral (N), anti-trade-biased (A) are the following:

$$\frac{\mathrm{d}\left(A^S/Y\right)}{\mathrm{d}Y} \lesseqgtr 0 \text{ according as } \eta_{sY} \lesseqgtr 1 \text{ i.e. according as } \mu_{sY} \lesseqgtr \alpha_{sY}. \tag{27.7}$$

Ultra-pro-trade-biased (*UP*) or ultra-anti-trade-biased (*UA*) production effects occur when $\mu_{sY} < 0$ or $\mu_{sY} > 1$, respectively. Table 27.1 lists the (intervals of) values for the parameters corresponding to the various cases.

Let us now consider the ratio of the *demand for imports* to income, $(A^D - A^S)/Y$, and calculate its derivative. We have

$$\begin{aligned}\frac{\mathrm{d}\left[\left(A^D - A^S\right)/Y\right]}{\mathrm{d}Y} &= \frac{\mathrm{d}\left(A^D/Y\right)}{\mathrm{d}Y} - \frac{\mathrm{d}\left(A^S/Y\right)}{\mathrm{d}Y} \\ &= \frac{1}{Y}\left[\frac{A^D}{Y}\left(\eta_{dY} - 1\right) - \frac{A^S}{Y}\left(\eta_{sY} - 1\right)\right] \\ &= \frac{1}{Y}\left[\left(\mu_{dY} - \alpha_{dY}\right) - \left(\mu_{sY} - \alpha_{sY}\right)\right],\end{aligned} \tag{27.8}$$

where we have used (27.1), (27.2), and (27.5). The definition of the *total* effects of growth, states that growth has pro-trade-biased, neutral, anti-trade-biased total effects according as

$$\frac{\mathrm{d}\left[\left(A^D - A^S\right)/Y\right]}{\mathrm{d}Y} \gtreqless 0,$$

and that it has ultra-pro-trade-biased or ultra-anti-trade-biased total effects according as

$$\frac{\mathrm{d}\left(A^D - A^S\right)}{\mathrm{d}Y} > 1 \text{ or } < 0, \text{ respectively.}$$

Given this definition, Eq. (27.8), and Table (27.1), it is a simple exercise to derive the results listed in Table 13.1 in the text.

Table 27.1 Consumption and production effects of growth

Effects	Consumption parameters	Production parameters
UP	$\eta_{dY} > 1, \alpha_{dY} < 1 < \mu_{dY}$	$\eta_{sY} < 0, \mu_{sY} < 0 < \alpha_{sY} < 1$
P	$\eta_{dY} > 1, 0 < \alpha_{dY} < \mu_{dY} < 1$	$0 < \eta_{sY} < 1, 0 < \mu_{sY} < \alpha_{sY} < 1$
N	$\eta_{dY} = 1, 0 < \mu_{dY} = \alpha_{dY} < 1$	$\eta_{sY} = 1, 0 < \mu_{sY} = \alpha_{sY} < 1$
A	$0 < \eta_{dY} < 1, 0 < \mu_{dY} < \alpha_{dY} < 1$	$\eta_{sY} > 1, 0 < \alpha_{sY} < \mu_{sY} < 1$
UA	$\eta_{dY} < 0, \mu_{dY} < 0 < \alpha_{dY} < 1$	$\eta_{sY} > 1, \alpha_{sY} < 1 < \mu_{sY}$

We now prove the relation between the various growth rates. If we denote the growth rate of import demand by g_m, the new demand for imports will be $(1 + g_m)(A^D - A^S)$. With a similar notation we can write the new domestic demand for and domestic supply of commodity A (the importable) as $(1 + g_m)A^D$ and $(1 + g_s)A^S$, respectively. Then, by definition,

$$(1 + g_m)\left(A^D - A^S\right) = (1 + g_d)\,A^D - (1 + g_s)A^S, \tag{27.9}$$

whence

$$g_m = \frac{g_d A^D - g_s A^S}{A^D - A^S} = \frac{A^D}{A^D - A^S} g_d - \frac{A^S}{A^D - A^S} g_s. \tag{27.10}$$

By suitably adding and subtracting g_Y we have

$$\begin{aligned} g_m &= \frac{(g_d - g_Y + g_Y)\,A^D - (g_s - g_Y + g_Y)\,A^S}{A^D - A^S} \\ &= \frac{g_Y\left(A^D - A^S\right) + (g_d - g_Y)\,A^D - (g_s - g_Y)\,A^S}{A^D - A^S} \\ &= g_Y + \frac{A^D}{A^D - A^S}\,(g_d - g_Y) - \frac{A^S}{A^D - A^S}\,(g_s - g_Y)\,, \end{aligned} \tag{27.11}$$

which is the expression given in Eq. (13.4).

27.2 Comparative Statics of the Effects of Growth in General

Economic growth involves an upward shift of the transformation curve, that is, an increase in the production possibilities of both A and B for any given relative price $p = p_B/p_A$. Since we are not concerned here with the causes of growth, to examine its effects in general it is sufficient to introduce a shift parameter (γ) in the functions defining the quantities of A and B as a function of p along the transformation curve of the country, which we will assume to be country 1 (see Sect. 19.2)[1]

$$A_1 = A_1(p, \gamma), \quad B_1 = (p, \gamma), \tag{27.12}$$

where $\partial A_1/\partial \gamma > 0$ and $\partial B_1/\partial \gamma > 0$ are for the time being considered as exogenously given. To examine the effects of growth (increase in γ) on the terms of trade we start from the international equilibrium condition—see Eq. (19.28)—which becomes

$$E_{1A}(p, \gamma) - pE_{2B}(p) = 0. \tag{27.13}$$

If we differentiate totally with respect to γ, account being taken of the fact that, by the implicit-function theorem, p is a function of γ, we get

$$\frac{\partial E_{1A}}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}\gamma} + \frac{\partial E_{1A}}{\partial \gamma} - E_{2B}\frac{\mathrm{d}p}{\mathrm{d}\gamma} - p\frac{\mathrm{d}E_{2B}}{\mathrm{d}p}\frac{\mathrm{d}p}{\mathrm{d}\gamma} = 0, \tag{27.14}$$

whence

$$\left(E_{2B} + p\frac{\mathrm{d}E_{2B}}{\mathrm{d}p} - \frac{\partial E_{1A}}{\partial p}\right)\frac{\mathrm{d}p}{\mathrm{d}\gamma} = \frac{\partial E_{1A}}{\partial \gamma}. \tag{27.15}$$

We now divide through by E_{2B} (which equals E_{1A}/p, from (27.13) above) and obtain

$$\left(1 + \frac{p}{E_{2B}}\frac{\mathrm{d}E_{2B}}{\mathrm{d}p} - \frac{p}{E_{1A}}\frac{\partial E_{1A}}{\partial p}\right)\frac{\mathrm{d}p}{\mathrm{d}\gamma} = \frac{1}{E_{2B}}\frac{\partial E_{1A}}{\partial \gamma}. \tag{27.16}$$

If we solve for $\mathrm{d}p/\mathrm{d}\gamma$ and use the definitions of the elasticities given in (19.36) and (19.44), we obtain

$$\frac{\mathrm{d}p}{\mathrm{d}\gamma} = \frac{\partial E_{1A}/\partial \gamma}{E_{2B}(1 + \xi_1 + \xi_2)}, \tag{27.17}$$

where ξ_1 and ξ_2 are the elasticities of the demand for imports of the two countries. To determine the sign of $\mathrm{d}p/\,\mathrm{d}\gamma$ we must determine the sign of the fraction on the right-hand side of (27.17). Let us begin by observing that, thanks to Samuelson' s correspondence principle (see, for example, Gandolfo, 2009, chap. 20), it is possible

[1]For simplicity and in accordance with the notation used in the Chap. 19, we henceforth omit the superscript S, so that A and B with no superscript will indicate the quantities supplied (produced), whilst we maintain the superscript D to denote the quantities demanded. It is as well to inform the reader that in what follows, we shall make ample use of the model explained in Chap. 19.

to determine the sign of the denominator: in fact, if the equilibrium is stable, the stability condition (19.48) tells us that $E_{2B}\,(1+\xi_1+\xi_2)<0$.[2]

All that remains is to determine the sign of the numerator. We recall from (19.25) that

$$\begin{aligned} E_{1A}\,(p,\gamma) &= A_1^D\,(I_{1A},p) - A_1\,(p,\gamma)\,, \\ I_{1A} &= A_1 + pB_1, \end{aligned} \tag{27.18}$$

where I_{1A} is country l's real income measured in terms of commodity A. Let us note, incidentally, that demand is ultimately a function of γ for any given p, because I_{1A} is a function of γ through the quantities produced. If we differentiate the first equation in (27.18) with respect to γ we get

$$\frac{\partial E_{1A}}{\partial\gamma} = \frac{\partial A_1^D}{\partial I_{1A}}\frac{\partial I_{1A}}{\partial\gamma} - \frac{\partial A_1}{\partial\gamma}, \tag{27.19}$$

from which

$$\frac{\partial E_{1A}}{\partial\gamma} = \frac{\partial I_{1A}}{\partial\gamma}\,(\mu_{dY}-\mu_{sY})\,, \tag{27.20}$$

where $\mu_{dY} \equiv \partial A_1^D/\partial I_{1A}$ and $\mu_{sY} \equiv (\partial A_1/\partial\gamma)\,/\,(\partial I_{1A}/\partial\gamma) = \partial A_1/\partial I_{1A}$ are, respectively, country 1's marginal propensity to consume and marginal propensity to produce commodity A, already met in the previous section.

As $(\partial I_{1A}/\partial\gamma) = \partial A_1/\partial\gamma + p\,(\partial B_1/\partial\gamma)$ is assumed positive, the sign of the numerator will depend on the sign of $(\mu_{dY}-\mu_{sY})$; as the sign of the fraction is the opposite of the sign of the numerator, we finally have

$$\frac{\mathrm{d}p}{\mathrm{d}\gamma} \gtreqless 0 \text{ according as } \mu_{dY} \lesseqgtr \mu_{sY}\,. \tag{27.21}$$

Relation (27.21), together with Tables 27.1 and 13.1, enables us to immediately obtain the result explained graphically in Sect. 13.3, that is, $\mathrm{d}p/\mathrm{d}\gamma<0$ except in the case of globally *UA* growth. For example, growth with *UA* consumption effects $(\mu_{dY}<0)$ and P production effects $(0<\mu_{sY}<\alpha_{sY})$, which has a *UA* total effect, implies $\mu_{dY}<\mu_{sY}$ and so $\mathrm{d}p/\mathrm{d}\gamma>0$. As another example, consider growth with N consumption effects $(\mu_{dY}=\alpha_{dY})$ and P production effects $(0<\mu_{sY}<\alpha_{sY})$, which has a P total effect. Since A is the importable, the average propensity

[2] It should be recalled that this condition was derived in Chap. 19 on the basis of the assumption that $E_{2B}>0$, i.e. that B is country 2's importable (and, therefore, that A is country 1's importable). But the result is unchanged if we assumed the opposite pattern of trade ($E_{2B}<0$, i.e. B is country 2's exportable, etc.). In this case, in fact, the expression under consideration would become $E_{2B}\,(1+\varepsilon_1+\varepsilon_2)$; given (19.45) and account being taken that $E_{2B}<0$, this expression is negative if the expression in the text is negative.

to consume is higher than the average propensity to produce, that is, $\alpha_{dY} > \alpha_{sY}$ and so, in our case, $\mu_{dY} > \mu_{sY}$, whence $\mathrm{d}p/\mathrm{d}\gamma < 0$.

We leave the other cases as an exercise and pass on to the problem of *immiserizing growth*.

27.2.1 Immiserizing Growth

To avoid the problems inherent in the use of social indifference curves we shall use an alternative way of measuring the improvement or impairment in social welfare, that is, we shall consider the situation as better (worse) if the new national income due to growth enables the society, account being taken of the change in prices, to purchase the *same* bundle of commodities as before *plus* (*minus*) something else. In other words, the situation is better (worse) if the new national income is higher (lower) than the cost (*at the new prices*) of the same bundle of commodities purchased before growth or, equivalently, if the increase in national income is higher (lower) than the increase in the cost of the p re-growth bundle of commodities, where both income and cost are measured in terms of one of the commodities (for example A) taken as numéraire.

Let us begin by calculating the increase in income, which is $\mathrm{d}I_{1A}/\mathrm{d}\gamma$.

It must be stressed that this is a *total* derivative, which takes all the effects of γ on output and p into account, and not the partial derivative previously used. Thus we have

$$\frac{\mathrm{d}I_{1A}}{\mathrm{d}\gamma} = \frac{\partial I_{1A}}{\partial \gamma} + \frac{\partial I_{1A}}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}\gamma} = \frac{\partial I_{1A}}{\partial \gamma} + B_1\frac{\mathrm{d}p}{\mathrm{d}\gamma}, \tag{27.22}$$

as $\partial I_{1A}/\partial p = B_1$ by (19.22). The pre-growth bundle of goods has a cost, at the pre-growth relative price, of $A_1^D + pB_1^D$; by keeping A_1^D and B_1^D unchanged and letting p vary we get the change in cost, $B_1^D(\mathrm{d}p/\mathrm{d}\gamma)$, so that the post-growth situation will be better or worse according as

$$\frac{\mathrm{d}I_{1A}}{\mathrm{d}\gamma} = \frac{\partial I_{1A}}{\partial \gamma} + B_1\frac{\mathrm{d}p}{\mathrm{d}\gamma} \gtrless B_1^D\frac{\mathrm{d}p}{\mathrm{d}\gamma}, \tag{27.23}$$

that is,

$$\frac{\partial I_{1A}}{\partial \gamma} - \left(B_1^D - B_1\right)\frac{\mathrm{d}p}{\mathrm{d}\gamma} = \frac{\partial I_{1A}}{\partial \gamma} - E_{1B}\frac{\mathrm{d}p}{\mathrm{d}\gamma} \gtrless 0. \tag{27.24}$$

If we use the value of $\mathrm{d}p/\mathrm{d}\gamma$ found above—see Eq. (27.17)—and recall that, by Eqs. (19.27), $-E_{1B} = E_{2B}$ in equilibrium, we get

$$\frac{\partial I_{1A}}{\partial \gamma} + \frac{\partial E_{1A}/\partial \gamma}{1+\xi_1+\xi_2} \gtrless 0, \tag{27.25}$$

whence, by substituting the value of $\partial E_{1A}/\partial\gamma$ from (27.20) and collecting terms, we arrive at

$$\frac{\partial I_{1A}}{\partial\gamma}\left(1+\frac{\mu_{dY}-\mu_{sY}}{1+\xi_1+\xi_2}\right)\gtrless 0. \tag{27.26}$$

As $\partial I_{1A}/\partial\gamma$ is assumed positive, the condition for a *worsening* (immiserizing growth) is

$$1+\frac{\mu_{dY}-\mu_{sY}}{1+\xi_1+\xi_2}<0. \tag{27.27}$$

Since we have assumed that the equilibrium is stable, that is $1+\xi_1+\xi_2<0$ from (19.49), we can rewrite (27.27) as

$$1+\xi_1+\xi_2+\mu_{dY}-\mu_{sY}>0. \tag{27.28}$$

This condition may, in general, be either realized or not, so that immiserizing growth remains a possibility to be further investigated by an examination of the causes of growth (see Sects. 27.3 and 27.4); this examination will also enable us to obtain exact expressions for the various derivatives $\partial A_1/\partial\gamma$ etc. and in particular for μ_{sY}. However, it is now possible to show the *necessary condition* for immiserizing growth, which is that the terms of trade move against the growing country. In fact, if these were to improve or to remain unchanged, given Eq. (27.21) we have $\mu_{dY}-\mu_{sY}<0$ and so, as $1+\xi_1+\xi_2<0$, condition (27.28) is *not* satisfied. This shows that the deterioration in the terms of trade is necessary, though not sufficient, for growth to be immiserizing.

27.3 Changes in Factor Endowments, Rybczynski's Theorem, and the Terms of Trade

Let us assume that the total amount of labour existing in the economy, L, increases. Given the assumption of continuous full employment of all factors of production, the expansion factor γ is identified with L. Thus we have to calculate $\partial I_{1A}/\partial L$ and $\partial A_1/\partial L$ to determine μ_{sY} and then $\mathrm{d}p/\mathrm{d}L$, as made clear in (27.20) and (27.21). In what follows we shall drop the subscript 1 for brevity of notation, as it is understood that all magnitudes refer to country 1.

If we consider the relations, derived from Eqs. (19.17)

$$\begin{aligned} L_A+L_B&=L,\\ \varrho_A L_A+\varrho_B L_B&=K, \end{aligned} \tag{27.29}$$

we can express L_A and L_B in terms of factor endowments and factor intensities, thus obtaining

$$L_A = \frac{\varrho_B L - K}{\varrho_B - \varrho_A}, \quad L_B = \frac{K - \varrho_A L}{\varrho_B - \varrho_A}. \tag{27.30}$$

The production functions in intensive form—see Eqs. (19.17)—are

$$A = L_A g_A(\varrho_A), \quad B = L_B g_B(\varrho_B), \tag{27.31}$$

where, thanks to the assumption of first-degree homogeneity, ϱ_A and ϱ_B depend solely on the relative price of factors, which is kept constant in computing the partial derivatives $\partial A/\partial L$ and $\partial B/\partial L$.

From (27.31) we therefore obtain

$$\frac{\partial A}{\partial L} = \frac{\partial L_A}{\partial L} g_A, \quad \frac{\partial B}{\partial L} = \frac{\partial L_B}{\partial L} g_B, \tag{27.32}$$

that is, given (27.30),

$$\frac{\partial A}{\partial L} = \frac{\varrho_B}{\varrho_B - \varrho_A} g_A, \quad \frac{\partial B}{\partial L} = \frac{-\varrho_A}{\varrho_B - \varrho_A} g_B, \tag{27.33}$$

so that the two productive levels move in opposite directions. If, for example, one assumes (as in the text) that sector A is labour intensive (so that sector B is capital intensive, whence $\varrho_B > \varrho_A$), it follows that the output of A increases and that of B decreases (the *Rybczynski theorem*). Besides, as $\mathrm{d}A/A = (\partial A/\partial L)\mathrm{d}L/A = [(\partial A/\partial L)/(A/L)](\mathrm{d}L/L)$, and account being taken of Eqs. (27.33), (27.31), and (27.30), we see that

$$\frac{\mathrm{d}A}{A} = \frac{\varrho_B}{\varrho_B - \varrho} \frac{\mathrm{d}L}{L}, \quad \varrho = K/L, \tag{27.34}$$

whence, as we assumed $\varrho_B > \varrho_A$ (and so $\varrho_B > \varrho > \varrho_A$), it follows that $\mathrm{d}A/A > \mathrm{d}L/L$. In other words, the output of the expanding sector (in our example, sector A) increases *more than proportionally* to the increase in the factor. Jones (1965) has called this the *magnification effect*.

Let us now go back to the main line and calculate $\partial I_A/\partial L$. Since $I_A = A + pB$, we have

$$\frac{\partial I_A}{\partial L} = \frac{\partial A}{\partial L} + p\frac{\partial B}{\partial L} = \frac{\varrho_B g_A - p\varrho_A g_B}{\varrho_B - \varrho_A}. \tag{27.35}$$

Now, from Eqs. (19.17) we get

$$pg_B = g_A + p\varrho_B g'_B - \varrho_A g'_A = g_A + p\varrho_B g'_B - \varrho_A p g'_B = g_A + pg'_B(\varrho_B - \varrho_A)$$

and so

$$\frac{\partial I_A}{\partial L} = \frac{\varrho_B g_A - \varrho_A g_A - \varrho_A p g_B' (\varrho_B - \varrho_A)}{\varrho_B - \varrho_A} = g_A - \varrho_A p g_B',$$

whence, as $pg_B' = g_A'$, we get

$$\frac{\partial I_A}{\partial L} = g_A - \varrho_A g_A', \tag{27.36}$$

which is certainly positive by (19.17).

Given (27.36) and (27.33), we can calculate

$$\mu_{sY} = \frac{\partial A/\partial L}{\partial I_A/\partial L} = \frac{\varrho_B g_A}{(\varrho_B - \varrho_A)\left(g_A - \varrho_A g_A'\right)}. \tag{27.37}$$

Since $g_A - \varrho_A g_A' > 0$, the sign of μ_{sY} depends on the sign of $\varrho_B - \varrho_A$ so that

$$\mu_{sY} \gtrless 0 \text{ according as } \varrho_B \gtrless \varrho_A\,, \tag{27.38}$$

that is, according as sector B's capital intensity is higher or lower than sector A's. Besides, it can be shown that, if $\varrho_B > \varrho_A$, then μ_{sY} is not only positive but also greater than one. In fact, we have

$$0 < \varrho_B g_A - \varrho_B \varrho_A g_A' - \varrho_A \left(g_A - \varrho_A g_A'\right) < \varrho_B g_A,$$

where the central expression is the denominator of the fraction in (27.37); the left-hand inequality derives from the assumption $\varrho_B > \varrho_A$, and the right-hand one is self-evident. It follows that the denominator under consideration, when positive, is certainly smaller than the numerator in (27.37), so that $\mu_{sY} > 1$. This is an important result because it enables us to determine the direction in which the terms of trade move and to *exclude the possibility of immiserizing growth when no good is inferior*. If we assume—as in the text—that A is the labour-intensive commodity, then $\varrho_A < \varrho_B$ and so $\mu_{sY} > 1$. Now, if no good is inferior, the marginal propensity to consume A must be smaller than one, that is, $\mu_{dY} < 1$ It follows that $\mu_{dY} < \mu_{sY}$ and so, according to (27.21), the relative price $p = p_B/p_A$ increases:

$$\frac{\mathrm{d}p}{\mathrm{d}L} > 0, \tag{27.39}$$

so that the terms of trade will move in favour of or against the country according as A is the importable or the exportable (*corollary of Rybczynski's theorem*). If we assume that A is the importable, the improvement in the terms of trade *excludes* the possibility of immiserizing growth, as the necessary condition (27.28) is *not* verified.

Immiserizing growth, therefore, requires as a necessary condition that the importable should *not* be intensive in the augmenting factor: only when $\varrho_A > \varrho_B$, in fact, $\mu_{sY} < 0$ and so $\mathrm{d}p/\mathrm{d}L < 0$, which is the necessary condition for immiserizing growth to occur.

So far we have examined the effects of an increase in the labour force; the same procedure can be used to analyse the effects of an increase in the stock of capital (this includes the case of a transfer from abroad). We only state the results, omitting all the intermediate steps, which are exactly like those detailed above in the case of an increase in L. We begin by

$$\frac{\partial A}{\partial K} = \frac{\partial L_A}{\partial K} g_A, \quad \frac{\partial B}{\partial K} = \frac{\partial L_B}{\partial K} g_B, \tag{27.40}$$

so that, given (27.30),

$$\frac{\partial A}{\partial K} = \frac{-g_A}{\varrho_B - \varrho_A}, \quad \frac{\partial B}{\partial K} = \frac{g_B}{\varrho_B - \varrho_A}, \tag{27.41}$$

and so $\partial A/\partial K \lesseqgtr 0$ and $\partial B/\partial K \gtreqless 0$ according as $\varrho_B \gtreqless \varrho_A$ (Rybczynski's theorem). We then calculate

$$\frac{\partial I_A}{\partial A} = \frac{\partial A}{\partial K} + p\frac{\partial B}{\partial K} = \frac{-g_A + pg_B}{\varrho_B - \varrho_A}, \tag{27.42}$$

whence, after suitable substitutions from (19.17), we get

$$\frac{\partial I_A}{\partial A} = g'_A > 0, \tag{27.43}$$

and so

$$\mu_{sY} = \frac{\partial A/\partial K}{\partial I_A/\partial K} = \frac{-g_A}{g'_A(\varrho_B - \varrho_A)} \gtreqless 0 \text{ according as } \varrho_B \lesseqgtr \varrho_A, \tag{27.44}$$

a result symmetrical with that obtained in the case of an increase in L, as can be arrived at intuitively.

Also in this case it is possible to show that μ_{sY}, when positive, is necessarily greater than one. From Eq. (19.17) we recall that $g_A - \varrho_A g'_A > 0$; besides, it is obvious that $g'_A \varrho_A > g'_A(\varrho_A - \varrho_B)$, so that $g_A > g'_A \varrho_A > g'_A(\varrho_A - \varrho_B)$. If we assume $\varrho_A - \varrho_B > 0$ (i.e., μ_{sY} positive), we can divide throughout by $g'_A(\varrho_A - \varrho_B)$ and obtain

$$\frac{g_A}{g'_A(\varrho_A - \varrho_B)} > 1, \tag{27.45}$$

as was to be demonstrated.

Also the results concerning the terms of trade are symmetrical with those obtained above in the case of an increase in L. In fact, if no good is inferior, $\mathrm{d}p/\mathrm{d}K \gtrless 0$ according as $\varrho_A \gtrless \varrho_B$, whence the corollary of Rybczynski's theorem and the usual conclusions on immiserizing growth follow.

27.3.1 Simultaneous Increase in Both Factors

We conclude this section by examining the effects of a *simultaneous increase in both factor endowments*. Let us consider the total differential

$$\mathrm{d}A = \frac{\partial A}{\partial L}\mathrm{d}L + \frac{\partial A}{\partial K}\mathrm{d}K, \tag{27.46}$$

whence, substituting from (27.33) and (27.41), and collecting terms

$$\mathrm{d}A = \frac{\varrho_B g_A \mathrm{d}L - g_A \mathrm{d}K}{\varrho_B - \varrho_A} = \frac{g_A \mathrm{d}L\,[\varrho_B - (\mathrm{d}K/\mathrm{d}L)]}{\varrho_B - \varrho_A}. \tag{27.47}$$

Similarly we obtain

$$\mathrm{d}B = \frac{g_B \mathrm{d}L\,[(\mathrm{d}K/\mathrm{d}L) - \varrho_A]}{\varrho_B - \varrho_A}. \tag{27.48}$$

In order to be able to analyse the signs of $\mathrm{d}A$ and $\mathrm{d}B$ we must know the changes in factor endowments. For this purpose we consider their proportional changes and introduce a parameter α such that

$$\frac{\mathrm{d}K}{K} = \alpha\frac{\mathrm{d}L}{L}, \tag{27.49}$$

where $\alpha \gtreqless 1$ according as the capital stock increases more than proportionally to, in the same proportion as, less than proportionally to the increase in the labour force. We then have

$$\frac{\mathrm{d}K}{\mathrm{d}L} = \alpha\varrho, \tag{27.50}$$

where $\varrho \equiv K/L$ is the initial factor endowment ratio which, as we recall from Eq. (4.3), is a weighted average of the factor intensities in the two sectors. By substituting (27.50) into (27.47) and (27.48), we obtain

$$\mathrm{d}A = \frac{g_A \mathrm{d}L\,(\varrho_B - \alpha\varrho)}{\varrho_B - \varrho_A}, \quad \mathrm{d}B = \frac{g_B \mathrm{d}L\,(\alpha\varrho - \varrho_A)}{\varrho_B - \varrho_A}. \tag{27.51}$$

It is possible, when $\alpha = 1$, to reach a definite conclusion, namely that an equiproportional increase in both factor endowments brings about an increase in the output of *both* commodities (conversely, as we have seen above, when only one factor increases, the commodity outputs move in opposite directions). In fact, by definition of average, ϱ is always situated between the minimum and the maximum term, that is

$$\begin{aligned} \varrho_A < \varrho < \varrho_B \text{ if } \varrho_B > \varrho_A, \\ \varrho_B < \varrho < \varrho_A \text{ if } \varrho_A > \varrho_B. \end{aligned} \tag{27.52}$$

Thus, letting $\alpha = 1$ in (27.51)—account being taken of (27.52)—we see that when $\varrho_B > \varrho_A$ we have $\varrho_B > \varrho$ and $\varrho > \varrho_A$, so that $\mathrm{d}A > 0$, $\mathrm{d}B > 0$. And when $\varrho_B < \varrho_A$ we have $\varrho_B < \varrho$ and $\varrho < \varrho_A$ so that $\mathrm{d}A > 0$, $\mathrm{d}B > 0$.

But when $\alpha \neq 1$ it is no longer possible to reach definite conclusions: the output of one commodity will certainly increase (this will be the commodity intensive in the factor with the higher proportional increase) but the other may either increase or decrease. Let us assume, for example, $\alpha > 1$ (the capital stock increases more than proportionally to the labour force) and $\varrho_B > \varrho_A$ (the capital-intensive commodity is B). Since from (27.52) we have $\varrho_A < \varrho$ it will also be true that $\varrho_A < \alpha\varrho$, hence $\mathrm{d}B > 0$. But, although $\varrho < \varrho_B$, $\alpha\varrho$ may in general be smaller than, equal to, or greater than ϱ_B, so that the sign of $\mathrm{d}A$ is indeterminate.

We next examine the effects on the terms of trade. For this purpose we must, first, calculate the variation in real income—at unchanged p—and then determine the marginal propensity to produce A. At unchanged p, we have the total differential

$$\mathrm{d}I_A = \mathrm{d}A + p\mathrm{d}B, \tag{27.53}$$

and substituting into it from (27.51) we get

$$\mathrm{d}I_A = \mathrm{d}L \frac{g_A\left(\varrho_B - \alpha\varrho\right) + pg_B\left(\alpha\varrho - \varrho_A\right)}{\varrho_B - \varrho_A}. \tag{27.54}$$

In the case in which $\alpha = 1$, we certainly have $\mathrm{d}I_A > 0$, whilst the sign of (27.57) is ambiguous when $\alpha \neq 1$. It is however possible to eliminate this ambiguity by rewriting $\mathrm{d}I_A$ in the form

$$\mathrm{d}I_A = \mathrm{d}L\left(g_A - \varrho_A g_A' + \alpha\varrho g_A'\right), \tag{27.55}$$

which can be arrived at either by transforming (27.54) by a suitable use of Eqs. (19.17), or by starting from the equivalent definition of the variation in real income

$$\mathrm{d}I_A = \frac{\partial I_A}{\partial L}\mathrm{d}L + \frac{\partial I_A}{\partial K}\mathrm{d}K = \mathrm{d}L\left(\frac{\partial I_A}{\partial L} + \alpha\varrho\frac{\partial I_A}{\partial K}\right), \tag{27.56}$$

and substituting from (27.36) and (27.43) into it.

Now, since $g_A - \varrho_A g'_A > 0$ and $\alpha\varrho g'_A > 0$, it follows from (27.55) that $\mathrm{d}I_A$ is certainly positive even if $\alpha \neq 1$.

We now calculate, from (27.51) and (27.54),

$$\mu_{sY} = \frac{\partial A}{\partial I_A} = \frac{g_A\left(\varrho_B - \alpha\varrho\right)}{g_A\left(\varrho_B - \alpha\varrho\right) + p g_B\left(\alpha\varrho - \varrho_A\right)}, \tag{27.57}$$

which has a generally indeterminate sign. In the particular case where $\alpha = 1$ it can easily be seen, by using (27.52), that the numerator is smaller in absolute value than the denominator, so that $0 < \mu_{sY} < 1$. If no good is inferior, $0 < \mu_{dY} < 1$ as well, so that—given (27.21)–the terms of trade can move in either direction and the phenomenon of immiserizing growth may appear.

27.4 Technical Progress

A possible way of representing technical progress in a sector, for example in that producing commodity A, is

$$A = f_A\left(\lambda K_A, \lambda' L_A\right), \tag{27.58}$$

where λ and λ' are parameters, initially equal to one, which increase when technical progress occurs (if it occurs continuously, λ and λ' will be continuous functions of time). This is called *factor-augmenting* disembodied technological change. Since f_A is, assumed, a first-degree homogeneous production function, we can write it in the intensive form

$$A = \lambda' L_A g_A\left(\frac{\lambda}{\lambda'}\varrho_A\right). \tag{27.59}$$

We then define technical progress as being neutral, capital-saving, or labour-saving according as λ increases in the same proportion as, more than proportionally or less than proportionally to the increase in λ', that is, according as the ratio λ/λ' remains unchanged, increases or decreases.[3]

Let us then assume, as in the text, that a technological change occurs in sector A, but not in sector B, so that the production function of the latter, $B = L_B g_B(\varrho_B)$, remains the same. The equilibrium conditions contained in the third and fourth equation of set (19.17) become, account being taken of Eq. (27.59),

[3]This definition is equivalent to the Hicks classification employed in the text, if the elasticity of substitution between the factors is smaller than one (see, for example, Vanek, 1966; see also Allen, 1967; Hicks, 1932; Johnson, 1955). To avoid unnecessary complications we assume this to be the case.

$$\begin{aligned} \lambda g_A' - p g_B' &= 0, \\ \lambda' g_A - \lambda \varrho_A g_A' - p\left(g_B - \varrho_B g_B'\right) &= 0. \end{aligned} \tag{27.60}$$

27.4.1 *Effects of Technical Progress on Factor Intensities and Rewards*

Let us begin by determining the changes in the factor intensities in the two sectors ϱ_A, ϱ_B. For this purpose we can consider Eqs. (27.60) as a set of implicit functions, so that, on the basis of the implicit function theorem, we can express ϱ_A and ϱ_B as differentiable functions of the parameters λ and λ' in a neighbourhood of the equilibrium point, provided that the Jacobian of (27.60) with respect to ϱ_A and ϱ_B, evaluated at the equilibrium point, is not zero. This Jacobian is

$$J = \begin{vmatrix} \lambda g_A'' \frac{\lambda}{\lambda'} & -p g_B'' \\ \lambda' g_A' \frac{\lambda}{\lambda'} - \lambda g_A' - \lambda \varrho_A g_A'' \frac{\lambda}{\lambda'} & -p g_B' + p g_B' + p \varrho_B g_B' \end{vmatrix}, \tag{27.61}$$

whence, noting that $\lambda = \lambda' = 1$ in the initial equilibrium point and simplifying,

$$J = -g_A'' g_B'' p \left(\varrho_B - \varrho_A\right), \tag{27.62}$$

which is different from zero since $\varrho_B \neq \varrho_A$. Thus there exist the functions

$$\begin{aligned} \varrho_A &= \varrho_A \left(\lambda, \lambda'\right), \\ \varrho_B &= \varrho_B \left(\lambda, \lambda'\right). \end{aligned} \tag{27.63}$$

If we differentiate (27.60) with respect to λ, account being taken of (27.63), we have

$$\begin{aligned} g_A' + \lambda g_A'' \left(\frac{\varrho_A}{\lambda'} + \frac{\lambda}{\lambda'} \frac{\partial \varrho_A}{\partial \lambda}\right) - p g_B'' \frac{\partial \varrho_B}{\partial \lambda} &= 0, \\ \lambda' g_A' \left(\frac{\varrho_A}{\lambda'} + \frac{\lambda}{\lambda'} \frac{\partial \varrho_A}{\partial \lambda}\right) - \varrho_A g_A' - \lambda \frac{\partial \varrho_A}{\partial \lambda} g_A' - \lambda \varrho_A g_A'' \left(\frac{\varrho_A}{\lambda'} + \frac{\lambda}{\lambda'} \frac{\partial \varrho_A}{\partial \lambda}\right) & \\ - p g_B' \frac{\partial \varrho_B}{\partial \lambda} + p g_B' \frac{\partial \varrho_B}{\partial \lambda} + p \varrho_B g_B'' \frac{\partial \varrho_B}{\partial \lambda} &= 0. \end{aligned} \tag{27.64}$$

Again noting that initially $\lambda = \lambda'$, we can simplify and rearrange terms, thus obtaining

$$\begin{aligned} g_A'' \frac{\partial \varrho_A}{\partial \lambda} - p g_B'' \frac{\partial \varrho_B}{\partial \lambda} &= -g_A' - g_A'' \varrho_A, \\ -\varrho_A g_A'' \frac{\partial \varrho_A}{\partial \lambda} + p \varrho_B g_B'' \frac{\partial \varrho_B}{\partial \lambda} &= g_A'' \varrho_A^2, \end{aligned}$$

which, solved for $\partial \varrho_A / \partial \lambda$ and $\partial \varrho_B / \partial \lambda$, yields

$$\begin{aligned}\frac{\partial \varrho_A}{\partial \lambda} &= \frac{-p\varrho_B g_B'' g_A' - p\varrho_B g_B'' g_A'' \varrho_A + p g_B'' g_A'' \varrho_A^2}{-g_A'' g_B'' p\,(\varrho_B - \varrho_A)} \\ &= -\frac{-p\varrho_A g_A'' g_B''\,(\varrho_B - \varrho_A) - p\varrho_B g_A' g_B''}{-g_A'' g_B'' p\,(\varrho_B - \varrho_A)} \\ &= -\left[\varrho_A + \frac{\varrho_B g_A'}{g_A''\,(\varrho_B - \varrho_A)}\right], \end{aligned} \tag{27.65}$$

and

$$\frac{\partial \varrho_B}{\partial \lambda} = -\frac{\varrho_A g_B'}{p g_B''\,(\varrho_B - \varrho_A)}. \tag{27.66}$$

By the same procedure (differentiate (27.60) with respect to λ', account being taken of (27.63), and solve the resulting system) one gets

$$\frac{\partial \varrho_A}{\partial \lambda'} = \varrho_A - \frac{g_A - \varrho_A g_A'}{g_A''\,(\varrho_B - \varrho_A)}, \tag{27.67}$$

$$\frac{\partial \varrho_B}{\partial \lambda'} = -\frac{g_B - \varrho_B g_B'}{p g_B''\,(\varrho_B - \varrho_A)}. \tag{27.68}$$

In general, both λ and λ' will increase as a consequence of technical progress; since we have assumed $\lambda = \lambda'$ initially, technical progress will be capital-saving-biased, neutral, labour-saving-biased according as

$$\mathrm{d}\lambda \gtreqless \mathrm{d}\lambda', \text{ that is } \mathrm{d}\lambda / \mathrm{d}\lambda' \gtreqless 1 \tag{27.69}$$

If we introduce a parameter β measuring the ratio between the two changes, $\beta = \mathrm{d}\lambda / \mathrm{d}\lambda'$, we can rewrite (27.69) as

$$\beta \gtreqless 1. \tag{27.70}$$

The total effect of technological change on ϱ_A, ϱ_B is obtained by calculating the total differentials of these. As regards ϱ_A we have

$$\mathrm{d}\varrho_A = \frac{\partial \varrho_A}{\partial \lambda}\mathrm{d}\lambda + \frac{\partial \varrho_A}{\partial \lambda'}\mathrm{d}\lambda', \tag{27.71}$$

that is, given the definition of β,

$$\mathrm{d}\varrho_A = \mathrm{d}\lambda' \left(\beta \frac{\partial \varrho_A}{\partial \lambda} + \frac{\partial \varrho_A}{\partial \lambda'}\right). \tag{27.72}$$

Since $d\lambda' > 0$, the sign of $d\varrho_A$ depends solely on the sign of the expression in parentheses; substituting Eqs. (27.65) and (27.67) into it we get

$$\begin{aligned}\beta\frac{\partial\varrho_A}{\partial\lambda}+\frac{\partial\varrho_A}{\partial\lambda'} &= -\beta\varrho_A-\frac{\beta\varrho_B g_A'}{g_A''(\varrho_B-\varrho_A)}+\varrho_A-\frac{g_A-\varrho_A g_A'}{g_A''(\varrho_B-\varrho_A)}\\ &= \varrho_A(1-\beta)-\frac{g_A-\varrho_A g_A'+\beta\varrho_B g_A'}{g_A''(\varrho_B-\varrho_A)}\\ &= \varrho_A(1-\beta)-\frac{p\left(g_B-\varrho_B g_B'\right)+\beta p\varrho_B g_B'}{g_A''(\varrho_B-\varrho_A)},\end{aligned} \tag{27.73}$$

where the last passage has been made possible by the equilibrium conditions (27.60) for $\lambda = \lambda'$.

Similarly, given

$$d\varrho_B = d\lambda'\left(\beta\frac{\partial\varrho_B}{\partial\lambda}+\frac{\partial\varrho_B}{\partial\lambda'}\right), \tag{27.74}$$

it follows that the sign of $d\varrho_B$ depends solely on the sign of the following expression

$$\begin{aligned}\beta\frac{\partial\varrho_B}{\partial\lambda}+\frac{\partial\varrho_B}{\partial\lambda'} &= -\frac{\beta\varrho_A g_B'}{p g_B''(\varrho_B-\varrho_A)}-\frac{g_B-\varrho_B g_B'}{p g_B''(\varrho_B-\varrho_A)}\\ &= -\frac{g_B-\varrho_B g_B'+\beta\varrho_A g_B'}{p g_B''(\varrho_B-\varrho_A)}.\end{aligned} \tag{27.75}$$

Let us note for future reference that the numerator of the last fraction in (27.73) and the numerator of the last fraction in (27.75) are in any case positive. In fact, $g_B - \varrho_B g_B'$ is the marginal productivity of labour in sector B and g_B' is the marginal productivity of capital in the same sector.

We can now consider the various types of technical progress, beginning with the neutral one. Since $\beta = 1$, in Eq. (27.73) only the last fraction remains,which becomes

$$\frac{p g_B}{-g_A''(\varrho_B-\varrho_A)}. \tag{27.76}$$

Since $-g_A'' > 0$ as $g_A'' < 0$ (decreasing marginal productivity), the sign of (27.76) depends exclusively on the sign of $(\varrho_B - \varrho_A)$. Therefore

$$d\varrho_A \gtreqless 0 \text{ according to whether } \varrho_B \gtreqless \varrho_A. \tag{27.77}$$

If we now examine (27.75) we can see that, as the numerator is positive and $g_B'' < 0$, its sign depends exclusively on $(\varrho_B - \varrho_A)$. Therefore

$$d\varrho_B \gtreqless 0 \text{ according to whether } \varrho_B \gtreqless \varrho_A. \tag{27.78}$$

It follows that ϱ_B and ϱ_A move in the same direction. If, for example (as assumed in the text), sector A, where technological change occurs, is intensive in capital, that is $\varrho_B < \varrho_A$, then factor intensities will decrease in both sectors.

When technical progress has a capital-saving bias, $\beta > 1$ and from Eqs. (27.73) and (27.75) it can easily be seen that both ϱ_B and ϱ_A decrease when $\varrho_A > \varrho_B$; when, conversely, $\varrho_B > \varrho_A$, the sign of $d\varrho_A$ is ambiguous, whilst $d\varrho_B$ is certainly negative.

Finally, when technical progress has a labour-saving bias, $\beta < 1$, and from Eqs. (27.73) and (27.75) it can be readily seen that both ϱ_B and ϱ_A increase when $\varrho_B > \varrho_A$; when, conversely, $\varrho_B < \varrho_A$, the sign of $d\varrho_A$ is ambiguous, whilst ϱ_B certainly decreases.

Let us now examine *real factor rewards*. Letting $w = p_L/p_A$ and $r = p_K/p_A$, from the equilibrium conditions we have

$$\begin{aligned} w &= \lambda' g_A - \lambda \varrho_A g'_A = p\left(g_B - \varrho_B g'_B\right), \\ r &= \lambda g'_A = p g'_B. \end{aligned} \tag{27.79}$$

It is clearly simpler to use the last expression in each of these two relations, so that

$$\begin{aligned} dw &= p\left(dg_B - g'_B d\varrho_B - \varrho_B dg'_B\right) \\ &= p\left(g'_B d\varrho_B - g'_B d\varrho_B - \varrho_B g'' d\varrho_B\right) = -p\varrho_B g''_B d\varrho_B, \\ dr &= p g''_B d\varrho_B. \end{aligned} \tag{27.80}$$

Since $g''_B < 0$, we see that w (and so p_L, since p_A is fixed by assumption) moves in the same direction as ϱ_B whilst r (and so p_K) moves in the opposite direction, so that the relative price of factors $w/r = p_L/p_K$ certainly moves in the same direction as ϱ_B.

27.4.2 *Effects of Technical Progress on Output Levels*

As regards the *changes in output levels*, we begin with observing that, as technical progress brings about an outward shift of the transformation curve, it is *not* possible for both output levels to decrease: if one decreases, the other must necessarily increase. Since, as we shall see, in most cases it is the output of B which decreases, whilst in the remaining cases the change in the output of B has an ambiguous sign (hence A has too), we shall restrict ourselves to deriving the formulae that give dB, partly because these are relatively simpler. Since $B = L_B g_B(\varrho_B)$, we have

$$dB = g_B dL_B + L_B g'_B d\varrho_B. \tag{27.81}$$

We now express $\mathrm{d}L_B$ in terms of $\mathrm{d}\varrho_A$ and $\mathrm{d}\varrho_B$. If we consider the total differentials of Eqs. (27.29) and recall that factor endowments are unchanged, we have

$$\begin{aligned} \mathrm{d}L_A + \mathrm{d}L_B &= 0, \\ \varrho_A \mathrm{d}L_A + L_A \mathrm{d}\varrho_A + \varrho_B \mathrm{d}L_B + L_B \mathrm{d}\varrho_B &= 0, \end{aligned} \tag{27.82}$$

whence, solving for $\mathrm{d}L_A$, $\mathrm{d}L_B$ we get

$$\mathrm{d}L_A = \frac{L_A \mathrm{d}\varrho_A + L_B \mathrm{d}\varrho_B}{\varrho_B - \varrho_A}, \quad \mathrm{d}L_B = -\frac{L_A \mathrm{d}\varrho_A + L_B \mathrm{d}\varrho_B}{\varrho_B - \varrho_A}. \tag{27.83}$$

By substituting Eq. (27.83) into (27.81) and rearranging terms we get

$$\mathrm{d}B = -\frac{g_B L_A}{\varrho_B - \varrho_A} \mathrm{d}\varrho_A - \frac{\left(g_B - \varrho_B g_B'\right) L_B + g_A' \varrho_A L_B}{\varrho_B - \varrho_A} \mathrm{d}\varrho_B. \tag{27.84}$$

It can be readily checked that the numerators of both fractions are positive: the first is obviously so; as regards the second it is sufficient to remember that $g_B - \varrho_B g_B'$ is the marginal productivity of labour in sector B. We can now examine the sign of $\mathrm{d}B$ in the various cases.

1. Neutral technical progress.

(1a) $\varrho_B > \varrho_A$.We know from the previous analysis—see (27.77) and (27.78)—that $\mathrm{d}\varrho_A > 0$,$\mathrm{d}\varrho_B > 0$, and so from (27.84) we get $\mathrm{d}B < 0$ (consequently $\mathrm{d}A$ will be positive).
(lb) $\varrho_B < \varrho_A$. In this case $\mathrm{d}\varrho_A < 0$,$\mathrm{d}\varrho_B < 0$, so that from (27.84) we again have $\mathrm{d}B < 0$ (and so $\mathrm{d}A > 0$).

2. Capital-saving technical progress.

(2a) $\varrho_B > \varrho_A$. In this case $\mathrm{d}\varrho_B < 0$, whilst $\mathrm{d}\varrho_A$ can have either sign. Therefore the sign of $\mathrm{d}B$ (and, consequently, that of $\mathrm{d}A$) is ambiguous.
(2b) $\varrho_B < \varrho_A$. Here we have $\mathrm{d}\varrho_A < 0$, $\mathrm{d}\varrho_B < 0$, so that $\mathrm{d}B < 0$ (and, consequently, $\mathrm{d}A > 0$).

3. Labour-saving technical progress.

(3a) $\varrho_B > \varrho_A$. In this case $\mathrm{d}\varrho_A > 0$, $\mathrm{d}\varrho_B > 0$, and so $\mathrm{d}B < 0$ (thus $\mathrm{d}A > 0$).
(3b) $\varrho_B < \varrho_A$. Here $\mathrm{d}\varrho_B < 0$, whilst the sign of $\mathrm{d}\varrho_A$ is ambiguous. Therefore $\mathrm{d}B$ (and, consequently, $\mathrm{d}A$) have ambiguous signs.

Before passing on to the analysis of the effects of the various types of technological change on the terms of trade it is as well to summarize in tabular form all the results reached up to this point. In Table 27.2 the sign + denotes an increase, the sign − a decrease, the question mark an ambiguous result.

Table 27.2 Effects of the various types of technological change in sector A

Types of progress	$\varrho_A > \varrho_B$							$\varrho_B > \varrho_A$						
	ϱ_A	ϱ_B	p_L	p_K	$\frac{p_L}{p_K}$	A	B	ϱ_A	ϱ_B	p_L	p_K	$\frac{p_L}{p_K}$	A	B
Neutral	−	−	−	+	−	+	−	+	+	+	−	+	+	−
Cap.-saving	−	−	−	+	−	+	−	?	+	+	−	+	?	?
Lab.-saving	?	−	−	+	−	?	?	+	+	+	−	+	+	−

27.4.3 Effects of Technical Progress on the Terms of Trade

We can now examine the effects of technical progress on the terms of trade; as usual we must first determine the change in income (at unchanged p) to determine the marginal propensity to produce commodity A. If we consider the total differential (in which p has been kept constant)

$$\mathrm{d}I_A = \mathrm{d}A + p\mathrm{d}B, \tag{27.85}$$

we only have to determine $\mathrm{d}A$, as $\mathrm{d}B$ is known from Eqs. (27.29) and (27.84). It is however possible to obtain exact results without having to calculate $\mathrm{d}A$. In fact, if we observe that technological change brings about an outward shift of the transformation curve, we realize that the isoincome which is tangent to the new transformation curve with the same slope as the one tangent to the previous transformation curve, is necessarily higher than the old one, so that $\mathrm{d}I_A > 0$. Consider now the relation

$$\mu_{sY} = \left(\frac{\mathrm{d}A}{\mathrm{d}I_A}\right)_{p=\mathrm{const}} = \left(\frac{\mathrm{d}A}{\mathrm{d}A + p\mathrm{d}B}\right)_{p=\mathrm{const}}, \tag{27.86}$$

from which it follows that, when $\mathrm{d}A > 0$ and $\mathrm{d}B < 0$, the numerator is certainly greater than the denominator, so that $\mu_{sY} > 1$. Now, as is clear from Table 27.2, all the cases in which technological change has unambiguous effects on productive levels are the ones in which A increases and B decreases. In the cases in which these effects are ambiguous it is, a fortiori, not possible to determine the value of μ_{sY}. Therefore, if we exclude inferior goods (so that $0 < \mu_{dY} < 1$), it follows from (27.21) that in all cases in which technical progress has unambiguous effects on productive levels (neutral technical progress, or factor-saving progress in the sector intensive in the saved factor) the relative price p_B/p_A increases and so p_A/p_B decreases. The terms of trade will therefore move against or in favour of the country according to whether the innovating sector (sector A in our case) is that producing the exportable or the importable. It is also clear that when A is the importable no immiserizing growth is possible, while it is possible when A is the exportable.

In all cases of factor-saving technical progress in the sector intensive in the other factor, the indeterminacy of μ_{sY} does not allow us to find the direction in which p moves, so that any result is possible.

References

Allen, R. G. D. (1967). *Macro-economy theory: A mathematical treatment* (chap. 13).
Gandolfo, G. (2009). *Economic dynamics*.
Hicks, J. R. (1932). *The theory of wages.*
Johnson, H. G. (1955). Economic expansion and international trade.
Jones, R. W. (1965). The structure of simple general equilibrium models.
Vanek, J. (1966). Towards a more general theory of growth with technological change.

Chapter 28
Appendix to Chapter 14

28.1 A Dynamic Model of Growth and Trade

If we consider the growth-trade relations in a dynamic context, the static equations of the basic neoclassical model treated in Chap. 19 must be supplemented by the following differential equations (a dot indicates the time derivative and the subscript $i = 1, 2$ refers to the countries)

$$\begin{aligned} \dot{L}_i/L_i &= n_i, \\ \dot{K}_i &= S_i, \\ S_i &= s_i I_{iA} = A^D_{i,} \; 0 < s_i < 1. \end{aligned} \tag{28.1}$$

The first equation expresses the assumption of an exogenously given growth rate of the labour force.

The second equation defines the increment in the stock of capital (investment) which equals saving, owing to the assumption that all saving gets automatically invested. Also note that we have assumed no technical progress and no capital depreciation; it would not be difficult to allow for factor-augmenting technological change and for "radioactive" depreciation. Factor-augmenting technical progress at constant rates could be taken into account by modifying the first two equations, for example $\dot{L}_i/L_i = n_i + \lambda_i$, $\dot{K}_i = S_i + \lambda_i' K_i$, where λ_i and λ_i' are the relevant rates. As regards radioactive depreciation, this implies depreciation allowances proportional to the existing capital stock, so that a quantity $\delta_i K_i, 0 < \delta_i < 1$, has to be subtracted from (gross) investment to obtain net investment. However, for simplicity's sake we ignore them.

The third equation is the saving function, proportional to national income (measured in terms of commodity A). Given the assumption that A is the investment good and that saving is automatically invested, saving coincides with the domestic demand for commodity A. In other words, these assumptions imply that the domestic demand for A, instead of having the general form $A = A_i^D(I_{iA}, p)$,

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_28,

has the particularly simple form $A_i^D = s_i I_{iA}$. It goes without saying that saving-investment will be higher (lower) than A's domestic output according to whether country i imports or exports this commodity. In what follows we assume that the structure of international trade is such that A is country 1's importable. To reconcile the assumption of internationally immobile factors of production (the stock of capital among them) with the fact that the investment good, as a produced commodity, is internationally traded, we must introduce the further assumption that this commodity, internationally mobile as a produced commodity subject to international trade, once installed as an addition to the capital stock of a country, becomes internationally immobile. In other words, stocks of existing capital are immobile, flows of investment goods are mobile and traded: to adapt terms used for other purposes in vintage growth theory, we might talk of ex ante international mobility, ex post immobility.

Before continuing, we note a feature of the steady-state growth path. In steady-state by definition all the physical variables (stock of capital, labour force, output, exports, imports etc.) must grow at the same rate, which coincides with the natural rate of growth, n_i. Now, if we exclude international lending, the usual conditions (19.27) for international equilibrium must hold, that is

$$E_{1A} + E_{2A} = 0, \tag{28.2}$$

whence

$$\dot{E}_{1A} + \dot{E}_{2A} = 0,$$

and so

$$E_{1A}\frac{\dot{E}_{1A}}{E_{1A}} + E_{2A}\frac{\dot{E}_{2A}}{E_{2A}} = 0, \tag{28.3}$$

which can be rewritten—since $E_{2A} = -E_{1A}$ by (28.2)—in the form

$$E_{1A}\left(\frac{\dot{E}_{1A}}{E_{1A}} - \frac{\dot{E}_{2A}}{E_{2A}}\right) = 0. \tag{28.4}$$

Similarly we obtain

$$E_{1B}\left(\frac{\dot{E}_{1B}}{E_{1B}} - \frac{\dot{E}_{2B}}{E_{2B}}\right) = 0. \tag{28.5}$$

Since, as we said above, in steady-state we have $\dot{E}_{1A}/E_{1A} = \dot{E}_{1B}/E_{1B} = n_1$ and $\dot{E}_{2A}/E_{2A} = \dot{E}_{2B}/E_{2B} = n_2$, from Eqs. (28.4) and (28.5) it follows that $n_1 = n_2$. In other words, in the model under consideration there can be no steady-state growth unless the two countries have the same natural rate of growth. This limitation can be removed if we introduce international lending and investment, but, to keep complications to a minimum, we shall not do so (the reader is referred, for example, to Kemp (1969b) and Bardhan (1970)), i.e. we shall assume that all of a country's stock of capital is owned by residents. Thus we shall assume that $n_1 = n_2 = n$.

As in any growth model we must distinguish between the momentary or short-run equilibrium and the long-run equilibrium. The momentary equilibrium is that occurring at a given point of time. At any moment of time, the labour force and the capital stock are given (the former exogenously, and the latter as a result of past accumulation), so that the equations of the static model described in Chap. 19 determine the equilibrium quantities and prices: this is the momentary equilibrium. The dynamic equations (28.1) determine the path over time of factor endowments and so the steady growth path (the long run equilibrium). Before examining this, it is as well to show further properties of momentary equilibrium not treated in Chap. 19.

28.2 Momentary Equilibrium

If we denote by $\varrho_i \equiv K_i/L_i$ the relative factor endowments in the two countries, the demand and supply functions, and so the excess demand functions, also depend on ϱ_i, as we have seen in Sect. 21.3 (see, in particular, Eqs. (27.46) ff.). We can therefore write $E_{iA} = E_{iA}(p, \varrho_i)$, and so the equation of international equilibrium (19.27) can be expressed as

$$E_{1A}(p, \varrho_1) + E_{2A}(p, \varrho_2) = 0, \tag{28.6}$$

by which—in accordance with the implicit function theorem—we can express p as a function of ϱ_1,ϱ_2, provided that the condition $\partial E_{1A}/\partial p + \partial E_{2A}/\partial p \neq 0$ is satisfied. We know from Eq. (19.34) that $\partial E_{1A}/\partial p > 0$, similarly we have $\partial E_{2A}/\partial p > 0$; as the condition is satisfied, there exists the differentiable function

$$p = p\left(\varrho_1, \varrho_2, \right), \tag{28.7}$$

of which we propose to calculate the partial derivatives. For this purpose we define the per-capita excess demands $e_{iA} = E_{iA}/L_i$, so that

$$L_{1A}e_{1A}\left(p, \varrho_1\right) + L_2 e_{2A}\left(p, \varrho_2\right) = 0. \tag{28.8}$$

If we divide through by $L_1 + L_2$ and define $\lambda_i = L_i/(L_1 + L_2)$, we get

$$\lambda_1 e_{1A}\left(p, \varrho_1\right) + \lambda_2 e_{2A}\left(p, \varrho_2\right) = 0, \tag{28.9}$$

where λ_1,λ_2 are constants owing to the assumption that L_1 and L_2 grow at the same rate n. From Eq. (28.9), account being taken of (28.7), we have

$$\frac{\partial p}{\partial \varrho_i} = -\frac{\partial e_{iA}/\partial \varrho_i}{\lambda_1 \dfrac{\partial e_{1A}}{\partial p} + \lambda_2 \dfrac{\partial e_{2A}}{\partial p}}. \tag{28.10}$$

Since $E_{iA} = L_i e_{iA}$, we have $L_i(\partial e_{iA}/\partial p) = \partial E_{iA}/\partial p$ and so the denominator in (28.10) is positive. As regards the numerator, we must find an explicit expression for the per-capita excess demand e_{iA}. If we recall that A is the investment good, the demand for which is proportional to income, we have

$$E_{iA} = s_i I_{iA} - L_{iA} g_{iA}\left(\varrho_{iA}\right), \tag{28.11}$$

and so

$$e_{iA} = s_i y_{iA} - l_{iA} g_{iA}\left(\varrho_{iA}\right), \quad y_{iA} \equiv I_{iA}/L_i, \quad l_{iA} \equiv L_{iA}/L_i. \tag{28.12}$$

Since each factor is paid its marginal product and the production functions are homogeneous of the first degree, Euler's theorem tells us that national income equals the sum of factor rewards,

$$I_{iA} = wL_i + rK_i, \tag{28.13}$$

where w is the *MPL* and r the *MPK*, both expressed in terms of commodity A. In per-capita terms,

$$y_{iA} = w + r\varrho_i, \tag{28.14}$$

and so

$$\frac{\partial y_{iA}}{\partial \varrho_i} = r = g'_{iA} > 0, \tag{28.15}$$

since g'_{iA} is the *MPK* in sector A.

For future reference, it should be noted that if we use the definition of income as value of output we have

$$y_{iA} = l_{iA} g_{iA}\left(\varrho_{iA}\right) + p l_{iB} g_{iB}\left(\varrho_{iB}\right), \tag{28.16}$$

whence

$$\frac{\partial y_{iA}}{\partial p} = l_{iB} g_{iB}\left(\varrho_{iB}\right) > 0. \tag{28.17}$$

It must be stressed that in (28.16) we assumed a situation of incomplete specialization, so that each country continues to produce both commodities. It is possible to carry out a complete taxonomy of all possible cases (for which see, e.g., Bardhan, 1970; Kemp, 1969b; Oniki & Uzawa, 1965), but we shall consider only the case of incomplete specialization in both countries. We now observe that from the full employment conditions we get

$$\begin{aligned} l_{iA} + l_{iB} &= 1, \\ \varrho_{iA} l_{iA} + \varrho_{iB} l_{iB} &= \varrho_i, \end{aligned} \tag{28.18}$$

from which

$$l_{iA} = \frac{\varrho_{iB} - \varrho_i}{\varrho_{iB} - \varrho_{iA}}, \quad l_{iB} = \frac{\varrho_i - \varrho_{iA}}{\varrho_{iB} - \varrho_{iA}}. \tag{28.19}$$

Equipped with these results we can calculate $\partial e_{iA}/\partial \varrho_i$. From Eq. (28.11) we have

$$\frac{\partial e_{iA}}{\partial \varrho_i} = s_i \frac{\partial y_{iA}}{\partial \varrho_i} - \frac{\partial l_{iA}}{\partial \varrho_i} g_{iA}, \tag{28.20}$$

and so, by substituting (28.15) into Eq. (28.20) and recalling Eqs. (28.19),

$$\frac{\partial e_{iA}}{\partial \varrho_i} = s_i g'_{iA} + \frac{1}{\varrho_{iB} - \varrho_{iA}} g_{iA}. \tag{28.21}$$

This derivative has an unambiguous sign only if $\varrho_{iB} > \varrho_{iA}$, that is, if the capital intensity is greater in the consumer good sector than in the capital good sector (this is the capital intensity condition, widely used in growth theory); for simplicity's sake, we assume that this condition holds,[1] so that $\partial e_{iA}/\partial \varrho_i$. It follows from this result that

$$\frac{\partial p}{\partial \varrho_i} < 0. \tag{28.22}$$

28.3 Long Run Equilibrium

We can now pass to the long-run equilibrium and consider the derivatives

$$\dot{\varrho}_i = \frac{\dot{K}_i L_i - \dot{L}_i K_i}{L_i^2} = \frac{\dot{K}_i}{L_i} - \frac{\dot{L}_i}{L_i} \varrho_i, \tag{28.23}$$

---

[1] In Sect. 14.1 we adopted the opposite assumption, and warned the reader that in that case stability requires other conditions (for example that concerning the elasticity of substitution). Since the formal proof of stability would be further complicated, we prefer here to adopt the traditional capital-intensity condition.

whence, by using (28.1) and remembering that y_{iA} denotes per capita income, we have the differential equation system

$$\begin{aligned}\dot{\varrho}_1 &= s_1 y_{1A}(p, \varrho_1) - n\varrho_1,\\ \dot{\varrho}_2 &= s_2 y_{2A}(p, \varrho_2) - n\varrho_2,\end{aligned} \tag{28.24}$$

the solution of which determines the time paths of the relative factor endowments in the two countries, $\varrho_i(t)$, and, consequently, of all the other variables. To examine this system it is convenient to rewrite it in the form

$$\begin{aligned}\frac{\dot{\varrho}_1}{\varrho_1} &= \frac{s_1 y_{1A}(p, \varrho_1)}{\varrho_1} - n = \phi_1(\varrho_1, \varrho_2),\\ \frac{\dot{\varrho}_2}{\varrho_2} &= \frac{s_2 y_{2A}(p, \varrho_2)}{\varrho_2} - n = \phi_2(\varrho_1, \varrho_2).\end{aligned} \tag{28.25}$$

Let us now calculate the derivatives of the functions ϕ_i which are

$$\begin{aligned}\frac{\partial \phi_1}{\partial \varrho_1} &= \frac{s_1}{\varrho_1^2}\left[\left(\frac{\partial y_{1A}}{\partial p}\frac{\partial p}{\partial \varrho_1} + \frac{\partial y_{1A}}{\partial \varrho_1}\right)\varrho_1 - y_{1A}\right]\\ &= \frac{s_1}{\varrho_1^2}\left[\frac{\partial y_{1A}}{\partial p}\frac{\partial p}{\partial \varrho_1} - \left(y_{1A} - \frac{\partial y_{1A}}{\partial \varrho_1}\varrho_1\right)\right]\\ &= \frac{s_1}{\varrho_1^2}\left[\frac{\partial y_{1A}}{\partial p}\frac{\partial p}{\partial \varrho_1} - \left(y_{1A} - g_{1A}'\varrho_1\right)\right],\end{aligned} \tag{28.26}$$

$$\frac{\partial \phi_1}{\partial \varrho_2} = \frac{s_1}{\varrho_1}\frac{\partial y_{1A}}{\partial p}\frac{\partial p}{\partial \varrho_2}, \tag{28.27}$$

$$\frac{\partial \phi_2}{\partial \varrho_1} = \frac{s_2}{\varrho_2}\frac{\partial y_{2A}}{\partial p}\frac{\partial p}{\partial \varrho_1}, \tag{28.28}$$

$$\begin{aligned}\frac{\partial \phi_2}{\partial \varrho_2} &= \frac{s_2}{\varrho_2^2}\left[\left(\frac{\partial y_{2A}}{\partial p}\frac{\partial p}{\partial \varrho_2} + \frac{\partial y_{2A}}{\partial \varrho_2}\right)\varrho_2 - y_{2A}\right]\\ &= \frac{s_2}{\varrho_2^2}\left[\frac{\partial y_{2A}}{\partial p}\frac{\partial p}{\partial \varrho_2} - \left(y_{2A} - g_{2A}'\varrho_2\right)\right].\end{aligned} \tag{28.29}$$

From (28.17) and (28.22) it immediately follows that $\partial\phi_1/\partial\varrho_2$ and $\partial\phi_2/\partial\varrho_1$ are both negative. As regards $\partial\phi_1/\partial\varrho_1$ and $\partial\phi_2/\partial\varrho_2$—if we note that $y_{iA} - g_{iA}'\varrho_i$ is positive as it represents (see Eq. (28.14)) the *MPL*—we see that they are negative as well. The reader can also check as an exercise that

$$\frac{\partial\phi_1/\partial\varrho_1}{\partial\phi_1/\partial\varrho_2} > \frac{\partial\phi_2/\partial\varrho_1}{\partial\phi_2/\partial\varrho_2}. \tag{28.30}$$

The steady growth solution of system (28.25) corresponds to $\dot{\varrho}_1 = \dot{\varrho}_2 = 0$, that is, to the singular point of the system. Thus we must consider the equations

$$\begin{aligned} \phi_1(\varrho_1, \varrho_2) &= 0, \\ \phi_2(\varrho_1, \varrho_2) &= 0, \end{aligned} \tag{28.31}$$

the solution of which determines the equilibrium values ϱ_1^E, ϱ_2^E. Now, each of the relations in system (28.31) determines a curve in the (ϱ_1, ϱ_2) plane, with slope

$$\left(\frac{\mathrm{d}\varrho_2}{\mathrm{d}\varrho_1}\right)_{\phi_1=0} = -\frac{\partial\phi_1/\partial\varrho_1}{\partial\phi_1/\partial\varrho_2}, \quad \left(\frac{\mathrm{d}\varrho_2}{\mathrm{d}\varrho_1}\right)_{\phi_2=0} = -\frac{\partial\phi_2/\partial\varrho_1}{\partial\phi_2/\partial\varrho_2}. \tag{28.32}$$

Since—given the signs of the partial derivatives and given (28.30)—in the relevant interval both curves are monotonically decreasing though with different slopes, they will cross only once in that interval.[2]

Having thus ascertained the existence and uniqueness of equilibrium, we examine its stability. For this purpose we observe that, as the ϱ_i can take on positive values only, we can introduce the following transformation of variables

$$x_i = \ln \varrho_i, \quad \text{hence } \varrho_i = e^{x_i}, \quad \dot{x}_i = \dot{\varrho}_i/\varrho_i, \tag{28.33}$$

so that system (28.25) can be rewritten as

$$\begin{aligned} \dot{x}_1 &= \frac{s_1 y_{1A}(p, e^{x_1})}{e^{x_1}} - n = \omega_1(x_1, x_2), \\ \dot{x}_2 &= \frac{s_2 y_{2A}(p, e^{x_2})}{e^{x2}} - n = \omega_2(x_1, x_2). \end{aligned} \tag{28.34}$$

Since the x_i variables are a monotonically increasing transformation of the ϱ_i, the stability of system (28.34) implies the stability of system (28.25). We could examine the stability by way of phase diagrams, but we prefer to use more powerful analytical methods. Let us first observe that

$$\frac{\partial\omega_j}{\partial x_i} = \frac{\partial\phi_j}{\partial\varrho_i}\frac{\partial\varrho_i}{\partial x_i} = \frac{\partial\phi_j}{\partial\varrho_i}\varrho_i, \quad i, j = 1, 2, \tag{28.35}$$

[2] That is, in the interval of incomplete specialization. For a more detailed analysis of the existence and uniqueness of equilibrium in this and all other possible cases, see, e.g., the already cited works of Kemp (1969b), Oniki and Uzawa (1965), and Bardhan (1970).

so that the Jacobian matrix $[\partial\omega_j/\partial x_i]$ of system (28.34) turns out to be

$$\mathbf{J} = \begin{vmatrix} \dfrac{\partial\omega_1}{\partial x_1} & \dfrac{\partial\omega_1}{\partial x_2} \\ \\ \dfrac{\partial\omega_2}{\partial x_1} & \dfrac{\partial\omega_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial\phi_1}{\partial\varrho_1}\varrho_1 & \dfrac{\partial\phi 1}{\partial\varrho_2}\varrho_2 \\ \\ \dfrac{\partial\phi_2}{\partial\varrho_1}\varrho_1 & \dfrac{\partial\phi_2}{\partial\varrho_2}\varrho_2 \end{vmatrix}. \tag{28.36}$$

Since, as we have seen above, $\partial\phi_j/\partial\varrho_i < 0$, $i, j = 1, 2$, and given (28.30), it follows that, everywhere in the relevant interval,

$$\begin{aligned} \frac{\partial\omega_1}{\partial x_1} + \frac{\partial\omega_2}{\partial x_2} &< 0, \\ |\mathbf{J}| &> 0, \\ \frac{\partial\omega_1}{\partial x_1}\frac{\partial\omega_2}{\partial x_2} &\neq 0. \end{aligned} \tag{28.37}$$

Therefore, the conditions of Olech's theorem (see, for example, Gandolfo, 2009, p. 376) are satisfied and, consequently, the equilibrium state is globally stable.

Let us note, as a conclusion, that we have examined the problem of the dynamic relations between trade and growth solely in the context of positive economics. Normative problems (including intertemporal welfare maximization) are surveyed in Bhagwati and Srinivasan (1983, chap. 31); see also Smith (1977) and Findlay (1984). For an alternative view to the traditional one, see Parrinello (1979) and Steedman (1979); see also Smith (1984).

References

Bardhan, P. K. (1970). *Economic growth, development, and foreign trade: A study in pure theory*.

Bhagwati, J. N., & Srinivasan, T. N. (1983). *Lectures on international trade*.

Findlay, R. (1984). Growth and development in trade models.

Gandolfo, G. (2009). *Economic dynamics*.

Kemp, M. C. (1969b). *The pure theory of international trade and investment*.

Oniki, H., & Uzawa, H. (1965). Patterns of trade and investment in a dynamic model of international trade.

Parrinello, S. (1979). Distribution, growth, and international trade.

Smith, M. A. M. (1977). Capital accumulation in the open two-sector economy.

Smith, A. (1984). Capital theory and trade theory.

Steedman, I. (1979). *Trade amongst growing economies*.

Chapter 29
Appendix to Chapter 15

29.1 Endogenous Growth and Traditional Trade Theory

The constant-returns-to scale production functions for the tradable goods A, B and for the R&D services Z can be written in the intensive form (see Sect. 19.2.1)

$$\begin{aligned} A(t) &= \lambda(t) L_A g_A(\rho_A). \\ B(t) &= \lambda(t) L_B g_B(\rho_B), \\ Z &= L_Z g_Z(\rho_Z), \end{aligned} \tag{29.1}$$

where $\rho_i \equiv K_i/L_i, i = A, B, Z$ are the factor intensities in the three sectors, and $\lambda(t)$ is an index of technological efficiency, assumed uniform for both tradables. The crucial assumption (Findlay, 1995) is that $\lambda(t)$ is endogenously determined by the per-capita output of R&D services. More precisely, the proportional rate of change of λ is assumed to depend on $z = Z/L = l_Z g_Z(\rho_Z)$, where $L = L_A + L_B + L_Z$ is the given and constant amount of labour existing in the economy and $l_Z = L_Z/L$ is the share of the labour force employed in the R&D sector:

$$\dot{\lambda}/\lambda = \phi(z), \quad \phi'(z) > 0, \quad \phi''(z) < 0. \tag{29.2}$$

A greater output of R&D services enhances the rate of technical progress, but this enhancement is subject to diminishing returns ($\phi''(z) < 0$).

Due to the small open economy assumption, the terms of trade or relative price of tradables $p = p_B/p_A$, is given from the outside. Let us take good A as the numéraire, so that $p = p_B$. At the initial time $t = 0$ we can set $A(0) = 1$ and, assuming that all three goods are produced, knowledge of the relative price p_B and of the production functions is sufficient to determine all factor intensities and rewards as well as the price of the non-traded good p_Z in terms of the numéraire (see Sects. 6.6 and 22.5) independently of demand conditions.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_29,

To solve the optimal resource allocation problem between tradables and R&D (since K and L are given and constant, more current output of tradables means less technical progress and vice versa) Findlay first defines the value of the per capita output of tradables. This can easily be done thanks to the assumption of a given p. Denoting such a value by $\nu(t)$, we have

$$\begin{aligned}\nu(t) &= [A(t) + p_B B(t)]/L = \lambda(t)\,[L_A g_A(\rho_A) + p_B L_B g_B(\rho_B)]/L \\ &= \lambda(t)\,[l_A g_A(\rho_A) + p_B l_B g_B(\rho_B)]\,,\end{aligned} \tag{29.3}$$

where $l_A = L_A/L, l_B = L_B/L$ are the shares of the total labour force allocated to tradables. For any given amount of labour L_Z allocated to the R&D sector, the remaining amount $L-L_Z$ will be optimally allocated to A and B so that (the optimal values of) l_A, l_B will be determined. Hence ν will change through time only because of technical progress λ, since the expression in square brackets is constant as long as L_Z is constant.

Equation (29.3) clearly shows the trade-off between technical progress and current output. If more resources are allocated to R&D, namely if more z is produced, less resources will be allocated to A and B. Hence at any given point in time per capita output $l_A g_A(\rho_A) + p_B l_B g_B(\rho_B)$ will negatively depend on z, so that we can write

$$l_A g_A(\rho_A) + p_B l_B g_B(\rho_B) = \upsilon(z), \quad \upsilon'(z) < 0, \quad \upsilon''(z) = 0. \tag{29.4}$$

Let us observe that $\upsilon''(z) = 0$ is not an assumption, for υ is a negative linear function of z. To show this, let us first observe that, since the relative price p_B is fixed, we can apply Hicks' theorem (1939, 1946) according to which, if the relative prices of a group of goods remain constant as the quantity of the goods themselves varies, the different goods in the group can be considered as a single whole, that is, as if they were a single good. We next observe that, at any given point in time, in our neoclassical perfectly competitive setting resources are optimally allocated, which implies that $(-\mathrm{d}\upsilon/\mathrm{d}z)$, namely the marginal rate of transformation between the composite commodity υ (tradables) and z, equals the price ratio p_B/p_Z (see Chap. 3). Since, as we have seen above, not only p_B but also p_Z is given from the outside, it follows that $\upsilon'(z) = \mathrm{d}\upsilon/\mathrm{d}z = -p_B/p_Z$ is a constant.

To ascertain the effect of R&D expenditure on the future value of the output of tradable goods we begin by differentiating $\nu(t)$ with respect to time. This gives

$$\begin{aligned}\dot{\nu} &= \dot{\lambda}\,[l_A g_A(\rho_A) + p_B l_B g_B(\rho_B)] \\ &= \lambda(t)\phi(z)\,[l_A g_A(\rho_A) + p_B l_B g_B(\rho_B)] \\ &= \lambda(t)\phi(z)\upsilon(z),\end{aligned} \tag{29.5}$$

where we have used (29.2) and (29.4). Setting $\lambda(t) = 1$ at $t = 0$ (see above) we can ascertain the marginal benefit from R&D by differentiating Eq. (29.5) with respect to z:

$$\left(\frac{\partial \dot{v}}{\partial z}\right)_{t=0} = \phi'(z)\upsilon(z) + \phi(z)\upsilon'(z) = \frac{\phi(z)\upsilon(z)}{z}\left[\frac{z}{\phi}\frac{\mathrm{d}\phi(z)}{\mathrm{d}z} + \frac{z}{\upsilon}\frac{\mathrm{d}\upsilon(z)}{\mathrm{d}z}\right], \tag{29.6}$$

where the two expressions in square brackets are the elasticities of $\phi(z)$ and $\upsilon(z)$ with respect to z.

Since this marginal benefit will accrue from now to infinity, its present value is simply

$$\frac{1}{\delta}\left(\frac{\partial \dot{v}}{\partial z}\right)_{t=0}. \tag{29.7}$$

Under perfect competition, the marginal cost of a unit of R&D output is simply p_Z. Hence marginal benefit and marginal cost are equated when

$$\frac{1}{\delta}\left(\frac{\partial \dot{v}}{\partial z}\right)_{t=0} = p_Z$$

or

$$\frac{\phi(z)\upsilon(z)}{z}\left[\frac{z}{\phi}\frac{\mathrm{d}\phi(z)}{\mathrm{d}z} + \frac{z}{\upsilon}\frac{\mathrm{d}\upsilon(z)}{\mathrm{d}z}\right] = \delta p_Z. \tag{29.8}$$

The second order condition for a maximum is

$$\left(\frac{\partial^2 \dot{v}}{\partial z^2}\right)_{t=0} = \left[\phi''(z)\upsilon(z) + 2\phi'(z)\upsilon'(z) + \phi(z)\upsilon''(z)\right] < 0, \tag{29.9}$$

which implies the concavity of the $\dot{v}$ function (curve *OF* shown in Fig. 15.1 in the text). Condition (29.9) is certainly satisfied given the signs of the various derivatives.

29.2 Endogenous Growth and Trade: Innovation, Imitation, and Product Cycles

The model described here is due to Grossman and Helpman (1991a, 1991b, 1991c).

29.2.1 Demand

There is a continuum of different products, indexed by $j \in [0, 1]$. At any moment only a subset of these products is available, represented by the number of brands that have been so far developed. The number of available brands is represented by the set $[0, n(t)]$, where $n(t)$ can be considered as the measure of products developed before time t. In each product line j there are several qualities ranked in increasing order of quality from 0 to m; the highest quality available is the state-of-the-art. The quality $q_i(j)$ is represented by a parameter $\lambda > 1$, and units of quality are chosen so that the lowest quality existing (quality 0) offers one unit of service, so that $q_0(j) = \lambda^0 = 1$, and $q_m(j) = \lambda^m$.

The representative household is assumed to maximize utility over an infinite horizon

$$U_t = \int_t^\infty \mathrm{e}^{-\rho(\tau-t)} \log D(\tau)\mathrm{d}\tau, \tag{29.10}$$

where $D(\tau)$ is an index of consumption at time τ and ρ is the subjective discount rate. The natural logarithm of $D(\tau)$ measures instantaneous utility at time τ.

To simplify the problem we assume functional separability, so that the household can solve its intertemporal maximization problem in two stages. In one stage it chooses the composition of any given level of expenditure so as to maximize instantaneous utility. In the other stage it optimizes the time path of spending. The two stages can be taken in any order. Let us begin with the intertemporal maximization problem.

The household is endowed with one unit of labour and possesses an amount of wealth $W(t)$. It can freely lend or borrow at the instantaneous interest rate $r(t)$. The household's intertemporal maximization problem subject to its intertemporal budget constraint can thus be written as

$$\begin{aligned} &\max U_t = \int_t^\infty \mathrm{e}^{-\rho(\tau-t)} \log D(\tau)\mathrm{d}\tau \\ &\text{sub} \\ &\int_t^\infty \mathrm{e}^{-[R(\tau)-R(t)]} P_D(\tau) D(\tau)\mathrm{d}\tau \le \int_t^\infty \mathrm{e}^{-[R(\tau)-R(t)]} w(\tau)\mathrm{d}\tau + W(t), \end{aligned} \tag{29.11}$$

where $R(\tau) = \int_0^\tau r(s)\mathrm{d}s$ is the discount factor from time τ to time zero, $P_D(\tau)$ is the aggregate price index corresponding to the quantity index $D(\tau)$ in the instantaneous budget constraint

$$D(\tau) = \frac{E(\tau)}{P_D(\tau)}, \tag{29.12}$$

where $E(\tau)$ is the given value of spending, and $w(\tau)$ is the wage rate. The intertemporal budget constraint requires the present value of spending not to exceed the present value of (labour) income plus the initial wealth. index corresponding to the quantity index $D(\tau)$.

The maximization problem (29.37) is a simple calculus-of-variations problem with an integral constraint. It can be solved (see, for example, Gandolfo, 2009, chap. 27, sect. 27.2.2; Kamien & Schwartz, 1991, chap. 7) by forming the Lagrangian

$$\Lambda = \mathrm{e}^{-\rho(\tau-t)} \log D(\tau) - \zeta(t) \left\{ \mathrm{e}^{-[R(\tau)-R(t)]} \left[P_D(\tau) D(\tau) - w(\tau) \right] \right\}, \qquad (29.13)$$

where ζ is a Lagrange multiplier, and then calculating the first-order condition

$$\frac{\partial \Lambda}{\partial D(\tau)} = \frac{\mathrm{e}^{-\rho(\tau-t)}}{D(\tau)} - \zeta(t) \mathrm{e}^{-[R(\tau)-R(t)]} P_D(\tau) = 0, \quad \forall \tau \geq t, \qquad (29.14)$$

whence

$$\frac{\mathrm{e}^{-\rho(\tau-t)}}{D(\tau)} = \zeta(t) \mathrm{e}^{-[R(\tau)-R(t)]} P_D(\tau). \qquad (29.15)$$

Let us now consider household spending as defined in (29.12), and differentiate it logarithmically with respect to time (τ). We have

$$\frac{\dot{E}}{E} = \frac{\dot{D}}{D} + \frac{\dot{P}_D}{P_D}.$$

Logarithmic differentiation of (29.15) with respect to τ yields

$$-\rho - \frac{\dot{D}}{D} = -\frac{\mathrm{d}}{\mathrm{d}t} R(t) + \frac{\dot{P}_D}{P_D}.$$

Now,

$$\frac{\mathrm{d}}{\mathrm{d}t} R(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau} \int_0^\tau r(s) \mathrm{d}s = r(\tau),$$

where we have used the rule for the differentiation of an integral with respect to a parameter that occurs in the limit(s) of integration.

Substituting these results into the expression for $\dot{E}/E$ we obtain

$$\frac{\dot{E}}{E} = r - \rho, \qquad (29.16)$$

namely in each instant aggregate expenditure must grow at a rate given by the difference between the interest rate and the subjective discount rate.

For convenience Grossman and Helpman (1991a) use as a normalization condition the condition that nominal spending remains constant,

$$E = 1. \qquad (29.17)$$

From (29.16) and (29.17) we get

$$r = \rho, \qquad (29.18)$$

namely the nominal interest rate is always equal to the subjective discount rate.

Let us now consider the stage in which the household chooses the composition of the given level of expenditure so as to maximize instantaneous utility.

Preferences embodied in instantaneous utility are assumed to be of the type

$$\log D(t) = \int_0^1 \log\left[\sum_m q_m(j)x_{mt}(j)\right] \mathrm{d}j, \tag{29.19}$$

where the summation extends over the set of qualities of product j that is available at time t. These preferences have the convenient property that vertically differentiated products in any industry are perfect substitutes for one another, once quantities are adjusted for quality differences (for example, one $1/\lambda^m$ units of the state-of-the-art product are equivalent to one unit of the lowest quality product). Further, products of different industries enter the utility index symmetrically, and the elasticity of substitution between every pair of product lines is constant and equal to one. We know that with this sort of preferences (see Sect. 23.2.1) the household's income will be equally divided between all available product lines; besides, in each line the household will purchase only one quality, namely the quality $\tilde{m}_t(j)$ that has the lowest price per unit of quality. Thus we have the demand functions

$$x_{mt}(j) = \begin{cases} \dfrac{E(t)}{p_{mt}(j)} & \text{for } m = \tilde{m}_t(j), \\ \\ 0 & \text{otherwise,} \end{cases}$$

where $p_{mt}(j)$ is the price of quality m at time t, and $E(t)$ denotes spending at time t.

When we consider a two-country world, the representative household is taken to be the same in both countries as is the interest rate, and the normalization condition on $E(t)$ also refers to world expenditure.

29.2.2 Supply: Product Quality, Innovation, and Imitation

Labour is the only factor of production in each region (North and South), and can be used for manufacturing or research.

29.2.2.1 The Manufacturing Sector

In manufacturing, production takes place under fixed coefficients, and units are chosen so that one unit of labour is needed to produce one unit of output, whatever the quality (the development of a higher-quality commodity requires R&D expenditure, on which more below, but when this higher-quality commodity is manufactured, its labour requirement is the same as that for lower qualities). Thus marginal and average cost is simply the wage rate. Letting w^N, w^S respectively denote the wage rate in North and South, and assuming that oligopolistic competition takes the form of price competition, the possibility of imitation by South requires $w^N > w^S$.

The three categories of firms have been described in the text (see Sect. 15.3), and their measure is denoted by n^{NN}, n^{NS}, n^S, normalized in such a way that

$$n^{NN} + n^{NS} + n^S = 1. \tag{29.20}$$

Let us now consider the profit rate for each category of firms and show that each good in the continuum has only one producer, beginning with a category (iii) firm. This is a Southern firm that has mastered via imitation the technique for producing some state-of-the-art product, and competes with the Northern firm that has developed this product. Assuming a Bertrand duopoly (hence price competition: on Bertrand duopoly in general see, for example, Mas-Colell, Whinston, & Green, 1995, chap. 12, sect. 12.C), the Southern firm perceives a perfectly elastic demand when it charges a price of w^N (the unit production cost of its Northern rival), a unit elastic demand at lower prices, and zero demand at higher prices. It follows that this firm maximizes profit by setting a limit price exactly equal to w^N, the unit cost of production of its Northern rival. The latter will of course have zero profit and hence leave the market. Also, no other Southern firm is willing to invest resources just to become the second regional producer of the commodity and earn zero profit (in fact, if two Southern firms learn to imitate the same state-of-the-art product, in their Bertrand competition each sets a price equal to its marginal cost w^S and earns zero profit).

Thus there will be only one producer of the commodity, a Southern firm, whose sales at the price w^N are $E/w^N = 1/w^N$ (given the normalization condition $E = 1$, see Sect. 29.2.1). Its profit rate is

$$\pi^S = \frac{w^N - w^S}{w^N}. \tag{29.21}$$

We next consider a category (ii) firm, namely a Northern firm that has just innovated and can produce the next generation commodity, while the Southern firm that is producing the previous state-of-the-art commodity finds that its product is now second-to-top. The Northern leader can charge a quality premium λ over the price of the second-to-top commodity. Since the Southern firm cannot profitably set a price below w^S, the Northern leader can capture the entire market by setting a price marginally below λw^S and hence selling $1/\lambda w^S$ units of output. This outcome will of course only be possible if the quality premium is sufficiently great so that the Northern leader can charge a price higher than its unit cost of production, namely it is required that $\lambda w^S > w^N$. This is assumed to be the case, for otherwise there would be no innovation at all.

Hence the profit rate of the Northern leader competing with a Southern firm will be

$$\pi^{NS} = \frac{\lambda w^S - w^N}{\lambda w^S}. \tag{29.22}$$

We finally consider a category (i) firm, namely a Northern leader that has just innovated and is competing with a Northern follower that can produce the second highest quality. In this case the limit price that maximizes the leader's profit and captures the entire market is λw^N, which entails an output of $1/\lambda w^N$. Hence the leader's profit rate will be

$$\pi^{NN} = \frac{\lambda w^N - w^N}{\lambda w^N}. \tag{29.23}$$

When a second Northern firm gains the ability to produce the same top-quality product, price competition between these two firms drives both to charge a price equal to w^N, hence zero profits. Since innovation and imitation are costly activities, there can be only one producer in this case as well.

29.2.2.2 Innovation

Innovation is assumed to take place only in North. New higher-quality products can be developed only through costly R&D activity. Broadly following Aghion and Howitt (1990, 1992), Grossman and Helpman model this activity as a risky process, in the sense that an entrepreneur who devotes resources to R&D has a probability of success (i.e., of actually developing a new higher-quality variety of the product) that is proportional to the scale of his efforts but smaller than unity. More precisely, a Northern entrepreneur can obtain a probability $\iota \mathrm{d}t$ of success in the time interval $\mathrm{d}t$ by devoting $a\iota$ units of labour to research during $\mathrm{d}t$.

We must however distinguish between leaders and followers. A Northern leader, through the development of the current top-quality product, has accumulated a stock of knowledge and product-specific information that can help to achieve a further technological advance. This knowledge is not available to other firms, hence a Northern follower will have to invest a greater amount of resources to try to upgrade the state-of-the-art product (i.e., to develop the next generation product). Let $a_L \iota < a_F \iota$ (where a_L, a_F are constant coefficients) be the amount of labour that a leader and a follower must respectively allocate to research in the time interval $\mathrm{d}t$ to achieve the probability $\iota \mathrm{d}t$ of success.

This research advantage entails that only leaders will invest resources in R&D to try and recapture the market when a Southern firm has successfully imitated the state-of-the-art product.

To show this, let us recall that successful imitation by a Southern firm crowds the Northern leader out of the market. When this occurs, neither the Northern leader nor Northern followers are producing in the product line under consideration. Hence a Northern firm that succeeds in upgrading the (South-produced) state-of-the-art product earns the same profit independently of whether it is a leader or a follower, as it makes the same two-step profit jump (from zero to the profit corresponding to the production of the next-generation commodity). Since followers undergo higher research costs, they will be crowded out by leaders.

Let us indicate by $\iota^S \mathrm{d}t$ the probability of success of a Northern firm that undertakes research because its state-of-the-art product has been copied by a Southern imitator (the superscript S serves to indicate that the research effort is targeted at upgrading a Southern product). This Northern firm undergoes a research cost of $w^N a_L \iota^S \mathrm{d}t$ and has an expected gain of $\upsilon^{NS} \iota^S \mathrm{d}t$, where υ^{NS} denotes the value of a category (ii) firm, namely a Northern firm that has a Southern firm as its closest competitor. Maximization of net expected value $\upsilon^{NS} \iota^S \mathrm{d}t - w^N a_L \iota^S \mathrm{d}t$ gives

$$\upsilon^{NS} \leq w^N a_L, \quad \iota^S \geq 0, \tag{29.24}$$

with $\upsilon^{NS} = w^N a_L$ if $\iota^S > 0$.

We now consider the case in which a state-of-the-art product has escaped imitation by South (imitation activity is not always successful). Then not only the Northern leader but also Northern followers have an incentive to undertake research leading to the development of the next generation of products, as in such a situation they stand to gain more from a research success than do leaders.

In fact, the Northern leader would pass from its current positive profit (deriving from the production of the state-of-the-art commodity) to the higher profit deriving from the production of the next-generation commodity, hence it would make a one-step jump in profits. A Northern follower would instead pass from its current situation of zero profit (as it is currently producing nothing) to the profit deriving from the production of the next-generation commodity, hence it would make a two-step jump in profits.

On the other hand, the leader has a research advantage over the follower hence lower research costs as shown above. Thus both leaders and followers may undertake research. It is assumed (Grossman & Helpman, 1991a, p. 315) that only followers do. Since no leader undertakes research, this implies that there will be at most a single quality step between any leader and its closest competitor: if, in fact, research by a follower is successful, this follower will become a leader making a two-step jump, and the previous leader will become a follower with a one-step gap (the previous state-of-the-art-product is now second highest).

If we denote by υ^{NN} the value of a Northern firm (a follower that may become a leader thanks to successful innovation) that has another Northern firm (the previous leader) as its nearest competitor, this firm undergoes a research cost of $w^N a_F \iota^N \mathrm{d}t$ and has an expected gain of $\upsilon^{NN} \iota^N \mathrm{d}t$. Maximization of net expected value gives

$$\upsilon^{NN} \leq w^N a_F, \quad \text{with equality if } \iota^N > 0. \tag{29.25}$$

29.2.2.3 Imitation

Southern firms cannot develop next generation products but can copy any state-of-the-art product developed by Northern firms. Imitation, however, is a costly and risky venture that requires investment of resources in research with a related probability of success (i.e., of succeeding in developing a marketable copy of the

Northern state-of-the-art product), and is modelled like innovation. A Southern imitator can achieve a probability $m\mathrm{d}t$ of success in the time interval $\mathrm{d}t$ by devoting $a_m m$ units of labour to research during $\mathrm{d}t$. Maximization of net expected value $\upsilon^S m\mathrm{d}t - w^S a_m m\mathrm{d}t$ gives

$$\upsilon^S \leq w^S a_m, \quad \text{with equality if } m > 0. \tag{29.26}$$

29.2.3 The No-Arbitrage Conditions

Let us consider a category (i) firm, namely a Northern leader facing competition by a Northern follower. In a small interval of time $\mathrm{d}t$ the owners of the firm collect profits $\pi^{NN}\mathrm{d}t$, have a capital gain (or loss) given by the normal change in the firm's value $\dot{\upsilon}^{NN}\mathrm{d}t$, and have a probability $(\iota^N + m)\mathrm{d}t$ of suffering the total loss of the firm. In fact, the leading firm loses its market—and hence its value drops to zero—both when another Northern entrepreneur develops the next-generation product (an event that has the probability $\iota^N\mathrm{d}t$ of occurring, as shown above) and when a Southern imitator successfully copies its product (an event that has the probability $m\mathrm{d}t$ of occurring). The expected value of such a loss is $\upsilon^{NN}(\iota^N + m)\mathrm{d}t$.

Expressing these magnitudes as instantaneous rates, namely dividing them by $\upsilon^{NN}\mathrm{d}t$, the standard no-arbitrage condition (namely the condition that the firm's equities yield a normal rate of return) requires that the rate of profit plus the rate of normal capital gain or loss, minus the expected rate of total loss be equal to the current rate of interest in North, namely

$$\frac{\pi^{NN}}{\upsilon^{NN}} + \frac{\dot{\upsilon}^{NN}}{\upsilon^{NN}} - (\iota^N + m) = r^N. \tag{29.27}$$

Similarly a category (ii) firm, namely a Northern firm that has just upgraded a state-of-the-art commodity produced in South, gives rise to the condition

$$\frac{\pi^{NS}}{\upsilon^{NS}} + \frac{\dot{\upsilon}^{NS}}{\upsilon^{NS}} - (\iota^N + m) = r^N, \tag{29.28}$$

since the probability of total loss is the same as in the previous case.

We finally have a category (iii) firm, namely a Southern firm that has successfully copied a state-of-the-art product previously manufactured in North. As seen above, in the time interval $\mathrm{d}t$ this firm faces a probability $\iota^S\mathrm{d}t$ of displacement by a Northern firm that succeeds in upgrading the product. The usual no-arbitrage condition gives

$$\frac{\pi^S}{\upsilon^S} + \frac{\dot{\upsilon}^S}{\upsilon^S} - \iota^s = r^S, \tag{29.29}$$

where r^S is the rate of interest in South.

29.2.4 The Labour Markets

Labour markets are in equilibrium at every moment of time. Let us begin with North. Northern labour is employed in manufacturing and research.

In Northern manufacturing, the n^{NN} firms that have a Northern firm as their nearest competitor produce $1/\lambda w^N$ each. Since we have assumed that one unit of labour is required per unit of output, total labour demand by these firms is $n^{NN}/\lambda w^N$. Similarly the labour demand by the n^{NS} Northern firm that have a Southern firm as their closest competitor is $n^{NS}/\lambda w^S$.

To determine labour demand in the Northern R&D sector, we first recall that n^S is the number of successful Southern imitators and hence of imitated products. It follows that the measure of Northern firms engaged in research aimed at developing the next generation product to replace each imitated product will be n^S; each of these firms will demand $a_L \iota^S$ units of labour for research (see above), hence a total demand of $n^S a_L \iota^S$. Similarly total labour demand for research by Northern followers targeted at upgrading each of n^N state-of-the art goods produced by Northern leaders will be $n^N a_F \iota^N$.

In conclusion, clearing of the Northern labour market requires

$$\frac{n^{NN}}{\lambda w^N} + \frac{n^{NS}}{\lambda w^S} + n^S a_L \iota^S + n^N a_F \iota^N = L^N, \tag{29.30}$$

where L^N is the exogenously given labour supply in North.

Let us now consider South. The manufacturing sector produces $1/w^N$ units each of n^S imitated products, hence a total labour demand of $n^S\ /w^N$ units of labour.

As regards the Southern R&D sector, research for imitation is aimed at imitating n^N state-of-the art goods (currently produced by Northern leaders) and is carried out using $a_m m$ units of labour for each targeted good, hence a total labour demand of $n^N a_m m$ units.

Thus market clearing in South, given an exogenous labour supply L^S, requires

$$\frac{n^S}{w^N} + n^N a_m m = L^S. \tag{29.31}$$

29.2.5 Steady-State Equilibrium

Since the amount of resources (labour) is assumed to be constant, growth is due solely to R&D in innovation and imitation. In the steady state, resources are allocated in unchanging shares to the various activities (manufacturing and R&D). This in turn implies constant numbers of commodities in the various product categories in all regions as well as constant rates of innovation and imitation, and constant relative prices. Constancy of relative prices necessitates all nominal variables to grow at the same rate. This requires that the value of all firms grows at the same rate as nominal expenditure. Since we have assumed above (see Sect. 29.2.1) the normalization condition $E(t) = 1$ for all t, it follows that the values of firms remain constant. The normalization condition also implies that the interest rate equals the households' discount rate ρ (see Sect. 29.2.1), hence

$$r^N = r^S = \rho. \tag{29.32}$$

If we use these facts and combine each no-arbitrage condition with the corresponding value-maximization condition—namely Eq. (29.27) with Eq. (29.25), and so on—we get the steady-state relationships

$$\frac{\pi^{NN}}{a_F w^N} \leq \iota^N + m + \rho, \qquad \text{with equality if } \iota^N > 0, \tag{29.33}$$

$$\frac{\pi^{NS}}{a_L w^N} \leq \iota^N + m + \rho, \qquad \text{with equality if } \iota^S > 0, \tag{29.34}$$

$$\frac{\pi^{S}}{a_m w^S} \leq \rho + \iota^S, \qquad \text{with equality if } m > 0. \tag{29.35}$$

Let us now consider the number and composition of products, that have to remain constant as said above.

In North, category (i), namely Northern leaders that compete with Northern followers, expands whenever a follower succeeds in innovating in an industry where formerly a Southern firm was second to top. This takes place at the success rate $\iota^N n^{NS}$. This category, however, shrinks whenever a Southern firm succeeds in imitating one of the n^{NN} state-of-the-art commodities, an occurrence that takes place at the success rate mn^{NN}. Thus a constant number and composition of Northern industries requires equal inflows and outflows, that is,

$$\iota^N n^{NS} = mn^{NN}. \tag{29.36}$$

As regards South, the number of firms increases whenever an entrepreneur successfully imitates one of the

$$n^N \equiv n^{NS} + n^{NN} \tag{29.37}$$

Northern products. This occurs at the success rate mn^N. On the other hand, Southern firms drop out of the market at the rate $\iota^S n^S$, as next-generation products are developed in the labs of the n^S former Northern leaders that were displaced by Southern imitation. The number of Southern products (firms) remains constant when inflows match outflows, namely

$$mn^N = \iota^S n^S. \tag{29.38}$$

Equations (29.33)–(29.38) describe the steady-state of the model, which, however, is not uniquely determined. There are, in fact, several types of possible steady-state equilibria depending on the values of parameters, in particular of the labour input coefficients in research (parameters a_m, a_L, a_F). If these are too great, R&D is too costly, and no innovation or imitation will take place (a stationary equilibrium). Or a_m might be too large with respect to a_L, a_F, so that no imitation will take place and there will only be innovation. These cases are not really interesting, hence we shall assume that parameter values are such that both innovation and imitation are present.

Even so, however, there are two possible cases. In one, both leaders and followers undertake R&D because followers are also efficient at innovation, though less so than leaders (the "efficient followers" case). In the other, only leaders engage in research because of their relative superiority in the research lab (the "inefficient followers" case). Both cases are examined below.

29.2.5.1 Efficient Followers

In this case, conditions (29.33)–(29.35) hold as equalities. Using them together with the other equilibrium relationships it is possible to derive reduced-form equations for the variables we are interested in. These are:

- The aggregate rate of product improvement, $\iota \equiv \iota^N n^N + \iota^S n^S$;
- The aggregate rate of technology transfer to South, $\mu \equiv m n^N$;
- The relative wage rate of South, $\omega \equiv w^S / w^N$.

Using these definitions and the steady-state relationships (29.36)–(29.38), we immediately get

$$\begin{aligned} \iota^S n^S &= m n^N = \mu, \\ \iota^N n^N &= \iota - \iota^S n^S = \iota - m n^N = \iota - \mu. \end{aligned} \tag{29.39}$$

From Eqs. (29.36) and (29.37) we have

$$\begin{aligned} \iota^N n^{NS} &= m(n^N - n^{NS}) \\ &= m n^N - m n^{NS}, \end{aligned}$$

hence

$$(\iota^N - m) n^{NS} = m n^{NS}.$$

Multiplying through by n^N, using Eqs. (29.39), and solving for n^{NS} we get

$$n^{NS} = n^N \mu / \iota. \tag{29.40}$$

This last equation, together with Eq. (29.37), yields

$$\begin{aligned} n^{NN} &= n^N - n^{NS} \\ &= n^N - n^N \mu/\iota \\ &= n^N (1 - \mu/\iota). \end{aligned} \tag{29.41}$$

Using Eqs. (29.39)–(29.41) and the definition of ω we can rewrite the Northern labour-market equilibrium condition (29.30) as

$$\frac{n^N (1 - \mu/\iota)}{\lambda w^N} + \frac{n^N \mu/\iota}{\lambda \omega w^N} + a_L \mu + a_F (\iota - \mu) = L^N. \tag{29.42}$$

As regards the Southern labour market, take Eq. (29.31), replace $n^N m$ with μ, and observe that $n^S \equiv 1 - n^N$ by (29.21) and (29.37). This gives

$$\frac{1 - n^N}{w^N} + a_m \mu = L^S. \tag{29.43}$$

The next step is to combine the profit expressions (29.21)–(29.23) with the no-arbitrage conditions (29.33)–(29.35) and the steady-state conditions (29.36) and (29.38).

We first rewrite the profit expressions as

$$\begin{aligned} \pi^S &= w^S (w^N / w^S - 1)/w^N, \\ \pi^{NS} &= (1 - w^N / \lambda w^S), \\ \pi^{NN} &= (1 - 1/\lambda). \end{aligned}$$

From the no-arbitrage conditions we have

$$\begin{aligned} \pi^S &= a_m w^S (\rho + \iota^S), \\ \pi^{NS} &= a_L w^N (\iota^N + m + \rho), \\ \pi^{NN} &= a_F w^N (\iota^N + m + \rho), \end{aligned}$$

hence

$$\begin{aligned} (w^N / w^S - 1)/w^N &= a_m (\rho + \iota^S), \\ (1 - w^N / \lambda w^S)/w^N &= a_L (\iota^N + m + \rho), \\ (1 - 1/\lambda)/w^N &= a_F (\iota^N + m + \rho). \end{aligned} \tag{29.44}$$

Given the definition of the relative wage rate ω we have

$$\begin{aligned} (1/\omega - 1)/w^N &= a_m (\rho + \iota^S), \\ (1 - 1/\lambda\omega)/w^N &= a_L (\iota^N + m + \rho), \\ (1 - 1/\lambda)/w^N &= a_F (\iota^N + m + \rho). \end{aligned} \tag{29.45}$$

Let us consider the first equation in (29.45) and multiply through by $(1-n^N)$. We get

$$\begin{aligned}\frac{(1/\omega-1)(1-n^N)}{w^N} &= a_m[(1-n^N)\rho+(1-n^N)\iota^S] \\ &= a_m[(1-n^N)\rho+\mu],\end{aligned}$$

since $(1-n^N)\iota^S = n^S\iota^S = \mu$ by the definitions.

We now consider the second and third equations in (29.45), multiply through by n^N and use the definitions, which give $\left(\iota^N + m + \rho\right) n^N = \iota + \rho n^N$. Putting these results together we can rewrite Eqs. (29.45) as

$$(1/\omega-1)(1-n^N)/w^N = a_m[(1-n^N)\rho+\mu], \tag{29.46}$$

$$(1-1/\lambda\omega)n^N/w^N = a_L\left(\iota+\rho n^N\right), \tag{29.47}$$

$$(1-1/\lambda)n^N/w^N = a_F\left(\iota+\rho n^N\right). \tag{29.48}$$

If we divide Eq. (29.47) by Eq. (29.48), we get

$$\frac{1-1/\lambda\omega}{1-1/\lambda} = \frac{a_L}{a_F},$$

hence

$$\frac{\lambda}{\lambda-1} - \frac{1}{(\lambda-1)}\omega^{-1} = \frac{a_L}{a_F},$$

from which

$$\omega^{-1} = (1-\lambda)\frac{a_L}{a_F} + \lambda. \tag{29.49}$$

Since λ, the quality index, is greater than 1, it follows that ω^{-1} is a decreasing linear function of a_L/a_F, which means that the relative wage of North is higher the lower is a_L relative to a_F, i.e., the more productive in R&D are leaders relative to followers.

29.2.5.2 Inefficient Followers

Since no research is carried out by followers, we have $\iota^N = 0$. Then conditions (29.36) and (29.38) imply $n^{NN} = 0$ and $\mu = \iota$, namely South learns to imitate precisely the same number of products that are improved in North per unit of time. The labour market clearing conditions (29.42) and (29.43) respectively simplify to

$$\frac{n^N}{\lambda \omega w^N} + a_L \iota = L^N, \tag{29.50}$$

$$\frac{1-n^N}{w^N} + a_m \iota = L^S. \tag{29.51}$$

As regards the no-arbitrage relations, (29.33) now holds as an inequality since $\iota^N = 0$. Conditions (29.34) and (29.35) hold as equalities and the same procedure followed above for Eqs. (29.46) and (29.47) gives, setting $\mu = \iota$,

$$(1/\omega - 1)(1-n^N)/w^N = a_m[(1-n^N)\rho + \iota], \tag{29.52}$$
$$(1 - 1/\lambda\omega)n^N/w^N = a_L \left(\iota + \rho n^N\right). \tag{29.53}$$

29.2.6 *Comparative Dynamics*

The interest of this model lies in the study of the effects of country size and research subsidies on the steady-state path. This is an exercise in comparative dynamics (see Gandolfo, 2009, chap. 20, sect. 20.6), and amounts to finding the partial derivatives of the steady-state values of the variables with respect to the size parameters and to parameters representing research subsidies.

As regards the former, country size is conveniently expressed by the labour force, hence our size parameters are L^N, L^S.

As regards research subsidies, we introduce parameters ϕ^N, ϕ^S representing the share of research costs subsidized by the government in North and South, respectively. Hence Northern and Southern entrepreneurs pay only the fraction $(1-\phi^N), (1-\phi^S)$ of R&D costs, respectively. Consequently, the no arbitrage conditions must be modified to take account of this lower cost, which can be done by multiplying the r.h.s. of (29.47) and (29.48) by $(1-\phi^N)$, and the r.h.s. of (29.46) by $(1-\phi^S)$.

Since the steady-state conditions are different according as followers are efficient or inefficient, these two cases must be treated separately.

29.2.6.1 Efficient Followers

We rewrite here the steady-state equations (29.42), (29.43), (29.47), and (29.46) as

$$\begin{aligned}
&\varphi_1(\iota,\mu,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\\
&\quad\equiv \frac{n^N(1-\mu/\iota)}{\lambda w^N}+\frac{n^N\mu/\iota}{\lambda\omega w^N}+a_L\mu+a_F(\iota-\mu)-L^N=0,\\
&\varphi_2(\iota,\mu,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\equiv\frac{1-n^N}{w^N}+a_m\mu-L^S=0,\\
&\varphi_3(\iota,\mu,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\\
&\qquad\equiv a_L\left(\iota+\rho n^N\right)(1-\phi^N)-(1-1/\lambda\omega)n^N/w^N=0,\\
&\varphi_4(\iota,\mu,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\\
&\quad\equiv a_m[(1-n^N)\rho+\mu](1-\phi^S)-(1/\omega-1)(1-n^N)/w^N=0.
\end{aligned}\tag{29.54}$$

The practical procedure usually followed is to differentiate these equations totally, and then compute the derivatives we are interested in as ratios of the relevant differentials. It is however more rigorous to proceed as follows. By the implicit function theorem we can express the variables ι,μ,n^N,w^N as differentiable functions of the parameters L^N,L^S,ϕ^N,ϕ^S in a neighbourhood of the equilibrium point provided that the Jacobian matrix of Eqs. (29.54) with respect to ι,μ,n^N,w^N is non-singular at the equilibrium point. We take as initial equilibrium point the steady-state where $\phi^N=\phi^S=0$, so as to determine the effects of the introduction of research subsidies with respect to the no-subsidy equilibrium.

The Jacobian matrix is

$$\mathbf{J}=\begin{bmatrix}
\frac{\partial\varphi_1}{\partial\iota} & \frac{\partial\varphi_1}{\partial\mu} & \frac{\partial\varphi_1}{\partial n^N} & \frac{\partial\varphi_1}{\partial w^N}\\
\frac{\partial\varphi_2}{\partial\iota} & \frac{\partial\varphi_2}{\partial\mu} & \frac{\partial\varphi_2}{\partial n^N} & \frac{\partial\varphi_2}{\partial w^N}\\
\frac{\partial\varphi_3}{\partial\iota} & \frac{\partial\varphi_3}{\partial\mu} & \frac{\partial\varphi_3}{\partial n^N} & \frac{\partial\varphi_3}{\partial w^N}\\
\frac{\partial\varphi_4}{\partial\iota} & \frac{\partial\varphi_4}{\partial\mu} & \frac{\partial\varphi_4}{\partial n^N} & \frac{\partial\varphi_4}{\partial w^N}
\end{bmatrix}.\tag{29.55}$$

Simple calculations yield

$$\frac{\partial\varphi_1}{\partial\iota}=b,\text{ where } b\equiv a_F-\frac{(1/\omega-1)n^N\mu}{\iota^2\lambda w^N}\text{ is assumed to be positive;}$$
$$\frac{\partial\varphi_1}{\partial\mu}=\frac{-n^N/\iota}{\lambda w^N}+\frac{n^N/\iota}{\lambda\omega w^N}+(a_L-a_F)=\frac{n^N}{\lambda w^N}(1/\omega-1)/\iota+(a_L-a_F).$$

Now, if we subtract (29.47) from (29.48) we have $(1-1/\lambda\omega)n^N/w^N-(1-1/\lambda)n^N/w^N=(a_L-a_F)\left(\iota+\rho n^N\right)$. Simple manipulations on the l.h.s. of this last expression yield $\frac{n^N}{\lambda w^N}(1-1/\omega)=(a_L-a_F)\left(\iota+\rho n^N\right)$. This can be substituted in the expression for $\partial\varphi_1/\partial\mu$ found above, thus obtaining

$$\frac{\partial\varphi_1}{\partial\mu}=\frac{-(a_L-a_F)}{\iota}\left(\iota+\rho n^N\right)+(a_L-a_F)=\frac{(a_F-a_L)\rho n^N}{\iota};$$
$$\frac{\partial\varphi_1}{\partial n^N}=\frac{(1-\mu/\iota)}{\lambda w^N}+\frac{\mu/\iota}{\lambda\omega w^N}=\frac{1+(\mu/\iota)(1/\omega-1)}{\lambda w^N}=\frac{\beta}{\lambda w^N},$$

where $\beta \equiv 1 + (\mu/\iota)(1/\omega - 1) > 0$;

$\dfrac{\partial \varphi_1}{\partial w^N} = -\dfrac{n^N(1-\mu/\iota)}{\lambda(w^N)^2} - \dfrac{n^N\mu/\iota}{\lambda\omega(w^N)^2} = -n^N\dfrac{1+(\mu/\iota)(1/\omega-1)}{\lambda(w^N)^2} = \dfrac{-n^N\beta}{\lambda(w^N)^2}$;

$\dfrac{\partial \varphi_2}{\partial \iota} = 0$; $\dfrac{\partial \varphi_2}{\partial \mu} = a_m$; $\dfrac{\partial \varphi_2}{\partial n^N} = \dfrac{-1}{w^N}$;

$\dfrac{\partial \varphi_2}{\partial w^N} = -(w^N)^{-2}(1-n^N) = -(w^N)^{-2}(-n^S)$;

$\dfrac{\partial \varphi_3}{\partial \iota} = a_L(1-\phi^N) = a_L$ (since we are evaluating the Jacobian at $\phi^N = 0$);

$\dfrac{\partial \varphi_3}{\partial \mu} = 0$;

$\dfrac{\partial \varphi_3}{\partial n^N} = a_L\rho - (1-1/\lambda\omega)/w^N$. Now from Eq. (29.47) we easily obtain

$-(1-1/\lambda\omega)/w^N = -a_L\rho - \dfrac{a_L\iota}{n^N}$, hence $\dfrac{\partial \varphi_3}{\partial n^N} = -\dfrac{a_L\iota}{n^N}$;

$\dfrac{\partial \varphi_3}{\partial w^N} = (w^N)^{-2}(1-1/\lambda\omega)n^N$;

$\dfrac{\partial \varphi_4}{\partial \iota} = 0$; $\dfrac{\partial \varphi_4}{\partial \mu} = a_m$; $\dfrac{\partial \varphi_4}{\partial n^N} = -a_m\rho + \dfrac{1/\omega - 1}{w^N}$. From Eq. (29.46) we get

$\dfrac{1/\omega - 1}{w^N} = a_m\rho + a_m\mu/(1-n^N) = a_m\rho + a_m\mu/n^S$ since $1 - n^N = n^S$.

Hence $\dfrac{\partial \varphi_4}{\partial n^N} = a_m\mu/n^S$;

$\dfrac{\partial \varphi_4}{\partial w^N} = (w^N)^{-2}(1/\omega - 1)n^S$.

These expressions can be substituted into (29.55) to obtain the explicit expression for the Jacobian. Using the property that multiplying one line (in our case the last column) of a determinant by the same constant is the same as multiplying the whole determinant by that constant, we can write the determinant of our Jacobian matrix as $|\mathbf{J}| = (w^N)^{-2}\Delta$ where

$$\Delta \equiv \begin{vmatrix} b & \dfrac{(a_F - a_L)\rho n^N}{\iota} & \dfrac{\beta}{\lambda w^N} & -\dfrac{n^N\beta}{\lambda} \\ 0 & a_m & -\dfrac{1}{w^N} & -n^S \\ a_L & 0 & -\dfrac{a_L\iota}{n^N} & \left(1-\dfrac{1}{\lambda\omega}\right)n^N \\ 0 & a_m & \dfrac{a_m\mu}{n^S} & (\dfrac{1}{\omega}-1)n^S \end{vmatrix}. \tag{29.56}$$

Expanding the determinant by the first column we obtain

$$
\begin{aligned}
\Delta = -ba_m \left[\frac{n^S a_L \iota}{\omega n^N} + (1 - 1/\lambda\omega) \frac{n^N a_m \mu}{n^S} + (1 - 1/\lambda\omega) \right] \\
-a_L \frac{(a_F - a_L)\rho n^N}{\iota} \left[\frac{n^S}{(1 - 1/\omega) w^N} + a_m \mu \right] \\
-a_L a_m \left[\frac{\beta n^S}{\lambda w^N (1/\omega - 1)} + \frac{a_m \mu n^N \beta}{n^S \lambda} \right] \\
-a_L a_m \left[\frac{\beta n^S}{\lambda w^N} + \frac{n^N \beta}{\lambda w^N} \right].
\end{aligned} \tag{29.57}
$$

If we remember that $\omega < 1$ (hence $(1/\omega - 1) > 0$), that $\lambda w^S > w^N$ (hence $\lambda\omega > 1$ and $(1 - 1/\lambda\omega) > 0$), and that $a_F - a_L > 0$, it is easy to check that all expressions in square brackets are positive and are multiplied by negative expressions, hence $\Delta < 0$.

We can now compute the partial derivatives of the variables with respect to the parameters (i.e., $\partial\iota/\partial L^N$, $\partial\iota/\partial L^S$, $\partial\iota/\partial\phi^N$, $\partial\iota/\partial\phi^S$, etc.) by applying the chain rule to Eqs. (29.54). We only calculate a subset of these 16 partial derivatives, beginning with the partial derivatives with respect to L^N. We have

$$
\begin{aligned}
\frac{\partial\varphi_1}{\partial\iota}\frac{\partial\iota}{\partial L^N} + \frac{\partial\varphi_1}{\partial\mu}\frac{\partial\mu}{\partial L^N} + \frac{\partial\varphi_1}{\partial n^N}\frac{\partial n^N}{\partial L^N} + \frac{\partial\varphi_1}{\partial w^N}\frac{\partial w^N}{\partial L^N} - \frac{\partial\varphi_1}{\partial L^N} = 0, \\
\frac{\partial\varphi_2}{\partial\iota}\frac{\partial\iota}{\partial L^N} + \frac{\partial\varphi_2}{\partial\mu}\frac{\partial\mu}{\partial L^N} + \frac{\partial\varphi_2}{\partial n^N}\frac{\partial n^N}{\partial L^N} + \frac{\partial\varphi_2}{\partial w^N}\frac{\partial w^N}{\partial L^N} - \frac{\partial\varphi_2}{\partial L^N} = 0, \\
\frac{\partial\varphi_3}{\partial\iota}\frac{\partial\iota}{\partial L^N} + \frac{\partial\varphi_3}{\partial\mu}\frac{\partial\mu}{\partial L^N} + \frac{\partial\varphi_3}{\partial n^N}\frac{\partial n^N}{\partial L^N} + \frac{\partial\varphi_3}{\partial w^N}\frac{\partial w^N}{\partial L^N} - \frac{\partial\varphi_3}{\partial L^N} = 0, \\
\frac{\partial\varphi_4}{\partial\iota}\frac{\partial\iota}{\partial L^N} + \frac{\partial\varphi_4}{\partial\mu}\frac{\partial\mu}{\partial L^N} + \frac{\partial\varphi_4}{\partial n^N}\frac{\partial n^N}{\partial L^N} + \frac{\partial\varphi_4}{\partial w^N}\frac{\partial w^N}{\partial L^N} - \frac{\partial\varphi_4}{\partial L^N} = 0,
\end{aligned} \tag{29.58}
$$

or

$$
\mathbf{J}\mathbf{v}_{L^N} = \boldsymbol{\kappa}_{L^N}, \tag{29.59}
$$

where $\mathbf{J}$ is the Jacobian matrix found above, $\mathbf{v}_{L^N} \equiv \left\{ \frac{\partial\iota}{\partial L^N} \; \frac{\partial\mu}{\partial L^N} \; \frac{\partial n^N}{\partial L^N} \; \frac{\partial w^N}{\partial L^N} \right\}$ is the column vector of the partial derivatives of the endogenous variables with respect to the parameter L^N, and $\boldsymbol{\kappa}_{L^N} \equiv \left\{ \frac{\partial\varphi_1}{\partial L^N} \; \frac{\partial\varphi_2}{\partial L^N} \; \frac{\partial\varphi_3}{\partial L^N} \; \frac{\partial\varphi_4}{\partial L^N} \right\} = \left\{ 1 \; 0 \; 0 \; 0 \right\}$ is the vector of known terms.

The solution to the linear system (29.59) exists, since $\mathbf{J}$ is non-singular, and is

$$
\mathbf{v}_{L^N} = \mathbf{J}^{-1}\boldsymbol{\kappa}_{L^N}. \tag{29.60}
$$

Similarly we have

$$\begin{aligned} \mathbf{v}_{L^S} &= \mathbf{J}^{-1}\boldsymbol{\kappa}_{L^S}, \\ \mathbf{v}_{\phi^N} &= \mathbf{J}^{-1}\boldsymbol{\kappa}_{\phi^N}, \\ \mathbf{v}_{\phi^S} &= \mathbf{J}^{-1}\boldsymbol{\kappa}_{\phi^S}. \end{aligned} \tag{29.61}$$

Using for example Cramer's rule we have

$$\begin{aligned} \frac{\partial \iota}{\partial L^N} &= \frac{\begin{vmatrix} 1 & \frac{(a_F - a_L)\rho n^N}{\iota} & \frac{\beta}{\lambda w^N} & -\frac{n^N\beta}{\lambda} \\ 0 & a_m & -\frac{1}{w^N} & -n^S \\ 0 & 0 & -\frac{a_L\iota}{n^N} & \left(1 - \frac{1}{\lambda\omega}\right)n^N \\ 0 & a_m & \frac{a_m\mu}{n^S} & (\frac{1}{\omega} - 1)n^S \end{vmatrix}}{\Delta} \\ &= -\frac{a_m}{\Delta}\left[\frac{a_L\iota}{\omega}\frac{n^S}{n^N} + \left(1 - \frac{1}{\lambda\omega}\right)n^N\left(\frac{1}{w^N} + \frac{a_m\mu}{n^S}\right)\right] > 0, \end{aligned} \tag{29.62}$$

where the sign is obvious if we observe that the expression in square brackets is positive given that $\lambda\omega > 1$.

It would be tedious to repeat this step-by-step procedure for all the other partial derivatives, hence we only state the results (that can easily be obtained as shown in detail for $\partial\iota/\partial L^N$), with the explanation of some less obvious transformations. Thus we have

$$\begin{aligned} \frac{\partial \iota}{\partial L^S} = \frac{\rho n^N}{\Delta}\Bigg\{&(a_F - a_L)\left[a_L\frac{n^S}{n^N}(1/\omega - 1) + (1 - 1/\lambda\omega)\frac{a_m\mu}{\iota}\frac{n^N}{n^S}\right] \\ &-\frac{a_m}{\rho}\left[(1 - 1/\lambda\omega)\beta/\lambda w^N - \beta a_L\iota/\lambda n^N\right]\Bigg\}, \end{aligned}$$

hence

$$\frac{\partial \iota}{\partial L^S} = \frac{\rho n^N}{\Delta}\left\{(a_F - a_L)\left[a_L\frac{n^S}{n^N}(\frac{1}{\omega} - 1) + (1 - \frac{1}{\lambda\omega})\frac{a_m\mu}{\iota}\frac{n^N}{n^S}\right] - a_m\frac{a_L\beta}{\lambda}\right\} \gtrless 0, \tag{29.63}$$

where we have used the transformation

$$(1 - 1/\lambda\omega)/\lambda w^N - a_L\iota/\lambda n^N = a_L\rho/\lambda$$

that derives from (29.47).

Next we have

$$\begin{aligned}\frac{\partial \iota}{\partial \phi^N} &= \frac{a_L(\iota+\rho n^N)}{\Delta}\left\{-a_m\left[\frac{\beta}{\lambda w^N}(\frac{1}{\omega}-1)n^S+\frac{n^N\beta}{\lambda}\frac{a_m\mu}{n^S}+\frac{\beta}{\lambda w^N}n^S\right.\right.\\ &\left.\left.+\frac{1}{w^N}\frac{n^N\beta}{\lambda}\right]+\frac{(a_F-a_L)\rho n^N}{\iota}\left[-\frac{1}{w^N}(\frac{1}{\omega}-1)n^S+a_m\mu\right]\right\},\end{aligned}$$

hence

$$\frac{\partial \iota}{\partial \phi^N} = \frac{a_m b_N}{\Delta}\left\{\frac{\beta}{\lambda}\left[\frac{n^S}{w^N}(1-\frac{1}{\omega})-a_m\mu\frac{n^N}{n^S}-\frac{n^N}{w^N}\right]+(a_L-a_F)n^N n^S\frac{\rho^2}{\iota}\right\}>0, \tag{29.64}$$

where $b_N \equiv a_L(\iota+\rho n^N)$, and where we have used the transformation

$$-\frac{1}{w^N}(\frac{1}{\omega}-1)(1-n^N)+a_m\mu = -a_m\rho(1-n^N),$$

that derives from (29.46).

The expression for $\partial\iota/\partial\phi^N$ is clearly positive because the expression in braces is negative given that $\omega<1$ and $a_L<a_F$. We now compute

$$\begin{aligned}\frac{\partial \iota}{\partial \phi^S} &= -\frac{a_m[\rho(1-n^N)+\mu]}{\Delta}\left\{\frac{(a_F-a_L)\rho n^N}{\iota}\left[-\frac{1}{w^N}\left(1-\frac{1}{\lambda\omega}\right)n^N-\frac{a_L\iota}{n^N}n^S\right]\right.\\ &\left.-a_m\left[\frac{\beta}{\lambda w^N}\left(1-\frac{1}{\lambda\omega}\right)n^N-\frac{a_L\iota\beta}{\lambda}\right]\right\}\\ &= \frac{\rho b_S}{\Delta}\left\{(a_F-a_L)\left[\frac{(n^N)^2}{\iota w^N}\left(1-\frac{1}{\lambda\omega}\right)+a_L n^S\right]+\frac{\beta}{\lambda}a_m a_L n^N\right\}<0,\end{aligned} \tag{29.65}$$

where we have used the transformation

$$\frac{1}{w^N}\left(1-\frac{1}{\lambda\omega}\right)n^N - a_L\iota = a_L\rho n^N$$

that derives from (29.47), and $b_S \equiv a_m[\rho(1-n^N)+\mu]$.

We next have

$$\frac{\partial\mu}{\partial L^N} = \frac{1}{\Delta}a_L\left[-(1/w^N)(1/\omega-1)n^S+a_m\mu\right].$$

Now, $(1/w^N)(1/\omega-1) = a_m\rho + a_m\iota/n^S$ by (29.46), hence

$$\frac{\partial\mu}{\partial L^N} = -\frac{1}{\Delta}a_L a_m\left[\rho n^S+(\iota-\mu)\right]>0, \tag{29.66}$$

given that $\iota-\mu>0$.

We then calculate

$$\begin{aligned}\frac{\partial\mu}{\partial L^S} &= \frac{1}{\Delta}\left\{b\left[(a_L\iota/n^N)(1-1/\omega)n^S-(a_m\mu/n^S)(1-1/\lambda\omega)n^N\right]\right.\\ &\qquad \left.-a_L\left[(\beta/\lambda w^N)(1/\omega-1)n^S+(a_m\mu/n^S)\beta n^N/\lambda\right]\right\}\\ &= \left\{-\frac{1}{\Delta}\frac{a_m\mu n^N}{n^S}\left[\frac{a_L\beta}{\lambda}+b\left(1-\frac{1}{\lambda\omega}\right)\right]\right.\\ &\qquad \left.+\left(\frac{1}{\omega}-1\right)n^S\left[\frac{a_L\beta}{\lambda w^N}+\frac{ba_L\iota}{n^N}\right]\right\} \end{aligned}\tag{29.67}$$

which shows that $\partial\mu/\partial L^S > 0$.

The effect of ϕ^N on μ is given by

$$\begin{aligned}\frac{\partial\mu}{\partial\phi^N} &= \frac{a_L(\iota+\rho n^N)b\left[\frac{(1/\omega-1)n^S}{w^N}-a_m\mu\right]}{\Delta}\\ &= \frac{1}{\Delta}(b_N b a_m \rho n^S) < 0,\end{aligned}\tag{29.68}$$

where we have used the transformation

$$\frac{(1/\omega-1)n^S}{w^N}-a_m\mu = a_m\rho n^S,$$

that derives from (29.46).

The effect of ϕ^S on μ is given by

$$\begin{aligned}\frac{\partial\mu}{\partial\phi^S} &= \frac{b_S}{\Delta}\left\{b\left[-(1-1/\lambda\omega)n^N/w^N-a_L\iota n^S/n^N\right]\right.\\ &\qquad \left.+a_L\left[-\beta n^S/\lambda w^N-\beta n^N/\lambda w^N\right]\right\}\\ &= -\frac{b_S}{\Delta}\left\{b\left[\left(1-\frac{1}{\lambda\omega}\right)\frac{n^N}{w^N}+\frac{a_L\iota n^S}{n^N}\right]+\frac{a_L\beta}{\lambda w^N}\right\} > 0,\end{aligned}\tag{29.69}$$

where we have used the fact that $n^S+n^N=1$.

We finally compute

$$\frac{\partial n^N}{\partial L^S} = \frac{n^N}{\Delta}\left[ba_m\left(1-\frac{1}{\lambda\omega}\right)+a_L\left(\frac{1}{\omega}-1\right)n^S(a_F-a_L)\frac{\rho}{\iota}+a_La_m\frac{\beta}{\lambda}\right] < 0. \tag{29.70}$$

The signs of the various partial derivatives give us the comparative dynamics results we are looking for. Strategic trade policy is a particularly interesting case.

Strategic Trade Policy Let us consider, for example, subsidies to innovation and imitation.

The introduction of a small research subsidy in North enhances R&D for both leaders and followers, hence the aggregate rate of innovation increases ($\partial\iota/\partial\phi^N > 0$). At the same time, the rate of imitation in South is adversely affected ($\partial\mu/\partial\phi^N < 0$). This is due to the following chain of effects. As North devotes more resources to R&D, less resources are devoted to the Northern manufacturing sector, which implies that the number of products manufactured in North declines while those manufactured in South increase. Thus the manufacturing sector expands in South and less resources are devoted there to research.

The effects of the introduction of a small subsidy to imitative research in South are exactly the opposite: the rate of imitation in South is favourably affected ($\partial\mu/\partial\phi^S > 0$) while the rate of innovation in North is adversely affected ($\partial\iota/\partial\phi^S < 0$).

These results point to a possible conflict between the two governments in their efforts to promote domestic growth via R&D, a result in contrast with a previous finding by the same authors in the context of another model (Grossman & Helpman, 1991a , chap. 11), where a subsidy to research by either government is favourable to technological progress in both regions.

We now examine the effects of size, as represented by L^N, L^S. An increase in the size of South increases both the rate of technological progress there ($\partial\mu/\partial L^S > 0$) and the intensity of imitation (i.e., $m = \mu/n^N$) increases, as can be seen from the fact that as μ increases, n^N decreases $\partial n^N/\partial L^S < 0$). The effect on the rate of technological progress in North is however uncertain ($\partial\iota/\partial L^S \gtrless 0$), while an increase in the size of North has a favourable effect on both the Northern and Southern rates of technological progress ($\partial\iota/\partial L^N > 0, \partial\mu/\partial L^N > 0$).

29.2.6.2 Inefficient Followers

In the case of inefficient followers, the steady-state equations are (29.50), (29.51), (29.53), and (29.52). In this case, since μ disappears from the equations, we take as endogenous variable the relative wage rate ω. Thus we have the system of implicit functions

$$
\begin{aligned}
&\gamma_1(\iota,\omega,n^N,w^N;L^N,L^S,\phi^N,\phi^S) \equiv \frac{n^N}{\lambda\omega w^N} + a_L\iota - L^N &= 0,\\
&\gamma_2(\iota,\omega,n^N,w^N;L^N,L^S,\phi^N,\phi^S) \equiv \frac{1-n^N}{w^N} + a_m\iota - L^S &= 0,\\
&\gamma_3(\iota,\omega,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\\
&\qquad \equiv -a_L\left(\iota+\rho n^N\right)(1-\phi^N) + (1-1/\lambda\omega)n^N/w^N &= 0,\\
&\gamma_4(\iota,\omega,n^N,w^N;L^N,L^S,\phi^N,\phi^S)\\
&\equiv -a_m[(1-n^N)\rho+\iota](1-\phi^S) + (1/\omega-1)(1-n^N)/w^N &= 0.
\end{aligned}
\tag{29.71}
$$

The partial derivatives making up the Jacobian evaluated at $\phi^N = \phi^S = 0$ are

$$\frac{\partial \gamma_1}{\partial \iota} = a_L; \frac{\partial \gamma_1}{\partial \omega} = -\frac{n^N}{\lambda w^N}\omega^{-2}; \frac{\partial \gamma_1}{\partial n^N} = \frac{1}{\lambda \omega w^N}; \frac{\partial \gamma_1}{\partial w^N} = -\frac{n^N}{\lambda \omega}\left(w^N\right)^{-2};$$
$$\frac{\partial \gamma_2}{\partial \iota} = a_m; \frac{\partial \gamma_2}{\partial \omega} = 0; \frac{\partial \gamma_2}{\partial n^N} = \frac{-1}{w^N}; \frac{\partial \gamma_2}{\partial w^N} = -n^S\left(w^N\right)^{-2};$$
$$\frac{\partial \gamma_3}{\partial \iota} = -a_L; \frac{\partial \gamma_3}{\partial \omega} = \frac{n^N}{\lambda w^N}\omega^{-2};$$
$$\frac{\partial \gamma_3}{\partial n^N} = -a_L\rho + (1 - 1/\lambda\omega)/w^N = \frac{a_L\iota}{n^N} \text{ using Eq. (29.53)};$$
$$\frac{\partial \gamma_3}{\partial w^N} = (1/\lambda\omega - 1)n^N\left(w^N\right)^{-2};$$
$$\frac{\partial \gamma_4}{\partial \iota} = -a_m; \frac{\partial \gamma_4}{\partial \omega} = -\frac{n^S}{w^N}\omega^{-2};$$
$$\frac{\partial \gamma_4}{\partial n^N} = \rho - (1/\omega - 1)/w^N = -\frac{a_m\iota}{n^S} \text{ using Eq. (29.52)};$$
$$\frac{\partial \gamma_4}{\partial w^N} = (1 - 1/\omega)n^S\left(w^N\right)^{-2}.$$

Using the property that multiplying one line (in our case the third column and then the fourth) of a determinant by the same constant is the same as multiplying the whole determinant by that constant. we can write the determinant of our Jacobian matrix as $|\mathbf{J}| = \omega^{-2}(w^N)^{-2}\tilde{\Delta}$, where

$$\tilde{\Delta} \equiv \begin{vmatrix} a_L & -\dfrac{n^N}{\lambda w^N} & \dfrac{1}{\lambda \omega w^N} & -\dfrac{n^N}{\lambda\omega} \\ a_m & 0 & -\dfrac{1}{w^N} & -n^S \\ -a_L & \dfrac{n^N}{\lambda w^N} & \dfrac{a_L\iota}{n^N} & \left(\dfrac{1}{\lambda\omega} - 1\right)n^N \\ -a_m & -\dfrac{n^S}{w^N} & -\dfrac{a_m\iota}{n^S} & \left(1 - \dfrac{1}{\omega}\right)n^S \end{vmatrix}. \tag{29.72}$$

If we expand $\tilde{\Delta}$ by the second row and then expand each of the resulting third-order determinants by the first row, we get, after simple manipulations (to obtain the expression in the third square brackets use the transformation $n^N(1/\lambda\omega - 1)/w^N + a_L\iota = -a_L\rho n^N$, that derives from Eq. (29.53))

$$\begin{aligned}\tilde{\Delta} &= \frac{a_m n^N}{\lambda w^N}\left[\frac{a_L\iota}{n^N}\left(1 - \frac{1}{\omega}\right)n^S + \left(\frac{1}{\lambda\omega} - 1\right)\frac{n^N a_m\iota}{n^S}\right] \\ &+ \frac{a_m n^N}{\lambda w^N}\left[\frac{1}{\lambda\omega w^N}(1 - \frac{1}{\omega})n^S - \frac{n^N a_m\iota}{n^S}\right]\end{aligned}$$

$$
\begin{aligned}
&+a_m \frac{n^S}{\lambda\omega w^N}\left[-a_L \rho n^N\right] \\
&+\frac{a_L}{w^N}\left[\frac{n^N n^S}{w^N}\left(\frac{1}{\lambda\omega}-1\right)\right] \\
&+\frac{a_L}{w^N}\left[-\frac{n^N n^S}{\lambda w^N}\right] \\
&+\frac{a_m}{w^N}\left[-\frac{n^N}{\lambda w^N}\right] \\
&+n^S a_L\left[-\frac{n^S a_L \iota}{w^N n^N}-\frac{n^S}{w^N \lambda\omega w^N}\right] \\
&+n^S a_m\left[-\frac{a_L \iota}{\lambda w^N}-\frac{n^N}{\lambda w^N \lambda\omega w^N}\right].
\end{aligned} \tag{29.73}
$$

If we recall that $\omega < 1, \lambda > 1, \lambda\omega > 1$, it is easy to see that all expressions in square brackets are negative, hence $\tilde{\Delta} < 0$.

We now proceed to calculate a subset of the partial derivatives of the endogenous variables with respect to the parameters by the same procedure whose details we have illustrated in the case of efficient followers. As before, all these derivatives are calculated at the point where $\phi^N = \phi^S = 0$.

Let us begin with the effects of size. We have

$$
\frac{\partial n^N}{\partial L^N} = \frac{1}{\tilde{\Delta}}\left\{a_m\left[\frac{n^N n^S}{w^N}\left(\frac{1}{\lambda}-1\right)\right] - n^S\left[a_L \frac{n^S}{w^N} + a_m \frac{n^N}{\lambda w^N}\right]\right\} > 0 \tag{29.74}
$$

and

$$
\frac{\partial n^N}{\partial L^S} = \frac{1}{\tilde{\Delta}}\left[a_L \frac{n^S n^N}{w^N}\left(1+\frac{1}{\lambda}\right) + a_m \frac{\left(n^N\right)^2}{\lambda w^N}\right] < 0, \tag{29.75}
$$

which (together with the identity $n^N + n^S = 1$) show that an increase in the size of a country has a favourable effect on the number of products manufactured in that country and, of course, the opposite effect on the other country.

Finally, nothing can be said on the effect of size on the relative wage rate, since the sign of the relevant partial derivatives is ambiguous.

Strategic Trade Policy We now turn to the effects of the introduction of research subsidies. We have

$$
\frac{\partial n^N}{\partial \phi^N} = \frac{a_L \left(n^S\right)^2 b_N}{w^N \tilde{\Delta}} < 0, \tag{29.76}
$$

$$\frac{\partial n^N}{\partial \phi^S} = -\frac{a_m \left(n^N\right)^2 b_S}{\lambda w^N \tilde{\Delta}} > 0, \tag{29.77}$$

$$\frac{\partial \omega}{\partial \phi^N} = \frac{\omega^2 a_m b_N}{\tilde{\Delta}} \left\{ \left[\frac{a_L}{a_m} (\frac{1}{\omega} - 1) \frac{n^S}{w^N} - a_L \iota \right] + \frac{n^S}{\lambda \omega^2 w^N} + \frac{n^N}{\lambda \omega} \frac{a_m \iota}{n^S} + \frac{n^N}{\lambda \omega w^N} \right\}.$$

Using (29.52) and (29.51) we get

$$\frac{\partial \omega}{\partial \phi^N} = \frac{\omega^2 a_m b_N \left(\rho a_L n^S + n^S/\lambda \omega^2 w^N + L^S n^N/\lambda \omega n^S\right)}{\tilde{\Delta}} < 0. \tag{29.78}$$

$$\frac{\partial \omega}{\partial \phi^S} = \frac{a_L b_S}{\omega^{-2} \tilde{\Delta}} \left\{ \left[\frac{n^N}{w^N} + \frac{a_L \iota}{n^N} n^S \right] + \left[\frac{a_m}{a_L} \frac{1}{\lambda \omega} \left(1 - \frac{1}{\lambda \omega}\right) \frac{n^N}{w^N} - \frac{a_m \iota}{\lambda \omega} \right] + \frac{n^S}{\lambda \omega w^N} \right\},$$

hence

$$\frac{\partial \omega}{\partial \phi^S} = -\frac{\omega^2 a_L b_S \left(\rho a_m n^N/\lambda \omega + n^N/w^N + L^N n^S/n^N\right)}{\tilde{\Delta}} > 0, \tag{29.79}$$

where we have used (29.53) and (29.50).

The signs of the partial derivatives can be interpreted in the following way. The introduction of a subsidy to R&D in North causes the industries there to devote more resources to research, hence less resources are employed in manufacturing. The number of products manufactured in North thus decreases ($\partial n^N/\partial \phi^N < 0$) while the relative wage of North increases ($\partial \omega/\partial \phi^N < 0$). A subsidy to imitative research in South has exactly the opposite result on product shares and the relative wage.

For an extension of this model see Lai (1995).

References

Aghion, P., & Howitt, P. (1990, 1992). A model of growth through creative destruction.
Findlay, R. (1995). *Factor proportions, trade, and growth* (chap. 3).
Gandolfo, G. *Economic dynamics* (chaps. 20 and 21, sect. 21.3.2; chap. 27, sect. 27.2.2).
Grossman, G. M., & Helpman, E. (1991a). *Innovation and growth in the global economy*.
Grossman, G. M., & Helpman, E. (1991b). Endogenous product cycles.

Grossman, G. M., & Helpman, E. (1991c). Quality ladders and product cycles.
Hicks, J. R. (1939, 1946). *Value and capital*. Mathematical appendix (sect. 10).
Kamien, M. I., & Schwartz, N. L. (1991). *Dynamic optimization* (chap. 7).
Lai, E. L. C. (1995). The product cycle and the world distribution of income: A reformulation.
Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*.

Chapter 30
Appendix to Chapter 16

Since the very beginning the literature on the new economic geography has made wide use of numerical methods to explore the models. In this appendix we provide two simple examples. In the first example we use numerical methods to draw the phase line which then allows performing the traditional topological analysis of global stability. The second example deals with the relationship between theory and empirics and consists in comparing the (numerical) solution values obtained from the model with what is observed in (or estimated from) the data.

30.1 Numerical Phase Lines

Consider the Core-Periphery model studied in Sect. 16.3. After operating the substitutions discussed below Eqs. (16.4) and (16.5) these equations become

$$w_1 = \frac{(w_1)^{\frac{1}{1-\mu}}\left[\frac{(1-\gamma)}{2} + \lambda_1 \gamma w_1\right]}{\lambda_1 (w_1)^{\frac{1}{1-\mu}} + (1-\lambda_1)\left(\frac{w_2}{\tau}\right)^{\frac{1}{1-\mu}}} + \frac{\left(\frac{w_1}{\tau}\right)^{\frac{1}{1-\mu}}\left[\frac{(1-\gamma)}{2} + (1-\lambda_1)\gamma w_2\right]}{\lambda_1 \left(\frac{w_1}{\tau}\right)^{\frac{1}{1-\mu}} + (1-\lambda_1)(w_2)^{\frac{1}{1-\mu}}} \tag{30.1}$$

$$w_2 = \frac{\left(\frac{w_2}{\tau}\right)^{\frac{1}{1-\mu}}\left[\frac{(1-\gamma)}{2} + \lambda_1 \gamma w_1\right]}{\lambda_1 (w_1)^{\frac{1}{1-\mu}} + (1-\lambda_1)\left(\frac{w_2}{\tau}\right)^{\frac{1}{1-\mu}}} + \frac{(w_2)^{\frac{1}{1-\mu}}\left[\frac{(1-\gamma)}{2} + (1-\lambda_1)\gamma w_2\right]}{\lambda_1 \left(\frac{w_1}{\tau}\right)^{\frac{1}{1-\mu}} + (1-\lambda_1)(w_2)^{\frac{1}{1-\mu}}}. \tag{30.2}$$

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_30,

Table 30.1 Numerical solutions of the core-periphery model

$\mu = 3/2,\quad \tau = 1/10,\quad \gamma = 4/10$		
λ_1	Nominal wages	Real wage differential
$\lambda_1 = 0.1$	$w_1 = 2.341826921, w_2 = 0.8509081199$	$\overset{\bullet}{\lambda}_1 = 0.2777197895$
$\lambda_1 = 0.2$	$w_1 = 1.838861373, w_2 = 0.7902846567$	$\overset{\bullet}{\lambda}_1 = 0.2558537633$
$\lambda_1 = 0.3$	$w_1 = 1.474281847, w_2 = 0.7967363511$	$\overset{\bullet}{\lambda}_1 = 0.1948332076$
$\lambda_1 = 0.4$	$w_1 = 1.200024688, w_2 = 0.8666502081$	$\overset{\bullet}{\lambda}_1 = 0.1049309257$
$\lambda_1 = 0.5$	$w_1 = 1.000000000, w_2 = 1.000000000$	$\overset{\bullet}{\lambda}_1 = 0$

Substituting for the price index into expression (16.7) the law of motion (16.6) becomes

$$\overset{\bullet}{\lambda}_1 = \frac{w_1}{\left[\lambda_1 (w_1)^{\frac{1}{1-\mu}} + (1-\lambda_1)\left(\frac{w_2}{\tau}\right)^{\frac{1}{1-\mu}}\right]^{\gamma(1-\mu)}} - \frac{w_2}{\left[\lambda_1 \left(\frac{w_1}{\tau}\right)^{\frac{1}{1-\mu}} + (1-\lambda_1)(w_2)^{\frac{1}{1-\mu}}\right]^{\gamma(1-\mu)}}. \tag{30.3}$$

The system is at rest when real wages equalize. The system is also at rest when the real wage is higher in 1 and $\lambda_1 = 1$ or when the real wage is larger in 2 and $\lambda_1 = 0$; these two cases represent the "core-periphery situations". After assigning values to the three parameters τ, μ, γ and to the state variable λ_1, Eqs. (30.1) and (30.2) can be solved numerically for w_1 and w_2. These numerical solutions may be obtained by use of any mathematical software. Having solved for nominal wages we can compute the real wage differential. Note that the so obtained real wage differential depends on the value assigned to λ_1. Assigning a series of values to λ_1 gives the corresponding series of solution values for the real wage differential. This means that we can associate to any value of λ_1 a corresponding value of $\overset{\bullet}{\lambda}_1$, i.e., we can constructed a (numerically computed) "dotted" phase line. The line is "dotted" in the obvious sense that we can compute only a finite number (not a continuum) of correspondences between λ_1 and $\overset{\bullet}{\lambda}_1$.

As an example we show the solutions obtained by assigning $\mu = 3/2$, $\tau = 1/10$, and $\gamma = 4/10$ for each of the following nine values of λ_1: $\lambda_1 = 0.1,\ 0.2,\ \ldots,\ 0.9$. The solutions are obtained by use of any mathematical software. The solutions are shown in Table 30.1. Given the symmetric structure of the model, the solutions values of nominal wages for $\lambda_1 = 0.6, \ldots, 0.9$ are symmetric around 1 to those for $\lambda_1 = 0.1, \ldots, 0.4$. Likewise, the solution values of the real wage differential for $\lambda_1 = 0.6, \ldots, 0.9$ are symmetric around zero to those for $\lambda_1 = 0.1, \ldots, 0.4$. We recall from Eq. (30.3) that $\overset{\bullet}{\lambda}_1$is equal to the real wage differential. Then, by plotting the values for the real wage differential in the space $\lambda_1, \overset{\bullet}{\lambda}_1$ and then connecting the points we obtain a numerical approximation of the phase line. The simplest way of connecting any two consecutive points is by a straight line (linear interpolation). Of course, a variety of more precise methods exists (see Judd, 1998). Once the numerical approximation of the phase line is drawn we can study the dynamic properties of the model by use of standard topological analysis (see Gandolfo, 2009, sect. 21.3).

30.2 Calibration

Calibration is one way to help assessing the empirical validity of theoretical models. The idea is to find the range of reasonable parameter values and the right modeling structure that replicate the empirical observations. As an example of calibration we use the simple exercise in Brülhart, Carrère, and Trionfetti (2012), from which we draw this subsection. The example shows how the empirical analysis offers guidance for the calibration and, more interestingly, for the right choice of the model.

Their objective is to study how the geographical distribution of economic activity within Austria has adjusted to trade opening with Eastern Europe. The model initially used is a three-region extension of the housing congestion model Sect. 16.4.1. In this extension Austria is composed by two regions between which the economic activity is endogenously allocated. The third region is the Eastern Europe. Trade opening occurs between Austria and Eastern Europe and triggers an endogenous reallocation of economic activity towards the eastern region of Austria since these regions are geographically closer to the new source of demand, Eastern Europe. Interestingly, this model does not replicate the data very well. Specifically the model predicts too high labour mobility with respect to what is observed in the data. Guided by this observation, the model was extended to include a migration inertia modeled along the lines of Tabuchi and Thisse (2002) and Murata (2003). This new model extension replicates the data with much greater precision.

30.2.1 The Model

The world economy consists of two countries and three regions. Regions I and B belong to country A and the third region constitutes a one-region country named R. Regions names are obvious mnemonics for the I(nterior) region and the Eastern B(order) region of A(ustria) and the third region represents the R(est of the world) with which trade opening takes place. There are iceberg trade costs: for a unit of good sent from region i to region j only a fraction $\tau_{ij} \in (0, 1)$ arrives in j. It is assumed that $\tau_{ij} = \tau_{ji}\ \forall\ i, j$ and $\tau_{ii} = 1,\ \forall\ i$. The geographical structure of the three-region model is represented by the following assumptions on trade costs:

$$\tau_{IR} = \tau_{IB}\tau_{BR},$$

which means that for a variety of the M-good to be transported between I and R it has to transit through B. Thus, the B(order) is nearer to R than the I(nterior) region and according to theory the increase in demand coming from R after trade opening should cause a geographical reallocation of economic activity from I to B. The task of the exercise, however, is not only to verify that the data confirms this theoretical prediction. The task is also and principally to replicate in the model the magnitude of the observed geographical reallocation of economic activity.

Continuing with the model description, labour is mobile within countries but immobile between countries. Individuals derive utility from consumption of goods and also from the pleasure of residing in a region. The component of utility that is associated with consumption is a Cobb-Douglas defined over M and H with expenditure share on M equal to γ. Demand for any domestic and any imported variety of good M are, respectively, $q_{ii}^d = (p_{ii})^{1-\sigma} (P_i)^{\sigma-1} \gamma E_i$ and $q_{ji}^d = \left(p_{ji}\right)^{1-\sigma} (P_i)^{\sigma-1} \gamma E_i$ where the first subscript refers to the region where the variety is produced and the second subscript refers to the region where the variety is consumed. Total indirect utility, V_i^k, is given by the sum of the real wage, ω_i, which represents the indirect utility derived from consumption and is common to all individuals in a given region, and utility derived from the idiosyncratic appreciation that each individual k associates with region i, ξ_i^k; that is:

$$V_i^k = \omega_i + \xi_i^k.$$

The letter ξ_i^k denotes a random variable that is identically and independently distributed across individuals according to a double exponential (Gumbel) distribution with zero mean and variance $\pi^2 \chi^2/6$ (the letter π here is the trigonometric $\pi \approx 3.14$) where $\chi \geq 0$ is a parameter. For notational convenience let $\bar{A}$ be the set of regions in A, i.e., $\bar{A} = \{I, B\}$. Given this distribution, the probability that an individual will choose to reside in region i of country A is given by the logit formula

$$\Pr{}_i\left(\omega_i, \chi\right) = \frac{\exp\left(\frac{\omega_i}{\chi}\right)}{\sum\limits_{i \in \bar{A}} \exp\left(\frac{\omega_i}{\chi}\right)}, \tag{30.4}$$

When $\chi \to 0$, individuals tend to have identical preferences and choose their region of residence solely according to the indirect utility derived from consumption of M and H. This is the preference structure of the model we considered in Sect. 16.4.1. As χ increases, idiosyncratic location preferences become more important and in the extreme case of $\chi \to \infty$ they are all that matters for workers' location choices.

The stock of H in each region is constant. The relative number of varieties, the price of H, total expenditure, and the real wage are given, respectively, by expressions (16.8)–(16.11) in Sect. 16.4.1. The zero profit condition determines the equilibrium size of firm output $q^* = (F/a)(\sigma - 1)$. Product-market equilibrium requires equality of supply and demand for any variety of M produced in each region. For notational convenience let $\bar{W}$ be the set of all regions, i.e., $\bar{W} = \{I, B, R\}$. Using the expressions for optimal prices, the price index for tradeable, total expenditure, the relative number of varieties, and equilibrium output of any variety, the system of goods market equilibrium equations is

$$1 = \sum_{j \in \bar{W}} \frac{\left(\tau_{ij}\right)^{\sigma-1} (w_i)^{-\sigma}}{\sum_{j \in \bar{W}} \lambda_i \left(\tau_{ij}\right)^{\sigma-1} (w_i)^{1-\sigma}} w_j L_j \quad . \qquad i = I, B. \tag{30.5}$$

A spatial equilibrium is defined by the condition that *net* migration flows be zero:

$$L_B \operatorname{Pr}_I - L_I \operatorname{Pr}_B = 0. \tag{30.6}$$

The first summand in Eq. (30.6) is the migration flow from region B to region I, and the second summand is the migration flow from region I to region B. The spatial equilibrium requires them to be equal. Lastly,

$$\sum_{i \in \bar{A}} L_i = L_A, \tag{30.7}$$

where L_A is the exogenously given population in A.

We choose w_R as numéraire and set $w_R = 1$. The equilibrium is characterized by a vector of wages $\left[w_I^*, w_B^*\right]$ and labour allocations $\left[L_I^*, L_B^*\right]$ that satisfies the system of four independent equations composed by the product-market equilibrium equations (30.5), the spatial equilibrium equation (30.6) and the labour market equilibrium equation (30.7). Equilibrium values of all other endogenous variables can be computed from the equilibrium values of wages and labour allocations.

Taking the natural logarithm of both sides of (30.6) and rearranging gives:

$$\omega_B - \omega_I = \chi \ln \left(\frac{L_B}{L_I} \right). \tag{30.8}$$

We therefore use the two Eqs. (30.5), (30.7), and (30.8) to obtain numerical solutions for w_I, w_B, L_I, and L_B.

30.2.2 Calibration of the Model

The starting point is the empirical observation. Let $\overset{\circ}{L}_i$ be the growth rate of employment between steady states. That is, the difference between the steady state equilibrium value of employment before the shock and the new steady state equilibrium value of employment on which the economy settles after the shock. Let $\overset{\circ}{w}_i$ be the analogous definition referred to wages. The estimated variable is the ratio between the difference in growth rates of employment and the difference in growth rates of wages:

$$\rho \equiv \frac{\overset{\circ}{L}_B - \overset{\circ}{L}_I}{\overset{\circ}{w}_B - \overset{\circ}{w}_I}. \tag{30.9}$$

Let $\hat{\rho}$ denote the estimated value of ρ. The empirical estimations give $\hat{\rho} \approx 3$ (more precisely the hypothesis $\hat{\rho} = 3$ is never rejected). We therefore take $\hat{\rho} = 3$ as the number to be replicated by the model.

To calibrate the model, we need to decide on the values of the following parameters: housing stocks (in each region), H_i, population in A and R, the elasticity of substitution among differentiated goods, $\sigma \equiv \mu/(\mu - 1)$, the expenditure share γ, the location preference parameter χ, and trade costs between regions. While for qualitative results these values can be chosen without any constraint, in the calibration for empirical purposes these values must satisfy two requirements: (1) they must be reasonable and (2) they must be such that the resulting values of the endogenous variables replicate the empirical evidence. The first criterion puts a constraint on the range in which parameter values may lie. Typically it is required that parameter values are in line with measures of them taken form independent sources. For example, in relation to the present calibration, it is reasonable to use the estimated value of σ found in the empirical literature. These estimates have found values that range in the interval from 3 to 6. The measurement for $(1-\gamma)$ are less numerous in the literature but the few existing studies find that the expenditure share on housing is approximately 0.25.[1] We therefore take the range 3–6 as reasonable values that may be assigned to σ and we take 0.25 as a reasonable value for $(1-\gamma)$. Coming to H_i, there is no data that suggest what reasonable values of the housing stocks could be but there is data on population. By inputting into the model the observed regional population before the shock and then solving the model for H_i we obtain consistent values of H_i.[2] These values are then used as exogenous regional stocks of H which remain constant throughout the simulation of trade liberalization. The distribution of the total stock of housing between A and R is instead arbitrarily assigned by choosing $H_R = \frac{H}{3}$ and $L_R = \frac{L}{3}$ and normalizing total stock of housing and labour by setting $H = L = 1$. Hence, A is twice the size of R. This is totally arbitrary but the lack of data does not allow to do any better. Thus, as a robustness check it is necessary to study the sensitivity of the results of different parametrization of H_R and L_R. We anticipate here that

[1] See Brülhart et al. (2012) for references on the empirical studies that have estimated σ and $(1-\gamma)$ as well as for references on the migration inertia reported below.

[2] In our data set region B accounted for 5.1 % of Austrian population prior to liberalization. Their implied housing stock in our calibrations ranges from 6 to 9 % of the total for country A.

Table 30.2 Baseline model: simulated values of ρ

	$\sigma = 3$	$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
$(1-\gamma) = 0.20$	10.33	9.60	9.23	9.00
$(1-\gamma) = 0.25$	7.70	7.16	6.88	6.71
$(1-\gamma) = 0.30$	5.97	5.54	5.33	5.20
$(1-\gamma) = 0.40$	3.82	3.55	3.43	3.33
$(1-\gamma) = 0.50$	2.54	2.36	2.27	2.21

robustness checks have shown that the simulated value of ρ is pretty insensitive to changes in the population and housing distribution between A and R.[3] Coming to the simulation of trade opening the following parameter values for trade costs are chosen to simulate external trade liberalization: $\tau_{IB} = 0.9$ (very low trade costs within A) and $\tau_{BR} = 0.1, 0.2, \ldots, 0.9$ (falling trade costs between A and R). The model is solved for each of the nine levels of τ_{BR} and we compute the relative change in steady state equilibrium nominal wages, $\overset{\circ}{w}_i$, and employment, $\overset{\circ}{L}_i$, for each 0.1 increment of trade cost reduction.[4] It turns out that ρ varies only trivially across pairs of trade costs for which it is calculated. We will therefore report averages of the eight computed ratios.

One last important matter concerns the stability of the equilibrium. The range of reasonable parameter values turns out to be such that the (only) equilibrium is globally stable for any value of trade costs. This means that the exogenous shock considered in the exercise will move the position of the equilibrium but will not move the economy on a different equilibrium.

Armed with the set of reasonable values to be assigned to the exogenous variables and to parameters, except for χ which we discuss below, and knowing that the equilibrium we shock is stable we can now come to confrontation with data. To show how empirical evidence is a guidance to calibration we solve the model first for $\chi \to 0$. The results are shown in Table 30.2 which reports the simulated values of ρ. It is clear that reasonable parameter values for $(1-\gamma)$ and σ do not easily replicate $\hat{\rho} = 3$. Only expenditure shares on housing well above 0.25 or unreasonably high values of σ yield corresponding simulated values of ρ near 3. The model simulated in the absence of migration inertia therefore is not able to satisfactorily replicate the data. In particular, the simulated values of ρ are too high with respect to $\hat{\rho}$ which, by inspection of expression (30.9), means that the model allows for too

[3]This insensitivity is not surprising. By increasing the size of R, for instance, trade liberalization becomes more important for both I and B, but more so for B. Yet, I is not a measure of the locational attractiveness of B relative to I; rather, it captures whether that increased attractiveness manifests itself more in terms of employment growth or in terms of nominal wage growth. This ratio is largely insensitive to the overall attractiveness of B with respect to I.

[4]$\overset{\circ}{w}_i$ and $\overset{\circ}{L}_i$ are growth rates between steady states. Their empirical counterparts are the average or cumulative growth rates over the entire pre- and post-liberalization subperiods, assuming that these subperiods are sufficiently long to capture the full transition between steady states.

Table 30.3 Extended model: implied immobility for $\rho = 3$

	$\sigma = 3$	$\sigma = 4$	$\sigma = 5$	$\sigma = 6$
$(1-\gamma) = 0.20$	25 [19]	29 [22]	30 [23]	31 [25]
$(1-\gamma) = 0.25$	24 [14]	28 [16]	31 [17]	32 [18]
$(1-\gamma) = 0.30$	20 [10]	25 [12]	27 [12]	29 [12]

much labour interregional mobility with respect to what is implied by the data. Therefore, introducing migration inertia, here in the form of idiosyncratic location preferences, promises to be a suitable way of adapting the model to replicate the empirical evidence. We therefore set $\chi > 0$.

There is no data to offer guidance about plausible values of χ. Still it is necessary to constrain χ in reasonable ranges to make the calibration meaningful. One way of doing so is to take account that when χ is positive the spatial equilibrium gives rise to real wage differences which do not induce migration because they are compensated by location preferences. Therefore, one way of constraining χ in a reasonable range is to find the values of χ such that the resulting equilibrium real wage differences between regions appear to be reasonable. But what are reasonable real wage differences that do not induce migration? There is only few evidence, mainly based on surveys, on this type of migration inertia. One study reports that 34 % of EU15 unemployed and 25 % of Czech unemployed stated in 2002 that they would not move under any circumstances even if a job became available elsewhere. One other study has found that the percentage of Italian unemployed refusing to move out of their town of residence if a job were available elsewhere ranges from 20.7 % (Northern male university graduates) to 61 % (Southern low-education females). These surveys would then serve as measures of reasonable migration inertia.

Solving the model for χ such that $\rho = 3$ gives the results shown in Table 30.3. Each cell reports the implied percentage real-wage difference between regions of A and, in brackets, the implied share of A's population that prefers not to migrate at the prevailing real-wage difference. It is clear that allowing for heterogenous location preferences aligns the simulated value of ρ with the estimated ρ under more reasonable parameter values. For instance, the necessary degree of preference heterogeneity when $\sigma = 4$ and $(1-\mu) = 0.25$ is such that 16 % of the population would not move even if the real wage were 28 % higher in the other region. In light of the available European evidence on the issue, this does not appear to be an excessive dose of assumed intrinsic migration inertia.

References

Brülhart, M., Carrère, C., & Trionfetti, F. (2012). How wages and employment adjust to trade liberalization: Quasi-experimental evidence from Austria.

Gandolfo, G. (2009). *Economic dynamics*.

Judd, K. L. (1998). *Numerical methods in economics*.

Murata, Y. (2003). Product diversity, taste heterogeneity, and geographic distribution of economic activities: Market vs. non-market interactions.

Tabuchi, T., & Thisse, J.-F. (2002). Taste heterogeneity, labor mobility and economic geography.

Chapter 31
Appendix to Chapter 17

31.1 A Model of Economic Geography, International Trade, and Globalization

The following model is due to Krugman and Venables (1995).

The world consists of two countries (or regions: henceforth we shall use 'region' and 'country' interchangeably), North and South, assumed to be identical in technology, endowments, and preferences. Both countries produce two kinds of commodities: an agricultural good (produced under constant returns to scale with labour as the sole input) and a variety of manufacturing goods (produced under increasing returns to scale using labour and intermediate goods). Manufacturing goods can be used both as final goods by consumers and as intermediate goods in the manufacturing sector.

Given the identity of the two regions, we shall describe the equations concerning North, as analogous equations hold in South. Variables with a superscript asterisk refer to South.

31.1.1 The Demand Side

The representative consumer receives only labour income, and has preferences that can be represented by a Cobb-Douglas expenditure function $Q_A^{(1-\gamma)} Q_M^{\gamma} V$, where V is utility, Q_A is the price of the agricultural good, Q_M is the price index for manufacturing goods, and γ is the share of manufactures in the expenditure of consumers. Given L, the country's labour endowment, and assuming that the representative consumer receives only labour income at the wage rate w, the budget constraint is

$$wL = Q_A^{(1-\gamma)} Q_M^{\gamma} V. \tag{31.1}$$

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5_31,

To determine Q_M, we begin by observing that, in equilibrium, all varieties of differentiated goods produced by the manufacturing sector are sold at the same price p (see Sect. 23.2.2). Because of increasing returns to scale coupled with the consumers' love for variety and the unlimited number of potential varieties, no firm will try to produce the same variety already produced by another firm (see Sect. 23.2.1). Hence the number of available varieties coincides with the number of firms in operation, that will usually be very large.

Now, as is typical of monopolistic competition models of international trade (see Sect. 23.2.3), there will be intra-industry trade in manufactures; we assume that manufacturing products of South sold in North incur iceberg transport costs at the rate $t \geq 1$. This means that a fraction $1/t$ of the good exported by South arrives to North, hence a consumer price in North of p^*t.

If we then aggregate all varieties of differentiated manufacturing products by a CES subutility function we obtain Q_M, which takes the form

$$Q_M = \left[np^{1-\sigma} + n^*(p^*t)^{1-\sigma}\right]^{1/(1-\sigma)}, \tag{31.2}$$

where n, n^* denote the number of varieties produced in North and South, respectively, and $\sigma > 1$ is the price elasticity of demand for a single variety. Let us note for future reference that Q_M is a decreasing function of the number of varieties.

31.1.2 The Supply Side

Agriculture is perfectly competitive and produces under constant returns to scale with labour as the sole input. The agricultural good can be costlessly traded, and is taken as the numéraire ($Q_A = 1$). Without loss of generality we can choose units so that one unit of labour produces one unit of output, which gives the equilibrium condition

$$w \geq 1, \tag{31.3}$$

where the equality sign must hold if the agricultural good is produced. Hence the wage rate equals unity if the country produces agriculture, and exceeds it only if there is no agricultural production.

The manufacturing sector uses labour and manufacturing intermediates to produce the composite final consumer good and intermediates. To keep the dimensionality of the model low, the major simplifying assumption is made that manufacturing output is an all-purpose composite commodity, that can be used both as composite final consumer good and as composite intermediate good. Thus the intermediate's price index is Q_M, as defined in Eq. (31.2).

Production is carried out by the representative firm combining labour and the intermediate with a Cobb-Douglas technology with shares $(1 - \mu)$ and μ, respectively, with α units of the combined input used as fixed cost, and β units per unit output as variable cost. Since each firm produces output for both the domestic market (y) and exports (x), we can write each firm's total cost as

$$TC = w^{1-\mu} Q_M^{\mu} [\alpha + \beta(y + x)], \tag{31.4}$$

whence the marginal cost

$$MC = w^{1-\mu} Q_M^{\mu} \beta. \tag{31.5}$$

How do increasing returns to scale manifest themselves in these functions? Simply through the reduction in Q_M (and hence in total and marginal cost) due to the increase in n, the number of domestic products (= the number of domestic firms). Let us note, incidentally, that the reduction in cost due to the increase in the number of firms is typical of monopolistically competitive models with constant-elasticity demand functions, where all scale economies work through changes in the variety of goods produced.

31.1.3 Equilibrium

To characterize equilibrium we first note that, with free entry and exit of firms, a zero-profit situation obtains. We next define E, the total value of Northern expenditure on manufactured commodities, which is given by consumers' expenditure (a proportion γ of the wage bill) and intermediate demand, which is a proportion μ of costs (and hence of revenue, since there are no profits)

$$E = \gamma w L + \mu(x + y) p n. \tag{31.6}$$

Let us now consider price determination. It is well known that, under monopolistic competition, the equilibrium excess-price over marginal cost equals the reciprocal of the elasticity of demand, namely $(p - MC)/p = 1/\sigma$, whence the firm's price-setting rule

$$p = \frac{\sigma}{\sigma - 1} MC,$$

that is,

$$p = \frac{\sigma}{\sigma - 1} w^{1-\mu} Q_M^{\mu} \beta. \tag{31.7}$$

We now note that the demand for a single variety is

$$y = p^{-\sigma} Q_M^{\sigma-1} E, \qquad x = p^{-\sigma} t^{1-\sigma} \left(Q_M^*\right)^{\sigma-1} E^*, \tag{31.8}$$

in North and South, respectively.

The zero profit condition means $p(y+x) = TC$. Substituting TC from (31.4) and p from (31.7) we get

$$y + x = (\sigma - 1)\alpha/\beta. \tag{31.9}$$

Without loss of generality we can choose units of measurement such that $(\sigma - 1)\alpha/\beta = 1$, whence, at the zero-profit equilibrium,

$$y + x = 1. \tag{31.10}$$

Substitution of (31.8) in (31.10) gives

$$p^{-\sigma}\left[Q_M^{\sigma-1}E + t^{1-\sigma}\left(Q_M^*\right)^{\sigma-1}E^*\right] = 1, \tag{31.11}$$

whence

$$p^{\sigma} = Q_M^{\sigma-1}E + t^{1-\sigma}\left(Q_M^*\right)^{\sigma-1}E^*. \tag{31.12}$$

The equilibrium values of the endogenous variables Q_M, w, p, n, E, and of the analogous variables in the other country, are determined by Eqs. (31.2), (31.3), (31.6), (31.7), and (31.11), and analogous equations for the other region.

We now note that n, the number of firms (= varieties) in the manufacturing sector, influences firms' profitability in three ways.

(a) As shown by Eq. (31.2), an increase in n reduces Q_M. This shifts each firm's demand curve down—see Eq. (31.8)—and reduces firms' profitability, see Eq. (31.11). This is the standard channel. The two other are related to a positive μ, namely are operative only if manufacturing uses manufactures as input (see above).
(b) The reduction in Q_M due to the increase in n causes a decrease in total and marginal cost—see Eqs. (31.4) and (31.5)—and hence an increase in firms' profits. This is a *cost,* or *forward linkage* between firms.
(c) An increase in n increases total expenditure on manufactures—see Eq. (31.6)—that in turn raises demand and profits of each firm, as shown by Eqs. (31.8) and (31.11). This is a *demand*, or *backward linkage* between firms.

31.2 The Dynamics of the Model and the Emergence of a Core-Periphery Pattern

The dynamics of the model could be examined in terms of the number of firms in the manufacturing sector in the two countries, on the basis of the assumption that firms enter in the sector if profits are positive, exit in the opposite case. Hence the dynamic system

$$\begin{aligned} \dot{n} &= f(\pi), \quad \operatorname{sgn} f(\pi) = \operatorname{sgn} \pi, f(0) = 0, f'(0) > 0, \\ \dot{n}^* &= f^*(\pi^*), \operatorname{sgn} f^*(\pi) = \operatorname{sgn} \pi^*, f^*(0) = 0, f^{*\prime}(0) > 0, \end{aligned} \tag{31.13}$$

where the dot over a variable denotes its time derivative.

Since profits are total revenue minus total cost, using (31.4), (31.7), and (31.8) it is easy to see that profits (given transportation cost) ultimately depend on the number of firms, namely

$$\begin{aligned} \pi &= \pi(n, n^*; t), \\ \pi^* &= \pi^*(n, n^*; t), \end{aligned} \tag{31.14}$$

so that, substituting into (31.13), we can write our dynamic system as

$$\begin{aligned} \dot{n} &= \varphi(n, n^*; t), \\ \dot{n}^* &= \varphi^*(n, n^*; t). \end{aligned} \tag{31.15}$$

This is the typical form of a planar system involving a parameter (t), which can give rise to bifurcations (for a treatment of bifurcation theory see Gandolfo, 2009, chap. 24).

Loosely speaking, we have a bifurcation when, given a dynamic system involving a parameter, the passage of the parameter through a critical value causes a qualitative change in the nature of singular point of the system (for example, from stability to instability). The value(s) of the parameter at which such a change occurs are called bifurcation values.

A bifurcation point consistent with the formation of a core-periphery pattern due to the decline in transportation costs requires the system to be stable for $t > t_0$ (i.e., for cost of transport higher than a critical value t_0) and unstable for $t < t_0$.

However, the study of bifurcations in a 2×2 system is rather complicated, hence it is convenient to transform the model so as to reduce its dynamics to a single differential equation. This can be done by concentrating on manufacturing equilibrium in North.

31.2.1 The Manufacturing Sector

We first observe that in manufacturing, since proportion μ of costs is spent on intermediates (see above, Eq. (31.6)), the remaining proportion $(1 - \mu)$ is spent for the wage bill, which is

$$wL_M = (1 - \mu)np(y + x), \tag{31.16}$$

where L_M is manufacturing employment. Assuming an initially symmetric equilibrium, where both countries produce both commodities and have a wage equal to unity, the proportions L_M/L and L_A/L are equal to γ and $1-\gamma$, respectively. Given our choice of the units of measurement—see Eq. (31.10)—we have

$$wL_M = (1-\mu)np. \tag{31.17}$$

Let us now consider the ratios of Northern to Southern endogenous variables, defined as

$$\widetilde{Q}_M \equiv \frac{Q_M}{Q_M^*},\ \widetilde{p} \equiv \frac{p}{p^*},\ \widetilde{E} \equiv \frac{E}{E^*},\ \widetilde{w} \equiv \frac{w}{w^*}. \tag{31.18}$$

For future reference we also define

$$\tau \equiv t^{1-\sigma}, \tag{31.19}$$

and observe that, since $\sigma > 1$, $t \in (1,\infty)$ implies $\tau \in (0,1)$.

From Eq. (31.2) and its analogous for the other region we have

$$\widetilde{Q}_M = \frac{\left[np^{1-\sigma} + n^*(p^*t)^{1-\sigma}\right]^{1/(1-\sigma)}}{\left[n^*(p^*)^{1-\sigma} + n(pt)^{1-\sigma}\right]^{1/(1-\sigma)}},$$

whence

$$\widetilde{Q}_M^{1-\sigma} = \frac{np^{1-\sigma} + n^*(p^*t)^{1-\sigma}}{n(pt)^{1-\sigma} + n^*(p^*)^{1-\sigma}} = \frac{(np)p^{-\sigma} + n^*p^*\tau(p^*)^{-\sigma}}{np\tau p^{-\sigma} + n^*p^*(p^*)^{-\sigma}}.$$

Substituting $np = L_M w/(1-\mu), n^*p^* = L_M^* w^*/(1-\mu)$, that are derived from Eq. (31.17) and its South analogous, we obtain, after simple manipulations,

$$\widetilde{Q}_M^{1-\sigma} = \frac{L_M \widetilde{w}\widetilde{p}^{-\sigma} + \tau L_M^*}{\tau L_M \widetilde{w}\widetilde{p}^{-\sigma} + L_M^*}. \tag{31.20}$$

Consider now $\widetilde{E}$. From Eq. (31.6) and its analogous for the other region we have, using $np = L_M w/(1-\mu), n^*p^* = L_M^* w^*/(1-\mu)$ and taking (31.10) into account,

$$\widetilde{E} = \frac{\gamma wL + \frac{\mu}{1-\mu}L_M w}{\gamma w^* L^* + \frac{\mu}{1-\mu}L_M^* w^*} = \widetilde{w}\frac{\gamma(1-\mu)L + \mu L_M}{\gamma(1-\mu)L + \mu L_M^*}, \tag{31.21}$$

where we have used the fact that $L^* = L$ by our initial assumptions.

As regards $\widetilde{p}$, from Eq. (31.7) and its analogous we have

$$\widetilde{p} = \frac{\frac{\sigma}{\sigma-1}w^{1-\mu}Q_M^{\mu}\beta}{\frac{\sigma}{\sigma-1}w^{*1-\mu}Q_M^{*\mu}\beta} = \widetilde{w}^{1-\mu}\widetilde{Q}_M^{\mu}. \tag{31.22}$$

We finally have, using Eq. (31.12), its analogous for the other region, and the definition of τ,

$$\widetilde{p}^{\sigma} = \frac{Q_M^{\sigma-1}E + \tau\left(Q_M^*\right)^{\sigma-1}E^*}{\left(Q_M^*\right)^{\sigma-1}E^* + \tau\left(Q_M\right)^{\sigma-1}E} = \frac{\widetilde{Q}_M^{\sigma-1}\widetilde{E} + \tau}{\tau\widetilde{Q}_M^{\sigma-1}\widetilde{E} + 1}. \tag{31.23}$$

System (31.20)–(31.23) can be reduced to a two-equation system by eliminating $\widetilde{Q}_M$ and $\widetilde{E}$. From Eq. (31.22) we obtain

$$\widetilde{p}^{\frac{1-\sigma}{\mu}}\widetilde{w}^{(1-\sigma)(\mu-1)/\mu} = \widetilde{Q}_M^{1-\sigma}, \tag{31.24}$$

hence, substituting $\widetilde{Q}_M^{1-\sigma}$ from (31.20) and rearranging terms

$$\widetilde{p}^{\frac{1-\sigma}{\mu}}\widetilde{w}^{(1-\sigma)(\mu-1)/\mu} - \frac{L_M\widetilde{w}\widetilde{p}^{-\sigma} + \tau L_M^*}{\tau L_M\widetilde{w}\widetilde{p}^{-\sigma} + L_M^*}$$
$$\equiv \varphi_1\left(\widetilde{p}, \widetilde{w}, L_M, L_M^*\right) = 0. \tag{31.25}$$

Consider now Eq. (31.23), which yields $\widetilde{p}^{\sigma}\left[\tau\widetilde{Q}_M^{\sigma-1}\widetilde{E} + 1\right] = \widetilde{Q}_M^{\sigma-1}\widetilde{E} + \tau$, whence solving for $\widetilde{Q}_M^{\sigma-1}\widetilde{E}$ we get

$$\widetilde{Q}_M^{\sigma-1}\widetilde{E} = \frac{\tau - \widetilde{p}^{\sigma}}{\tau\widetilde{p}^{\sigma} - 1}. \tag{31.26}$$

Substitution from (31.24) and (31.21) into (31.26) yields

$$\widetilde{p}^{(\sigma-1)/\mu}\widetilde{w}^{(\sigma-1)(\mu-1)/\mu}\widetilde{w}\frac{\gamma(1-\mu)L + \mu L_M}{\gamma(1-\mu)L + \mu L_M^*} - \frac{\tau - \widetilde{p}^{\sigma}}{\tau\widetilde{p}^{\sigma} - 1}$$
$$\equiv \varphi_2\left(\widetilde{p}, \widetilde{w}, L_M, L_M^*\right) = 0. \tag{31.27}$$

Equations (31.25) and (31.27) are a set of two implicit functions in the four variables $\widetilde{p}, \widetilde{w}, L_M, L_M^*$. According to the implicit function theorem (see, for example, Gandolfo, 2009, chap. 20, sect. 20.2) we can express $\widetilde{p}, \widetilde{w}$ as differentiable functions of L_M, L_M^* provided that the Jacobian of the set is non-singular at the equilibrium point.

For this purpose we first observe that, by simple inspection, if $L_M = L_M^*$, Eqs. (31.25) and (31.27) are satisfied for $\widetilde{p} = \widetilde{w} = 1$. This we take as our (symmetric) equilibrium.

Let us now compute the Jacobian

$$\mathbf{J} \equiv \begin{bmatrix} \dfrac{\partial \varphi_1}{\partial \widetilde{p}} & \dfrac{\partial \varphi_1}{\partial \widetilde{w}} \\ \dfrac{\partial \varphi_2}{\partial \widetilde{p}} & \dfrac{\partial \varphi_2}{\partial \widetilde{w}} \end{bmatrix}, \tag{31.28}$$

where the partial derivatives are evaluated at the equilibrium point. Simple calculations yield

$$\mathbf{J} = \begin{bmatrix} \dfrac{1-\sigma}{\mu} + \dfrac{\sigma(1-\tau)}{1+\tau} & \dfrac{(\mu-1)(1-\sigma)}{\mu} + \dfrac{\tau-1}{1+\tau} \\ \dfrac{\sigma-1}{\mu} + \dfrac{\sigma(1+\tau)}{\tau-1} & \dfrac{\sigma(\mu-1)+1}{\mu} \end{bmatrix}, \tag{31.29}$$

from which we obtain, expanding the determinant of $\mathbf{J}$ and rearranging terms,

$$|\mathbf{J}| = \frac{2\tau}{\mu(1+\tau)(1-\tau)} \left\{2\sigma(\sigma-1)(1-\mu) + (\tau-1)\left[\sigma(\mu+1)-1\right]\right\}. \tag{31.30}$$

If we take the parameter restrictions into account (i.e., $\sigma > 1$, $0 < \mu < 1$, $\tau \in (0,1)$), we see that the fraction is positive, while the expression $\{2\sigma(\sigma-1)(1-\mu) + (\tau-1)[\sigma(\mu+1)-1]\}$ contains one positive and one negative term. To determine the sign of this expression, we first observe that it is a monotonically increasing function of τ. Hence if it is positive for $\tau = 0$ it will be positive for all positive τ. For $\tau = 0$ the expression becomes, after simple manipulations,

$$2\sigma(\sigma-1)(1-\mu) - [\sigma(\mu+1)-1] = (2\sigma-1)[\sigma(1-\mu)-1], \tag{31.31}$$

which will be positive when

$$\sigma(1-\mu) > 1 \quad \text{or,} \quad \frac{\sigma-1}{\sigma} > \mu, \tag{31.32}$$

a condition assumed by Krugman and Venables (1995, p. 878). This assumption turns out to be crucial—see below, Eqs. (31.43) and (31.44)—hence it is interesting to discuss its economic meaning. The condition requires either σ to be sufficiently high or μ to be sufficiently low.

Thus the first cause of the possible violation of condition (31.32) is that demand is insufficiently elastic (σ too low); the second cause (μ too high) is too high a share of intermediates in manufacturing costs, namely too strong backward and forward linkages. Now, when condition (31.32) is satisfied at the equilibrium point, a lower σ means a higher n (the equilibrium number of varieties produced), and so a lower σ means stronger economies of scale (economies of scale are positively related to n: see above, p. 619). Thus, in equilibrium, if σ is small economies of scale will be very high, and σ may become so small as to reverse inequality (31.32), which means that condition (31.32) will be violated when economies of scale are too strong (Krugman & Venables, 1995, p. 870).

Turning back to the mathematics, under (31.32) we have $|\mathbf{J}| \neq 0$, and there exist the differentiable functions

$$\begin{aligned} \widetilde{p} &= \widetilde{p}(L_M, L_M^*), \\ \widetilde{w} &= \widetilde{w}(L_M, L_M^*). \end{aligned} \tag{31.33}$$

We now consider a small change in manufacturing employment in North, $\mathrm{d}L_M$, with associated change $\mathrm{d}L_M^*$ in the opposite direction in the neighbourhood of the symmetric equilibrium. From the second function in (31.33) we have

$$\mathrm{d}\widetilde{w} = \frac{\partial \widetilde{w}}{\partial L_M}\mathrm{d}L_M + \frac{\partial \widetilde{w}}{\partial L_M^*}\mathrm{d}L_M^*.$$

Since we have assumed $\mathrm{d}L_M + \mathrm{d}L_M^* = 0$, we have

$$\mathrm{d}\widetilde{w} = \left(\frac{\partial \widetilde{w}}{\partial L_M} - \frac{\partial \widetilde{w}}{\partial L_M^*} \right) \mathrm{d}L_M,$$

whence

$$\frac{\mathrm{d}\widetilde{w}}{\mathrm{d}L_M} = \frac{\partial \widetilde{w}}{\partial L_M} - \frac{\partial \widetilde{w}}{\partial L_M^*}, \tag{31.34}$$

that gives the total effect on wages of the assumed change in manufacturing employment.

The comparative statics method (Gandolfo, 2009, chap. 20, sect. 20.2) gives us the way of rigorously computing $\partial\widetilde{w}/\partial L_M$, $\partial\widetilde{w}/\partial L_M^*$. Thus we have

$$\frac{\partial \widetilde{w}}{\partial L_M} = \frac{\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial \widetilde{p}} & -\dfrac{\partial \varphi_1}{\partial L_M} \\ \dfrac{\partial \varphi_2}{\partial \widetilde{p}} & -\dfrac{\partial \varphi_2}{\partial L_M} \end{vmatrix}}{|\mathbf{J}|} \tag{31.35}$$

and

$$\frac{\partial \widetilde{w}}{\partial L_M^*} = \frac{\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial \widetilde{p}} & -\dfrac{\partial \varphi_1}{\partial L_M^*} \\ \dfrac{\partial \varphi_2}{\partial \widetilde{p}} & -\dfrac{\partial \varphi_2}{\partial L_M^*} \end{vmatrix}}{|\mathbf{J}|}, \tag{31.36}$$

where $\partial\varphi_1/\partial L_M, \partial\varphi_2/\partial L_M, \partial\varphi_1/\partial L_M^*, \partial\varphi_2/\partial L_M^*$ are computed from Eqs. (31.25) and (31.27), and are evaluated at the symmetric equilibrium point. They turn out to be

$$\begin{aligned}\frac{\partial\varphi_1}{\partial L_M} &= \frac{\tau-1}{(1+\tau)L_M} = \frac{\tau-1}{(1+\tau)\gamma L},\\ \frac{\partial\varphi_2}{\partial L_M} &= \frac{\mu}{\gamma L},\\ \frac{\partial\varphi_1}{\partial L_M^*} &= \frac{1-\tau}{(1+\tau)L_M} = \frac{1-\tau}{(1+\tau)\gamma L},\\ \frac{\partial\varphi_2}{\partial L_M^*} &= -\frac{\mu}{\gamma L},\end{aligned} \tag{31.37}$$

where we have used the fact that $L_M = L_M^* = \gamma L$ at the symmetric equilibrium point. It is apparent from (31.37), given (31.35) and (31.36), that $\partial\widetilde{w}/\partial L_M = -\partial\widetilde{w}/\partial L_M^*$, hence (31.34) becomes

$$\frac{d\widetilde{w}}{dL_M} = 2\frac{\partial\widetilde{w}}{\partial L_M}. \tag{31.38}$$

Let us now calculate the numerator of the fraction in (31.35), call it N. Substituting the values of the partial derivatives found above and expanding the determinant we obtain, after rearrangement of terms,

$$N = \frac{1}{\gamma L\mu}\left\{(1-\mu)\left[\sigma(\mu-1)+1\right] + \tau(\mu+1)\left[\sigma(\mu+1)-1\right]\right\}. \tag{31.39}$$

Substitution of (31.39) and (31.30) in (31.35) and then in (31.38) yields, after rearrangement of terms,

$$\begin{aligned}\frac{d\widetilde{w}}{dL_M} &= \left(\frac{\tau-1}{\tau\gamma L}\right)\frac{(\mu-1)\left[\sigma(\mu-1)+1\right] - \tau(\mu+1)\left[\sigma(\mu+1)-1\right]}{2\sigma(\sigma-1)(1-\mu)+(\tau-1)\left[\sigma(\mu+1)-1\right]}\\ &= \left(\frac{\tau-1}{\tau\gamma L}\right)\frac{(1-\mu)\left[\sigma(1-\mu)-1\right] - \tau(\mu+1)\left[\sigma(\mu+1)-1\right]}{2\sigma(\sigma-1)(1-\mu)+(\tau-1)\left[\sigma(\mu+1)-1\right]}.\end{aligned} \tag{31.40}$$

We have already seen—see Eq. (31.30)—that the expression in the denominator of the second fraction on the r.h.s. is positive. Given that $(\tau-1)<0$, it follows that

$$\frac{d\widetilde{w}}{dL_M} \gtreqless 0 \text{ according as } (1-\mu)\left[\sigma(1-\mu)-1\right] - \tau(\mu+1)\left[\sigma(\mu+1)-1\right] \lesseqgtr 0, \tag{31.41}$$

i.e., recalling (31.32), according as

$$\tau^{-1} \lesseqqgtr \frac{(\mu+1)\left[\sigma(\mu+1)-1\right]}{(1-\mu)\left[\sigma(1-\mu)-1\right]}, \tag{31.42}$$

which implies, given the definition of τ, that

$$\frac{\mathrm{d}\widetilde{w}}{\mathrm{d}L_M} \gtreqqless 0 \text{ according as } t^{\sigma-1} \lesseqqgtr \left(\frac{1+\mu}{1-\mu}\right)\left(\frac{\sigma(1+\mu)-1}{\sigma(1-\mu)-1}\right). \tag{31.43}$$

Let us denote by t_0 the value of t at which the above inequality is satisfied as an equality, and observe that, given (31.32) and $0 < \mu < 1$, both fractions in (31.43) are greater than unity, hence—since $\sigma > 1$—the critical value t_0 will certainly be greater than unity, which means a positive level of transportation cost.

It should now be pointed out the crucial role of assumption (31.32). Suppose that the contrary is true, namely $\sigma(1-\mu)-1 < 0$ (and suppose that this has no effect on $|\mathbf{J}|$, which remains non zero). Then division by a negative quantity would imply that in passing from (31.41) to (31.42) the order of the inequalities would have to be reversed, whence

$$\frac{\mathrm{d}\widetilde{w}}{\mathrm{d}L_M} \gtreqqless 0 \text{ according as } t^{\sigma-1} \gtreqqless \left(\frac{1+\mu}{1-\mu}\right)\left(\frac{\sigma(1+\mu)-1}{\sigma(1-\mu)-1}\right). \tag{31.44}$$

This means that, since the r.h.s. of the inequality is negative, there exists no positive critical value t_0, hence $\mathrm{d}\widetilde{w}/\mathrm{d}L_M > 0$ and the model would be always unstable (see the next section), i.e., a core-periphery pattern would always emerge no matter how high transport costs are. In other words, the forces driving to industrial agglomeration in North would always predominate: this region would become a kind of black hole for world industry. In economic terms, increasing returns to scale are so strong (let us recall that the crucial inequality is reversed when σ is too low, namely economies of scale are too strong) that an increase in manufacturing employment in North always causes an increase in the manufacturing wage rate in North relative to South.

31.2.2 The Dynamics: Bifurcation Analysis

We now come to the dynamics proper. We have seen that $\widetilde{w}$, the manufacturing wage rate in North relative to South, is a function of manufacturing employment L_M, given transport costs. Since in the symmetric equilibrium the wage rate is taken to be unity (see above), such a function will determine the corresponding equilibrium manufacturing employment, say L_M^e. There may be more than one equilibrium value of L_M.

Any such equilibrium will be stable if actual manufacturing employment tends to increase (decrease) when it falls short of (exceeds) the equilibrium value considered, unstable in the opposite case.

This is little more than a tautology, thus we must look for the forces that cause manufacturing employment to change. These are undoubtedly given by the wage rate: more precisely, actual manufacturing employment in North will tend to increase (decrease) if the wage rate there happens to be higher (lower) than in South, whatever the cause that has displaced the symmetric equilibrium. Given the functional relation between $\widetilde{w}$ and L_M, the variation in manufacturing employment will bring about a change in the wage rate, which will in turn feed back on manufacturing employment, and so forth.

The formal counterpart of this dynamic behaviour is the differential equation

$$\dot{L}_M = f\left(\widetilde{w}\right),\ f(1) = 0,\ f'(1) > 0,\ \mathrm{sgn}\ f\left(\widetilde{w} - 1\right) = \ \mathrm{sgn}\ \left(\widetilde{w} - 1\right), \quad (31.45)$$

where $\widetilde{w} = h(L_M; t)$, hence

$$\dot{L}_M = f\left(h(L_M; t)\right). \quad (31.46)$$

Equation (31.46) is a one-parameter differential equation, where a codimension-one bifurcation may occur. This is indeed the case. In fact, the characteristic root of its linear approximation at the equilibrium point is

$$\lambda = k\frac{\mathrm{d}\widetilde{w}}{\mathrm{d}L_M},$$

where $k \equiv f'(1)$, and $\mathrm{d}\widetilde{w}/\mathrm{d}L_M$ is evaluated at the equilibrium point. Equilibrium will be stable (unstable) when $\lambda \lesseqgtr 0$, respectively. This is equivalent to $\mathrm{d}\widetilde{w}/\mathrm{d}L_M \lesseqgtr 0$, given that $k > 0$. Using (31.43), we see that there is a bifurcation point at $t = t_0$, at which the equilibrium from stable ($t > t_0$ implies $\mathrm{d}\widetilde{w}/\mathrm{d}L_M < 0$) becomes unstable ($t < t_0$ implies $\mathrm{d}\widetilde{w}/\mathrm{d}L_M > 0$).

This proves that as transportation costs decline, there is a critical point at which the core-periphery pattern emerges.

For numerical simulations of this model see Krugman and Venables (1995).

References

Gandolfo, G. (2009). *Economic dynamics*.
Krugman, P. R., & Venables, A. J. (1995). Globalization and the inequality of nations.

Bibliography

Acemoglu, D. (2002). Directed technical change. *Review of Economic Studies, 69*, 781–809.

Acemoglu, D. (2003). Patterns of skill premia. *Review of Economic Studies*, 70, 199–230.

Aghion, P., & Howitt, P. (1990). *A model of growth through creative destruction* (NBER Working Paper 3223).

Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica, 60*, 323–351.

Aghion, P., & Howitt, P. (1998). *Endogenous growth theory*. Cambridge, MA: MIT.

Allen, R. G. D. (1967). *Macro-economy theory: A mathematical treatment*. London: Macmillan.

Allen, W. R. (Ed.). (1965). *International trade theory: Hume to Ohlin*. New York: Random House.

Alonso-Villar, O. (2005). The effects of transport costs revisited. *Journal of Economic Geography, 5*, 589–604.

Amir, R. (2005). Supermodularity and complementarity in economics: An elementary survey. *Southern Economic Journal, 71*, 636–660.

Amiti, M., & Davis, D. R. (2011). Trade, firms, and wages: Theory and evidence. *Review of Economic Studies, 79*, 1–36.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *American Economic Review, 69*, 106–116.

Anderson, J. E. (1981). The Heckscher-Ohlin and Travis-Vanek theorems under uncertainty. *Journal of International Economics, 11*, 239–247.

Anderson, S. P., Schmitt, N., & Thisse, J.-F. (1995). Who benefits from antidumping legislation? *Journal of International Economics, 38*, 321–337.

Andersson, F., & Forslid, R. (2003). Tax competition and economic geography. *Journal of Public Economic Theory, 5*, 279–304.

Annicchiarico, B., Orioli, F., & Trionfetti, F. (2012). National oligopoly and economic geography. *Annals of Regional Science, 48*, 71–99.

Antràs, P., Garicano, L., & Rossi-Hansberg, E. (2006). Offshoring in a knowledge economy. *Quarterly Journal of Economics, 121*, 31–77.

Antràs, P., & Helpman, E. (2004). Global sourcing. *Journal of Political Economy, 112*, 552–580.

Aquino, A. (1981). Changes over time in the pattern of comparative advantage in manufactured goods: An empirical analysis for the period 1962–1974. *European Economic Review, 15*, 41–62.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk bearing. *Review of Economic Studies, 31*, 91–96.

Balassa, B. (1965). Tariff protection in industrial countries: An evaluation. *Journal of Political Economy, 73*, 573–594.

Baldwin, Robert E. (1971). *Non-tariff distortions of international trade*. Washington, DC: The Brookings Institution.

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5,

Baldwin, Robert E. (1982). The political economy of protectionism. In Bhagwati, J. N. (Ed.), *Import competition and response* (pp. 263–286). Chicago: Chicago University Press. (Reprinted in R. E. Baldwin, 1988, pp. 97–136)

Baldwin, Robert E. (1988). *Trade policy in a changing world economy*. London: Harvester-Wheatsheaf.

Baldwin, Robert E. (1992). Are economist's traditional trade policy views still valid? *Journal of Economic Literature, 30*, 804–829.

Baldwin, Richard E. (1994). *Towards an integrated Europe*. London: Centre for Economic Policy Research.

Baldwin, Richard E. (2001). The Core-Periphery model with forward-looking expectations. *Regional Science and Urban Economics, 31*, 21–49.

Baldwin, Robert E. (2008). *The development and testing of Heckscher-Ohlin trade models*. Cambridge, MA: MIT.

Baldwin, Richard E., Forslid, R., Martin, P., Ottaviano, G. I. P., & Robert-Nicoud, F. (2003). *Economic geography and public policy*. Princeton, NJ: Princeton University Press.

Baldwin, Richard E., & Krugman, P. R. (2004). Agglomeration, integration and tax harmonisation. *European Economic Review, 48*, 1–23.

Baldwin, Richard E., Martin, P., & Ottaviano, G. I. P. (2001). Global economic divergence, trade and industrialisation: The geography of growth takeoffs. *Journal of Economic Growth, 6*, 5–37.

Baldwin, Richard E., & Ottaviano, G. I. P. (2001). Multiproduct multinationals and reciprocal FDI dumping. *Journal of International Economics, 54*, 429–448.

Bandyopadhyay, S. (1997). Demand elasticities, asymmetry and strategic trade policy. *Journal of International Economics, 42*, 167–177.

Bardhan, P. K. (1970). *Economic growth, development, and foreign trade: A study in pure theory*. New York: Wiley.

Barker, J. (1977). International trade and economic growth: An alternative to the neoclassical approach. *Cambridge Journal of Economics, 1*, 153–172.

Barone, E. (1908). *Principi di economia politica*. Roma: Athenaeum.

Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth* (2nd ed.). New York: McGraw-Hill.

Basu, K. (2001). On the goals of development. In G. Meier & J. Stiglitz (Eds.), *Frontiers of development economics: The future in perspective*. Washington, DC: World Bank and Oxford University Press.

Batra, R. N., & Casas, F. R. (1973). Intermediate products and the pure theory of international trade: A neo-Heckscher-Ohlin framework. *American Economic Review, 63*, 297–311.

Beccaria, C. (1764). Tentativo analitico sui contrabbandi, il Caffe' (Vol. 1). (Reprinted in P. Custodi (a cura di), 1804, *Scrittori italiani e classici di Economia Politica*, parte moderna, tomo XII, Milano, pp. 235–241)

Beckmann, M. J. (1987). Location of economic activity. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new Palgrave: A dictionary of economics* (Vol. 3). London: Macmillan.

Behrens, K., Lamorgese, A., Ottaviano, G. I. P., & Tabuchi, T. (2009). Beyond the home market effect: Market size and specialization in a multi-country world. *Journal of International Economics, 79*, 259–265.

Behrens, K., & Robert-Nicoud, F. (2010). Tempora mutantur: In search of a new testament for NEG, mimeo. Available at SSRN: http://ssrn.com/abstract=1593769

Belloc, M. (2006). Institutions and international trade: A reconsideration of comparative advantage. *Journal of Economic Surveys, 20*, 3–26.

Belloc, M., & Bowles, S. (2013). *Cultural-institutional persistence under autarchy, international trade, and factor mobility* (Working Paper 13-01-003). Santa Fe Institute.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic Perspectives, 21*, 105–130.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics, 4*, 283–313.

Bernard, A. B., Redding, S. J., & Schott, P. K. (2007). Comparative advantage and heterogenous firms. *Review of Economic Studies, 74*, 31–66.

Bernhardt, D. (1984). Dumping costs and uncertainty. *Journal of Economic Dynamics and Control, 8*, 349–370.

Bernhofen, D. M. (1995). Price dumping in intermediate good markets. *Journal of International Economics, 39*, 159–173.

Bhagwati, J. N. (1958). Immiserizing growth: A geometrical note. *Review of Economic Studies, 25*, 201–205. (Reprinted in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 300–305; in J. N. Bhagwati, Ed., 1981, pp. 301–306; and in J. N. Bhagwati, 1983a, pp. 301–306)

Bhagwati, J. N. (1964). The pure theory of international trade: A survey. *Economic Journal, 74*, 1–78. (Reprinted in J. N. Bhagwati, 1983b, pp. 313–432)

Bhagwati, J. N. (1965). On the equivalence of tariffs and quotas. In Various Authors (Eds.), *Trade, growth and the balance of payments: Essays in honour of G. Haberler* (pp. 53–67). Chicago: Rand McNally. (Reprinted in J. N. Bhagwati, 1983a, pp. 196–214)

Bhagwati, J. N. (Ed.). (1969). *International trade—Selected readings*. Harmondsworth, UK: Penguin.

Bhagwati, J. N. (1971). The generalized theory of distortions and welfare. In J. N. Bhagwati et al. (Eds.), pp. 69–90. (Reprinted in J. N. Bhagwati, Ed., 1981, pp. 171–189; and in J. N. Bhagwati, 1983a, pp. 73–96)

Bhagwati, J. N. (1972). The Heckscher-Ohlin theorem in the multi-commodity case. *Journal of Political Economy, 30*, 1052–1055. (Reprinted in J. N. Bhagwati, 1983b, pp. 462–465)

Bhagwati, J. N. (1973). The theory of immiserizing growth: Further application. In M. B. Connolly & A. K. Swoboda (Eds.), *International trade and money* (pp. 45–54). London: Allen & Unwin. (Reprinted in J. N. Bhagwati, 1983a, pp. 121–129)

Bhagwati, J. N. (Ed.). (1974). *Illegal transactions in international trade: Theory and measurement*. Amsterdam: North-Holland.

Bhagwati, J. N. (1981a). Alternative theories of illegal trade: Economic consequences and statistical detection. *Weltwirtschaftliches Archiv, 117*, 409–427. (Reprinted in J. N. Bhagwati, 1983a, pp. 325–342)

Bhagwati, J. N. (Ed.). (1981b). *International trade—Selected readings*. Cambridge, MA: MIT.

Bhagwati, J. N. (1982a). Directly unproductive, profit-seeking (DUP) activities. *Journal of Political Economy, 90*, 988–1002. (Reprinted in J. N. Bhagwati, 1983a, pp. 259–273)

Bhagwati, J. N. (1982b). Shifting comparative advantage, protectionist demands, and policy response. In J. N. Bhagwati (Ed.), *Import competition and response*. Chicago: University of Chicago Press.

Bhagwati, J. N. (1983a). *The theory of commercial policy—Essays in international economic theory* (R. C. Feenstra, Ed., Vol. 1). Cambridge, MA: MIT.

Bhagwati, J. N. (1983b). *International factor mobility—Essays in international economic theory* (R. C. Feenstra, Ed., Vol. 2). Cambridge, MA: MIT.

Bhagwati, J. N. (1987). VER, quid pro quo DFI and VIEs: Political-economy-theoretic analyses. *International Economic Journal, 1*, 1–14.

Bhagwati, J. N. (1995). U.S. trade policy: The infatuation with free trade agreements. In J. N. Bhagwati & A. O. Krueger (Eds.), *The dangerous drift to preferential trade agreements*. Washington, DC: American Enterprise Institute.

Bhagwati, J. N., Brecher, R. A., & Hatta, T. (1984). The paradoxes of immiserizing growth and donor-enriching "Recipient-immiserizing" transfers: A tale of two literatures. *Weltwirtschaftliches Archiv, 120*, 228–243.

Bhagwati, J. N., & Hansen, B. (1973). A theoretical analysis of smuggling. *Quarterly Journal of Economics, 87*, 172–187. (Reprinted in J. N. Bhagwati, Ed., 1974; and in J. N. Bhagwati, 1983a, pp. 286–301)

Bhagwati, J. N., & Hudec, R. E. (Eds.). (1996). *Fair trade and harmonization: Prerequisites for free trade?* (Economic analysis, Vol. I and Legal analysis, Vol. II). Cambridge, MA: MIT.

Bhagwati, J. N., Jones, J. F., Mundell, R. A., & Vanek, J. (Eds.). (1971). *Trade, balance of payments and growth: Essays in international economics in Honour of C. Kindleberger*. Amsterdam: North-Holland.

Bhagwati, J. N., & Panagariya, A. (Eds.). (1996). *The economics of preferential trading agreements*. Washington, DC: American Enterprise Institute.

Bhagwati, J. N., Panagariya, A., & Srinivasan, T. N. (1998). *Lectures on international trade* (2nd ed.). Cambridge, MA: MIT.

Bhagwati, J. N., & Srinivasan, T. N. (1974). An alternative proof of the Bhagwati-Hansen results on smuggling and welfare. In J. N. Bhagwati (Ed.), *Illegal transactions in international trade: Theory and measurement* (pp. 23–26). Amsterdam: North-Holland.

Bowen, H. P., Learner, E. E., & Sveinkauskas, L. (1987). Multifactor tests of the factor abundance theory. *American Economic Review, 77*, 791–809.

Bown, C. P. (2010). Temporary Trade Barriers Database (TTBD). Available at http://econ.worldbank.org/ttbd/

Brainard, S. L., & Martimort, D. (1997). Strategic trade policy with incompletely informed policymakers. *Journal of International Economics, 42*, 33–65.

Brander, J. A. (1981). Intra-industry trade in identical commodities. *Journal of International Economics, 11*, 1–14.

Brander, J. A., & Krugman, P. R. (1983). A "Reciprocal Dumping" model of international trade. *Journal of International Economics, 15*, 313–321.

Brander, J. A., & Spencer, B. J. (1984). Tariff protection and imperfect competition. In H. Kierzkowski (Ed.) (pp. 194–206).

Brander J. A., & Spencer, B. J. (1985). Export subsidies and international market share rivalry. *Journal of International Economics, 18*, 83–100.

Brecher, R. A., & Choudhri, E. U. (1982). The factor content of international trade without factor-price equalization. *Journal of International Economics, 12*, 277–283.

Brecher, R. A., & Choudhri, E. U. (1984). New products and the factor content of international trade. *Journal of Political Economy, 92*, 965–971.

Brems, H. (1980). *Inflation, interest and growth*. Lexington, MA: Lexington Books.

Brock, W. A., & Magee, S. P. (1978). The economics of special interest politics: The case of the tariff. *American Economic Review, 68*, 246–250 (Papers and Proceedings).

Brock, W. A., Magee, S. P., & Young, L. (1989). *Black hole tariffs and endogenous policy theory: Political economy in general equilibrium*. Cambridge, UK: Cambridge University Press.

Broda, C., & Weinstein, D. E. (2006). Globalization and the gains from variety. *Quarterly Journal of Economics, 121*, 541–585.

Brown, D. K., Deardoff, A. V., & Stern, R. M. (1996). In Bhagwati & Hudec (Eds.), *International labor standards and trade: A theoretical analysis* (pp. 227–280).

Brülhart, M., Carrère, C., & Trionfetti, F. (2012). *How wages and employment adjust to trade liberalization: Quasi-experimental evidence from Austria*. mimeo.

Brülhart, M., & Jametti, M. (2008). *Does tax competition tame the leviathan*? Université de Lausanne, mimeo.

Brülhart, M., Jametti, M., & Schmidheiny, K. (2012). Do agglomeration economies reduce the sensitivity of firm location to tax differentials? *Economic Journal, 122*, 1069–1093 .

Brülhart, M., & Parchet, R. (2011). *Alleged tax competition: The mysterious death of inheritance taxes in Switzerland*. Université de Lausanne, mimeo.

Brülhart, M., & Trionfetti, F. (2004). Public expenditure, international specialisation and agglomeration. *European Economic Review, 48*, 851–881.

Brülhart, M., & Trionfetti, F. (2009). A test of trade theories when expenditure is home biased. *European Economic Review, 53*, 830–845.

Bruno, M., & Sachs, J. (1982). Energy and resource allocation: A dynamic model of the "Dutch Disease". *Review of Economic Studies, 49*, 845–859.

Burmeister, E., & Dobell, A. R. (1970). *Mathematical theories of economic growth*. New York: Macmillan.

Burstein, A., & Vogel, J. (2012). *Globalization, technology, and the skill premium* (NBER Working Paper No. 16459), Cambridge, MA: National Bureau of Economic Research.

Bustos, P. (2011). Trade liberalization, exports and technology upgrading: Evidence on the impact of MERCOSUR on Argentinian firms. *American Economic Review*, 101, 304–340.

Casas, F. R. (1983). International trade with produced transport services. *Oxford Economic Papers, 35*, 89–109.

Casas, F. R., & Choi, E. K. (1984). Trade imbalance and the Leontief paradox. *Manchester-School, 52*, 391–401.

Casas, F. R., & Choi, E. K. (1985a). The Leontief paradox: Continued or resolved? *Journal of Political Economy, 93*, 610–615.

Casas, F. R., & Choi, E. K. (1985b). Some paradoxes of transport costs in international trade. *Southern Economic Journal, 51*, 983–987.

Casprini, F. (1979). L'incertezza nella teoria pura del commercio internazionale. In A. Vari (Ed.), *Economia internazionale: Teoria e prassi* (pp. 251–269). Milano: Giuffre'.

Caves, R. E. (1965). "Vent for Surplus" models of trade and growth. In Various authors (Eds.) *Trade, growth and the balance of payments: Essays in honour of G. Haberler* (pp. 95–115). Chicago: Rand McNally.

Caves, R. E. (1979). International cartels and monopolies in international trade. In R. Dornbusch & J. A. Frenkel (Eds.), *International economic policy* (pp. 39–73). Baltimore: Johns Hopkins.

Caves, R. E. (1981). Intra-industry trade and market structure in the industrial countries. *Oxford Economic Papers, 33*, 203–223.

Caves, R. E., Frankel, J. A., & Jones, R. W. (2006). *World trade and payments: An introduction* (10th ed.). Englewood Cliffs, NJ: Prentice Hall.

Caves, R. E., & Johnson, H. G. (Eds.). (1969). *Readings in international economics*. London: Allen & Unwin.

Chacholiades, M. (1978). *International trade theory and policy*. New York: McGraw-Hill.

Chacholiades, M. (1981). *Principles of international economics*. New York: McGraw-Hill.

Chacholiades, M. (1985). Circulating capital in the theory of international trade. *Southern Economic Journal, 52*, 1–22.

Chamberlin, E. H. (1933). *The theory of monopolistic competition*. Cambridge, MA: Harvard University Press.

Charlot, S., Gaigne, C., Robert-Nicoud, F., & Thisse, J.-F. (2006). Agglomeration and welfare: The core-periphery model in the light of Bentham, Kaldor, and Rawls. *Journal of Public Economics, 90*, 325–347.

Charos, E. N., & Simos, E. O. (1988). The effects of human capital and R&D upon international trade flows: Evidence from a multi-input, multi-output model for the United States. *Weltwirtschaftliches Archiv, 124*, 701–712.

Cheng, L. (1984). International competition in R&D and technological leadership: An examination of the Posner-Hufbauer hypothesis. *Journal of International Economics, 17*, 15–40.

Chipman, J. S. (1965a). A survey of the theory of international trade: Part 1, The classical theory. *Econometrica, 33*, 477–519.

Chipman, J. S. (1965b). A survey of the theory of international trade: Part 2, The neoclassical theory. *Econometrica, 33*, 685–760.

Chipman, J. S. (1966). A survey of the theory of international trade: Part 3, The modern theory. *Econometrica, 34*, 18–76.

Clement, M. O., Pfister, R. L., & Rothwell, K. J. (1967). *Theoretical issues in international economics*. Boston: Houghton Mifflin.

Clifton, D. S., Jr., & Marxsen, W. S. (1984). An empirical investigation of the Heckscher-Ohlin theorem. *Canadian Journal of Economics, 17*, 32–38.

Costinot, A. (2009). On the origins of comparative advantage. *Journal of International Economics, 77*, 255–264.

Combes, P.-P. (1997). Industrial agglomeration under cournot competition. *Annales d'Economie et de Statistique, 45*, 161–182.

Combes, P.-P., & Lafourcade, M. (2011). Competition, market access and economic geography: Structural estimations and predictions for France. *Regional Science and Urban Economics, 41*, 508–524.

Combes, P.-P., Lafourcade, M., & Mayer, T. (2005). The trade-creating effects of business and social networks: Evidence from France. *Journal of International Economics, 66*, 1–29.

Corden, W. M. (1966). The structure of a tariff system and the effective protective rate. *Journal of Political Economy, 74*, 221–237. (Reprinted in J. N. Bhagwati, Ed., 1969, pp. 284–308; and in J. N. Bhagwati, Ed., 1981, pp. 108–128)

Corden, W. M. (1969). Effective protective rates in the general equilibrium model: A geometric note. *Oxford Economic Papers, 21*, 135–141.

Corden, W. M. (1971a). *The theory of protection*. London: Oxford University Press.

Corden, W. M. (1971b). The effects of trade on the rate of growth. In J. N. Bhagwati et al. (Eds.), *Trade, balance of payments and growth* (pp. 117–143). Amsterdam: Nort-Holland.

Corden, W. M. (1974). *Trade policy and economic welfare*. Oxford, UK: Oxford University Press.

Corden, W. M. (1984a). Booming sector and Dutch disease economics: Survey and consolidation. *Oxford Economic Papers, 36*, 359–380.

Corden, W. M. (1984b). The normative theory of international trade. In R. W. Jones & P. B. Kenen (Eds.), *Handbook of international economics* (chap. 2). Amsterdam: North-Holland.

Corden, W. M., & Neary, J. P. (1982). Booming sector and de-industrialization in a small open economy. *Economic Journal, 92*, 825–848.

Cornes, R. (1992). *Duality and modern economics*. Cambridge, UK: Cambridge University Press.

Costinot, A., & Vogel, J. (2010). Matching and inequality in the world economy. *Journal of Political Economy, 118*, 747–786.

Crozet, M. (2004). Do migrants follow market potentials? An estimation of a new economic geography model. *Journal of Economic Geography, 4*, 439–458.

Crozet, M., & Trionfetti, F. (2008). Trade costs and the home market effect. *Journal of International Economics, 76*, 309–321.

Cuñat, A., & Melitz, M. (2007). Volatility, labour market flexibility, and the pattern of comparative advantage. *Journal of the European Economic Association, 10*, 225–254.

Cunynghame, H. (1904). *A geometrical political economy*. Oxford, UK: Clarendon.

Davies, S. W., & McGuinness, A. J. (1982). Dumping at less than marginal cost. *Journal of International Economics, 12*, 169–182.

Davis, D. R. (1995). Intra-industry trade: A Heckscher-Ohlin Ricardo approach. *Journal of International Economics, 39*, 201–226.

Davis, D. R. (1996). Critical evidence on comparative advantage? North-North trade in a multilateral world. *Journal of Political Economy, 102*, 1051–1060.

Davis, D. R. (1998). The home market, trade and industrial structure. *American Economic Review, 88*, 1264–1276.

Davis, D. R., & Weinstein, D. E. (1999). Economic geography and regional production structure, an empirical investigation. *European Economic Review, 43*, 379–407.

Davis, D. R., & Weinstein, D. E. (2001). An account of global factor trade. *American Economic Review, 91*, 1423–1453.

Davis, D. R., & Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review, 92*, 1269–1289.

Davis, D. R., & Weinstein, D. E. (2003). Market access, economic geography and comparative advantage, an empirical test. *Journal of International Economics, 59*, 1–23.

Davis, D. R., & Weinstein, D. E. (2008). A search for multiple equilibria in urban industrial structure. *Journal of Regional Science, 48*, 29–65.

Davis, D. R., Weinstein, D. E., Bradford, S. C., & Shimpo, K. (1997). Using international and Japanese data to determine when the factor abundance theory of trade works. *American Economic Review, 87*, 421–446.

Deardoff, A. V. (1979). Weak links in the chain of comparative advantage. *Journal of International Economics, 9*, 197–209. (Reprinted in J. N. Bhagwati, Ed., 1981, pp. 76–87)

Deardoff, A. V. (1982). The general validity of the Heckscher-Ohlin theorem. *American Economic Review, 72*, 683–694.

Deardoff, A. V. (1984). Testing trade theories and predicting trade flows. In R. W. Jones & P. B. Kenen (Eds.), *Handbook of international economics* (chap. 10). Amsterdam: North-Holland.

Deardoff, A. V. (1994). The possibility of factor price equalization, revisited. *Journal of International Economics, 36*, 167–175.

Deardorff, A. V. (1998). Determinants of bilateral trade, does gravity work in a neoclassical world? In J. A. Frankel (Ed.), *The regionalization of the world economy*. Chicago: University of Chicago Press and NBER.

Deardorff, A. V. (2000). Factor prices and the factor content of trade revisited: What's the use? *Journal of International Economics, 50*, 73–90.

Deardoff, A., & Stern, R. M. (1984). The effects of the Tokyo round on the structure of protection. In R. E. Baldwin & A. O. Krueger (Eds.), *The structure and evolution of recent U.S. trade policy* (pp. 361–388). Chicago: Chicago University Press for the NBER.

Debreu, G. (1959). *Theory of value*. New York: Wiley.

Diab, M. A. (1956). *The United States capital position in the structure of its foreign trade*. Amsterdam: North-Holland.

Diewert, W. E. (1974). Applications of duality theory. In M. D. Intriligator & D. A. Kendrick (Eds.), *Frontiers of quantitative economics* (chap. 3). Amsterdam: North-Holland.

Diewert, W. E. (1982). Duality approaches to microeconomic theory. In K. J. Arrow & M. D. Intriligator (Eds.), *Handbook of mathematical economics* (Vol. II, chap. 12). Amsterdam: North-Holland.

Dinopoulos, E. (1988). A formalization of the "Biological" model of trade in similar products. *Journal of International Economics, 25*, 95–110.

Dinopoulos, E., & Kreinin, M. (1989). Import quotas and VERs: A comparative analysis of a three-country framework. *Journal of International Economics, 26*, 169–178.

Dinopoulos, E., Syropoulos, C., Xu, B., & Yotov, Y. V. (2011). Intraindustry trade and the skill premium: Theory and evidence. *Journal of International Economics*, *84*, 15–25.

Dixit, A. K. (1981). The export of capital theory. *Journal of International Economics, 11*, 279–294.

Dixit, A. K. (1984). International trade policy for oligopolistic industries. In *Conference papers* (1983, Conference of the Royal Economic Society and the AUTE). *Supplement to the Economic Journal, 94*, 1–16.

Dixit, A. K., & Norman, V. (1980). *Theory of international trade: A dual, general equilibrium approach*. Cambridge, UK: Cambridge University Press.

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *American Economic Review, 67*, 297–308.

Dixit, A. K., & Woodland, A. (1982). The relationship between factor endowments and commodity trade. *Journal of International Economics, 13*, 201–214.

Djajić, S. (1983). Intermediate inputs and international trade: An analysis of the real and monetary aspects of a change in the price of oil. *Journal of International Money and Finance, 2*, 179–195.

Dorfman, R., Samuelson, P. A., & Solow, R. M. (1958). *Linear programming and economic analysis*. New York: McGraw-Hill.

Dornbusch, R., Fischer, S., & Samuelson, P. A. (1977). Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. *American Economic Review, 67*, 823–839.

Dreher, A. (2006). Does globalization affect growth? Evidence from a new index of globalization. *Applied Economics, 38*, 1091–1110.

Dreher, A., Gaston, N., & Martens, P. (2008). *Measuring globalization—Gauging its consequences*. New York: Springer.

Duesenberry, J. S. (1949). *Income, saving, and the theory of consumer behavior*. Cambridge, MA: Harvard University Press.

Dumas, B. (1980). The theorems of international trade under generalized uncertainty. *Journal of International Economics, 10*, 481–498.

Dunning, J. H. (1977). *Trade, location of economic activity and the MNE: A search for an eclectic approach*. In B. Ohlin, P. Hesselborn, & P. Wijkman (Eds.), *The international allocation of economic activity* (Proceedings of a nobel symposium held at Stockholm). London: Macmillan.

Dunning, J. H. (1982). International business in a changing world environment. *Banca Nazionale del Lavoro Quarterly Review, 143*, 351–374.

Eaton, J. (1979). The allocation of resources in an open economy with uncertain terms of trade. *International Economic Review, 20*, 391–403.

Eaton, J., & Kierzkowski, H. (1984). Oligopolistic competition, product variety and international trade. In H. Kierzkowski (Ed.), *Monopolistic competition and international trade* (pp. 69–83). Oxford, UK: Oxford University Press.

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica, 70*, 1741–1779.

Economides, N. (1984). Do increases in preference similarity (across countries) induce increases in trade? An affirmative example. *Journal of International Economics, 17*, 375–381.

Edgeworth, F. Y. (1894). The theory of international values, I, II, III. *Economic Journal, 4*, 35–50, 424–443, 606–638. (Reprinted in *Papers relating to political economy*, Vol. 11, pp. 3–60, by F. Y. Edgeworth, 1925, London: Macmillan)

Edgeworth, F. Y. (1905). Review of H. Cunynghame's book "A Geometrical Political Economy". *Economic Journal, 15*, 62–71. (Reprinted in F. Y. Edgeworth, 1925, op. cit., Vol. III, pp. 136–144)

El-Agraa, A. M. (1984). *Trade theory and policy*. London: Macmillan.

Ellis, H. S., & Metzler, L. A. (Eds.). (1949). *Readings in the theory of international trade*. Philadelphia: Blakiston.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy, 105*, 889–927.

Eltis, C. M. (1983). An alternative interpretation and empirical test of the linder hypothesis. *Quarterly Journal of Business and Economics, 22*, 53–62.

Epifani, P., & Gancia, G. (2008). The skill bias of world trade. *Economic Journal, 118*, 927–960.

Ethier, W. J. (1972). Nontraded goods and the Heckscher-Ohlin model. *International Economic Review, 13*, 132–147.

Ethier, W. J. (1979). The theorems of international trade in time-phased economies. *Journal of International Economics, 9*, 225–238.

Ethier, W. J. (1981). A reply to professors Metcalfe and Steedman. *Journal of International Economics, 11*, 273–277.

Ethier, W. J. (1982a). National and international returns to scale in the modern theory of international tade. *American Economic Review, 72*, 389–405.

Ethier, W. J. (1982b). The general role of factor intensity in the theorems of international trade. *Economics Letters, 10*, 337–342.

Ethier, W. J. (1984). Higher dimensional trade theory. In R. W. Jones & P. B. Kenen (Eds.) (chap. 3).

Ethier, W. J., & Horn, H. (1984). A new look at economic integration. In H. Kierzkowski (Ed.), *Monopolistic competition and international trade* (pp. 207–229). Oxford, UK: Oxford University Press.

Ethier, W. J., & Svensson, L. E. O. (1986). The theorems of international trade with factor mobility. *Journal of International Economics, 20*, 21–42.

Evenett, S. J., & Keller, W. (2002). On theories explaining the success of the gravity equation. *Journal of Political Economy, 110*, 281–316.

Falvey, R. E. (1981). Commercial policy and intra-industry trade. *Journal of International Economics, 11*, 495–511.

Falvey, R. E., & Kierzkowski, H. (1987). Product quality, intra-industry trade and (Im)perfect competition. In H. Kierzkowski (Ed.), *Protection and competition in international trade: Essays in honor of W.M. Corden* (pp. 143–161). Oxford, UK: Basil Blackwell.

Fausti, S. (1992). Smuggling and parallel markets for exports. *International Trade Journal, 6*, 443–470.

Feenstra, R. C., & Hanson, G. H. (1997). Foreign direct investment and relative wages: Evidence from Mexico's maquiladoras. *Journal of International Economics, 42*, 371–393.

Feenstra, R. C., & Hanson, G. H. (1999). The impact of outsourcing and high-technology capital on wages: Estimates for the U.S., 1979–1990. *Quarterly Journal of Economics, 114*, 907–940.

Feenstra, R. C., & Taylor, A. M. (2008). *International economics*. New York: Worth Publishers.

Findlay, R. (1970). *Trade and specialization*. Harmondsworth, UK: Penguin.

Findlay, R. (1984). Growth and development in trade models. In R. W. Jones & P. B. Kenen (Eds.), *Handbook of international economics* (chap. 4). Amsterdam: North Holland

Findlay, R. (1995). *Factor proportions, trade, and growth*. Cambridge, MA: MIT.

Findlay, R., & Grubert, H. (1959). Factor intensities, technological progress and the terms of trade. *Oxford Economic Papers, 11*, 111–121. (Reprinted in J. N. Bhagwati, Ed., 1969, pp. 327–340; and in J. N. Bhagwati, Ed., 1981, pp. 289–300)

Findlay, R., & Wellisz, S. (1983). Some aspects of the political economy of trade restrictions. *Kyklos, 36*, 469–481.

Fisher, F. M., & Hillman, A. L. (1984). The commodity pattern of trade and the Heckscher-Ohlin theorem in the presence of aggregate and commodity-specific factor-intensity reversals. *Journal of International Economics, 17*, 159–172.

Flam, H., & Helpman, E. (1987). Industrial trade policy under monopolistic competition. *Journal of International Economics, 22*, 79–102.

Ford, J. L. (1982). The Ricardian and Heckscher-Ohlin explanations of trade: A general proof of an equivalence theorem and its empirical implications. *Oxford Economic Papers, 34*, 141–145.

Ford, J. L. (1985). Comments on P.J. Lloyd "The Ricardian-Ohlin explanation of trade: A comment on a general theorem which is not general". *Oxford Economic Papers, 37*, 134–137.

Forslid, R., & Ottaviano, G. I. P. (2003). An analytically solvable core-periphery model. *Journal of Economic Geography, 3*, 229–240.

Frey, B. S. (1984). *International political economics*. Oxford, UK: Basil Blackwell.

Friedman, J. W. (1977). *Oligopoly and the theory of games*. Amsterdam: North-Holland.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics, 114*, 739–767.

Gabel, H. L., & Neven, D. (1988). *Fair trade in commercial aircraft: The case of Boeing vs. Airbus industry*. Fontainebleau, France: INSEAD.

Gabszewicz, J. J., Shaked, A., Sutton, J., & Thisse, J.-F. (1981). International trade in differentiated products. *International Economic Review, 22*, 527–534.

Gaisford, J. D. (1995). International factor mobility, the factor content of trade and Leontief paradoxes. *Journal of International Economics, 39*, 175–183.

Gale, D., & Nikaidô, H. (1965). The Jacobian matrix and global univalence of mappings. *Matematische Annalen, 159*, 81–93. (Reprinted in P. K. Newman, Ed., 1968, pp. 68–80)

Gandolfo, G. (1971). *Mathematical methods and models in economic dynamics*. Amsterdam: North-Holland.

Gandolfo, G. (2009). *Economic dynamics* (4th ed.). Berlin/Heidelberg, Germany; New York: Springer.

Garicano, L., & Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *Quarterly Journal of Economics, 121*, 1383–1435.

Ghosh, D. K. (1984). *Trade, distortion and growth*. New Delhi, India: Concept Publishing.

Giersch, H. (Ed.). (1979). *On the economics of intra-industry trade*. Tübingen, Germany: J.C.B. Mohr.

Goldberg, P., & Pavcnik, N. (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature, 45*, 39–82.

Graham, F. D. (1923). The theory of international values reexamined. *Quarterly Journal of Economics, 28*, 54–86. (Reprinted in H. S. Ellis & L. A. Metzler, Eds., 1949, pp. 301–330; and in W. R. Allen, Ed., 1965, chap. 6)

Greaney, T. M. (1996). Import now! An analysis of market-share Voluntary Import Expansions (VIEs). *Journal of International Economics, 40*, 149–163.

Green, H. A. J. (1964). *Aggregation in economic analysis*. Princeton, NJ: Princeton University Press.

Greenaway, D. (1983). *International trade policy: From tariffs to the new protectionism*. London: Macmillan.

Greenaway, D. (Ed.). (1985). *Current issues in international trade: Theory and policy*. London: Macmillan.

Greenaway, D., & Milner, C. (1986). *The economics of intra-industry trade*. Oxford, UK: Basil Blackwell.

Grinols, E. L. (1984). A thorn in the lion's paw: Has Britain paid too much for common market membership? *Journal of International Economics, 16*, 271–293.

Grinols, E. L. (1985). International trade with incomplete markets. *Economica, 52*, 245–255.

Grossman, G. M. (1983). Partially mobile capital: A general approach to two-sector trade theory. *Journal of International Economics, 15*, 1–17.

Grossman, G. M. (Ed.). (1992). *Imperfect competition and international trade*. Cambridge, MA: MIT.

Grossman, G. M., & Helpman, E. (1991a). *Innovation and growth in the global economy*. Cambridge, MA: MIT.

Grossman, G. M., & Helpman, E. (1991b). Endogenous product cycles. *Economic Journal, 101*, 1214–1229.

Grossman, G. M., & Helpman, E. (1991c). Quality ladders and product cycles. *Quarterly Journal of Economics, 106*, 557–586.

Grossman, G. M., & Helpman, E. (1991d). Growth and welfare in a small open economy. In E. Helpman & A. Razin (Eds.), *International trade and trade policy* (chap. 8). Cambridge, MA: MIT.

Grossman, G. M., & Razin, A. (1985). The pattern of trade in a Ricardian model with country-specific uncertainty. *International Economic Review, 26*, 193–202.

Grossman, G. M., & Richardson, J. D. (1985). *Strategic trade policy: A survey of the issues and early analysis* (Special papers in international economics, Vol.15). Princeton, NJ: International Finance Section, Princeton University.

Grossman, G. M., & Rossi-Hansberg, E. (2008). A simple theory of offshoring. *American Economic Review, 98*, 1978–1997.

Grubel, H. G. (1967). Intra-industry specialization and the pattern of trade. *Canadian Journal of Economics and Political Science, 23*, 347–388.

Grubel, H. G., & Lloyd, P. J. (1975). *Intra-industry trade: The theory and measurement of international trade in differentiated products*. London: Macmillan.

Guerrieri, P., & Padoan, P. C. (1996). Strategic trade policy: A review of the theoretical debate. *International Trade Journal, 10*, 85–103.

Gupta, S. D. (Ed.). (1997). *The political economy of globalization*. Boston/Dordrecht, The Netherlands/London: Kluwer.

Haberler, G. (1936). *The theory of international trade with its applications to commercial policy*. London: Hodge & Co. (English translation, revised by the author, of the 1933 German edition)

Haberler, G. (1990). Strategic trade policy and the new international economics: A critical analysis. In R. W. Jones & A. O. Krueger (Eds.), *The political economy of international trade* (pp. 25–30). Oxford, UK: Basil Blackwell.

Hadley, G., & Kemp, M. C. (1966). Equilibrium and efficiency in international trade. *Metroeconomica, 18*, 125–141.

Hahn, F. (1982). The neo-Ricardians. *Cambridge Journal of Economics, 6*, 353–374. (Reprinted in *Equilibrium and macroeconomics*, pp. 353–386, by F. Hahn, 1984, Oxford, UK: Blackwell)

Hamilton, C., & Svensson, L. E. O. (1983). Should direct or total factor intensities be used in tests of the factor proportions hypothesis? *Weltwirtschaftliches Archiv, 119*, 453–463.

Hamilton, C., & Svensson, L. E. O. (1984). Do countries' factor endowments correspond to the factor contents in their bilateral trade flows? *Scandinavian Journal of Economics, 86*, 84–97.

Hanink, D. M. (1990). Linder, again. *Weltwirtschaftliches Archiv, 126*, 257–267.

Hanson, G. H., & Xiang, C. (2004). The home market effect and bilateral trade patterns. *American Economic Review, 94*, 1108–1129.

Harcourt, G. C. (1972). *Some Cambridge controversies in the theory of capital*. Cambridge, UK: Cambridge University Press.

Harkness, J. P. (1978). Factor abundance and comparative advantage. *American Economic Review, 68*, 784–800.

Harkness, J. P. (1983). The factor-proportions model with many nations, goods and factors: Theory and evidence. *Review of Economics and Statistics, 65*, 293–305.

Harrigan, J. (1994). Scale economies and the volume of trade. *Review of Economics and Statistics, 76*, 321–328.

Harrigan, J. (1996). Openness to trade in manufactures in the OECD. *Journal of International Economics, 40*, 23–39.

Harrigan, J. (1997). Technology, factor supplies, and international specialization: Examining the neoclassical model. *American Economic Review, 87*, 475–494.

Harris, R. (1984). Applied general equilibrium analysis of small open economies with scale economies and imperfect competition. *American Economic Review, 74*, 1016–1032.

Harrod, R. F. (1953). Factor-price relations under free trade. *Economic Journal, 68*, 245–255.

Hartquist, D. A. (1987). Trade wars—Arming for battle. *North Carolina Journal of International Law and Commercial Regulation, 12*, 37–58.

Hartwick, J. M. (1979). Distribution of world income in the Ricardo-Mill model of international trade. *Journal of International Economics, 9*, 117–126.

Hatta, T. (1984). Immiserizing growth in a many-commodity setting. *Journal of International Economics, 17*, 335–345.

Hazari, B. R., & Sgro, P. M. (2001). *Migration, unemployment and trade*. Dordrecht, The Netherlands: Kluwer.

Hazari, B. R., Sgro, P. M., & Suh, D. C. (1981). *Non-traded and intermediate goods and the pure theory of international trade*. London: Croom-Helm.

Head, K., & Mayer, T. (2004). The empirics of agglomeration and trade. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of urban and regional economics* (Vol. 4, chap. 59). Amsterdam: North Holland.

Head, K., & Mayer, T. (2006). Regional wages and employment responses to market potential in the EU. *Regional Science and Urban Economics, 36*, 573–595.

Head, K., & Mayer, T. (2013). Gravity equations: Workhorse, toolkit, and cookbook. In E. Helpman, G. Gopinath, & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 4). Amsterdam: Elsevier (forthcoming).

Head, K., & Ries, J. (2001). Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade. *American Economic Review, 91*, 858–876.

Head, K., Mayer, T., & Ries, J. (2002). On the pervasiveness of home market effects. *Economica, 69*, 357–370.

Heckscher, E. F. (1949), 1919. The effect of foreign trade on the distribution of income. In H. S. Ellis & L. A. Metzler (Eds.), pp. 272–300, 1949. (English translation of the original 1919 article in *Ekonomisk Tidschrift, 21*, 497–512)

Helpman, E. (1981). International trade in the presence of product differentiation, economies of scale and monopolistic competition: A Chamberlin-Heckscher-Ohlin approach. *Journal of International Economics, 11*, 305–340.

Helpman, E. (1984a). The factor content of international trade. *Economic Journal, 94*, 84–94.

Helpman, E. (1984b). Increasing returns, imperfect markets, and trade theory. In R. W. Jones & P. B. Kenen (Eds.) (chap. 7).

Helpman, E. (1987). Imperfect competition and international trade: Evidence from fourteen industrialized countries. *Journal of the Japanese and International Economies, 1*, 62–81.

Helpman, E. (1990). *Monopolistic competition in trade theory* (Special paper in international economics, Vol. 16). Princeton, NJ: Princeton University, International Finance Section.

Helpman, E. (1998). The size of regions. In D. Pines, E. Sadka, & I. Zilcha (Eds.), *Topics in public economics: Theoretical and applied analysis* (pp. 33–54). Cambridge, UK: Cambridge University Press.

Helpman, E., Itskhoki, O., & Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica, 78*, 1239–1283.

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade*. Brighton, Sussex: Wheatsheaf Books.

Helpman, E., & Krugman, P. R. (1989). *Trade policy and market structure*. Cambridge, MA: MIT Press.

Helpman, E., & Razin, A. (1978). *A theory of international trade under uncertainty*. New York: Academic.

Helpman, E., & Razin, A. (Eds.). (1991). *International trade and trade policy*. Cambridge, MA: MIT Press.

Henderson, J. M., & Quandt, R. E. (1980). *Microeconomic theory: A mathematical approach* (3rd ed.). New York: McGraw-Hill.

Herberg, H. (1969). On the shape of the transformation curve in the case of homogeneous production functions. *Zeitschrift für die gesamte Staatswissenschaff, 125*, 202–210.

Herberg, H. (1990). Welfare effects of non-tariff barriers: A general equilibrium analysis. *Weltwirtschaftliches Archiv, 126*, 511–522.

Herberg, H., Kemp, M. C., & Tawada, M. (1982). Further implications of variable returns to scale. *Journal of International Economics, 13*, 65–84.

Hicks, J. R. (1932). *The theory of wages*. London: Macmillan. (2nd ed.: 1963)

Hicks, J. R. (1939, 1946). *Value and capital*. Oxford, UK: Oxford University Press.

Hicks, J. R. (1981). *Wealth and welfare* (Collected essays on economic theory, Vol. 1). Oxford, UK: Basil Blackwell.

Hill, J. K., & Mendez, J. A. (1983). Factor mobility and the general equilibrium model of production. *Journal of International Economics, 15*, 18–26.

Hillman, A. L., & Ursprung, H.-W. (1988). Domestic politics, foreign interests, and international trade policy. *American Economic Review, 7*, 719–745.

Hirsch, S. (1967). *Location of industry and international competitiveness*. London: Oxford University Press.

Hirsch, S. (1975). The product cycle model of international trade: A multi-country cross section analysis. *Oxford Bulletin of Economics and Statistics, 37*, 305–317.

Hirshleifer, J. (1970). *Investment, interest, and capital*. Englewood Cliffs, NJ: Prentice-Hall.

Hocking, R. D. (1980). Trade in motor cars between the major European producers. *Economic Journal, 90*, 504–519.

Hoff, K. (1994). A reexamination of the neoclassical trade model under uncertainty. *Journal of International Economics, 36*, 1–27.

Horiba, Y. (1974). General equilibrium and the Heckscher-Ohlin theory of trade: The multi-country case. *International Economic Review, 15*, 440–449.

Hotelling, H. (1929). Stability in competition. *Economic Journal, 34*, 41–57.

Houthakker, H. S. (1957). An international comparison of household expenditure patterns, commemorating the centenary of Engel's law. *Econometrica, 25*, 532–551.

Hufbauer, G. C. (1966). *Synthetic materials and the theory of international trade*. Cambridge, MA: Harvard University Press.

Hummels, D., & Levinsohn, J. A. (1995). Monopolistic competition and international trade: Reconsidering the evidence. *Quarterly Journal of Economics, 110*, 799–836.

Jensen, B. S., & Wong, K.-Y. (Eds.). (1998). *Dynamics, economic growth, and international trade*. Ann Arbor, MI: University of Michigan Press.

Johnson, H. G. (1950). Optimum welfare and maximum revenue tariffs. *Review of Economic Studies, 19*, 28–35. (Reprinted (in part) as Appendix to chap. II in H. G. Johnson, 1958)

Johnson, H. G. (1953). Optimum tariffs and retaliation. *Review of Economic Studies, XXI*, 142–153. (Reprinted in H. G. Johnson, 1958, chap. II)

Johnson, H. G. (1955). Economic expansion and international trade. *The Manchester School of Economic and Social Studies, 23*, 95–112. (Reprinted in H. G. Johnson, 1958, chap. III)

Johnson, H. G. (1957). Factor endowments, international trade and factor prices. *Manchester School of Economic and Social Studies, 25*, 270–283. (Reprinted in H. G. Johnson, 1958, pp. 17–30; and in R. E. Caves & H. G. Johnson, Eds., 1969, pp. 78–89)

Johnson, H. G. (1958). *International Trade and Economic Growth*. London: Allen & Unwin.

Johnson, H. G. (1959). Economic development and international trade. *Nationaløkonomisk Tidsskriff, 97*, 253–272. (Reprinted in H. G. Johnson, 1962, chap. IV)

Johnson, H. G. (1960). The economic theory of customs union. *Pakistan Economic Journal, 10*, 14–32. (Reprinted (with appendix) in H. G. Johnson, 1962, chap. III)

Johnson, H. G. (1962). *Money, trade and economic growth*. London: Allen & Unwin.

Johnson, H. G. (1969). The standard theory of tariffs. *Canadian Journal of Economics, 2*, 333–352.

Johnson, H. G. (1971a). Trade and growth: A geometric exposition. *Journal of International Economics, 1*, 83–101.

Johnson, H. G. (1971b). The theory of trade and growth: A diagrammatic analysis. In J. N. Bhagwati et al. (Eds.) (pp. 144–167).

Jones, A. J. (1979). The theory of economic integration. In J. K. Bowers (Ed.), *Inflation, development and integration* (Essays in honour of A. J. Brown, pp. 193–213). Leeds, UK: Leeds University Press.

Jones, K. (1984). The political economy of voluntary export restraint agreements. *Kyklos, 37*, 82–101.

Jones, R. W. (1956). Factor proportions and the Heckscher-Ohlin theorem. *Review of Economic Studies, 24*, 1–10. (Reprinted in R. W. Jones, 1979, chap. 1; and in J. N. Bhagwati, Ed., 1969, pp. 77–92)

Jones, R. W. (1961). Comparative advantage and the theory of tariffs: A multi-country multi-commodity model. *Review of Economic Studies, 28*, 161–175. (Reprinted in R. W. Jones, 1979, chap. 3)

Jones, R. W. (1965). The structure of simple general equilibrium models. *Journal of Political Economy, 73*, 557–572. (Reprinted in R. W. Jones, 1979, chap. 4; and in J. N. Bhagwati, Ed., 1981, chap. 3)

Jones, R. W. (1967). International capital movements and the theory of tariffs and trade. *Quarterly Journal of Economics, 81*, 1–38. (Reprinted in R. W. Jones, 1979, chap. 12)

Jones, R. W. (1971). A three-factor model in theory, trade, and history. In J. N. Bhagwati et al. (Eds.) (pp. 3–21). (Reprinted in R. W. Jones, 1979, chap. 6)

Jones, R. W. (1979). *International trade: Essays in theory*. Amsterdam: North-Holland.

Jones, R. W., & Dei, F. (1983). International trade and foreign investment: A simple model. *Economic Inquiry, 21*, 449–464.

Jones, R. W., & Kenen, P. B. (Eds.). (1984). *Handbook of international economics* (Vol. I). Amsterdam: North-Holland.

Jones, R. W., & Krueger, A. O. (Eds.). (1990). *The political economy of international trade*. Oxford, UK: Basil Blackwell.

Jones, R. W., & Neary, J. P. (1984). The positive theory of international trade. In R. W. Jones & P. B. Kenen (Eds.) (chap. 1).

Judd, K. L. (1998). *Numerical methods in economics*. Cambridge, MA: MIT Press.

Kamien, M. I., & Schwartz, N. L. (1991). *Dynamic optimization* (2nd ed.). Amsterdam: North-Holland.

Kemp, M. C. (1964). *The pure theory of international trade*. Englewood Cliffs, NJ: Prentice Hall.

Kemp, M. C. (1969a). *A contribution to the general equilibrium theory of preferential trading*. Amsterdam: North-Holland.

Kemp, M. C. (1969b). *The pure theory of international trade and investment*. Englewood Cliffs, NJ: Prentice Hall.

Kemp, M. C. (1976). *Three topics in the theory of international trade: Distribution, welfare and uncertainty*. Amsterdam: North-Holland.

Kemp, M. C., & Long, N. V. (1984). The role of natural resources in trade models. In R. W. Jones & P. B. Kenen (Eds.) (chap. 8).

Kierzkowski, H. (Ed.). (1984). *Monopolistic competition and international trade*. Oxford, UK: Oxford University Press.

Kierzkowski, H. (1985). Models of international trade in differentiated goods. In D. Greenaway (Ed.) (pp. 7–24).

Kierzkowski, H. (Ed.). (1987). *Protection and competition in international trade: Essays in honor of W.M. Corden*. London: Basil Blackwell.

Kind, H. J., Midelfart-Knarvik, K. H., & Schjelderup, G. (2000). Competing for capital in a 'Lumpy' world. *Journal of Public Economics, 78*, 253–274.

Kleinman, E., & Kop, Y. (1984). Who trades with whom—The income pattern of international trade. *Weltwirtschaftliches Archiv, 120*, 499–521.

Kohler, W. K. (1988). Modeling Heckscher-Ohlin comparative advantage in regression equations: A critical survey. *Empirica* (Austrian Economic Papers), *15*, 263–293.

Kol, J., & Tharakan, P. K. M. (1989). *Intra-industry trade: Theory, evidence and extensions*. New York: St. Martin's Press.

Komiya, R. (1967). Non-traded goods and the pure theory of international trade. *International Economic Review, 8*, 132–152.

Kowalczyk, C., & Wonnacott, R. J. (1992). *Hubs and spokes, and free trade in the americas* (NBER Working Paper 4198), Cambridge, MA: National Bureau of Economic Research.

Krauss, M. B. (1972). Recent developments in customs union theory: An interpretive survey. *Journal of Economic Literature, 10*, 413–436.

Krauss, M. B. (1979). *A geometric approach to international trade*. Oxford, UK: Basil Blackwell.

Kravis, I. B. (1956). "Availability" and other influences on the commodity composition of trade. *Journal of Political Economic, 64*, 143–155.

Kreinin, M. E. (1965). Comparative labor effectiveness and the Leontief scarce factor paradox. *American Economic Review, 55*, 131–140.

Kreps, D. M. (1990). *Microeconomic theory*. Princeton, NJ: Princeton University Press.

Krueger, A. O. (1974). The political economy of the rent-seeking society. *American Economic Review, 64*, 291–303.

Krueger, A. O. (2000). NAFTA's effects: A preliminary assessment. *World Economy, 23*, 761–775.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics, 9*, 469–479. (Reprinted in J. N. Bhagwati, Ed., 1981, pp. 88–100)

Krugman, P. R. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review, 70*, 950–959.

Krugman, P. R. (1981). Intraindustry specialization and the gains from trade. *Journal of Political Economy, 89*, 959–974.

Krugman, P. R. (1983). New theories of trade among industrial countries. *American Economic Review, 73*, 343–347 (Papers and Proceedings).

Krugman, P. R. (1990). *Rethinking international trade*. Cambridge, MA: MIT Press.

Krugman, P. R. (1991a). *Geography and trade*. Cambridge, MA: MIT Press.

Krugman, P. R. (1991b). Increasing returns and economic geography. *Journal of Political Economy, 99*, 483–499.

Krugman, P. R. (1991). History versus expectations. *Quarterly Journal of Economics, 106*, 651–667.

Krugman, P. R. (1993a). On the relationship between trade theory and location theory. *Review of International Economics, 1*, 110–122.

Krugman, P. R. (1993b). The hub effect: Or, threeness in interregional trade. In W. J. Ethier, E. Helpman, & J. P. Neary (Eds.), *Theory, policy and dynamics in international trade*. Cambridge, UK: Cambridge University Press.

Krugman, P. R. (1997). What should trade negotiators negotiate about? *Journal of Economic Literature, 35*, 113–120.

Krugman, P. R. (2000). Technology, trade and factor prices. *Journal of International Economics, 50*, 51–71.

Krugman, P. R., Obstfeld, M., & Melitz, M. J. (2011). *International economics: Theory and policy* (9th ed.). New York: Addison-Wesley.

Krugman, P. R., & Venables, A. J. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics, 110*, 857–880.

Laffer, A. B., & Miles, M. A. (1982). *International economics in an integrated world*. Glenview, IL: Scott, Foresman.

Lai, E. L. C. (1995). The product cycle and the world distribution of income: A reformulation. *Journal of International Economics, 39*, 369–382.

Laird, S., & Yeats, A. (1990). Trends in non-tariff barriers of developed countries, 1966–86. *Weltwirtschaftliches Archiv, 126*, 299–325.

Lancaster, K. (1957). The Heckscher-Ohlin trade model: A geometric treatment. *Economica, 24*, 19–39. (Reprinted in J. N. Bhagwati, Ed., 1969, pp. 49–74)

Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*, 132–157.

Lancaster, K. (1971). *Consumer demand: A new approach*. New York: Columbia University Press.

Lancaster, K. (1980). Intra-industry trade under perfect monopolistic competition. *Journal of International Economics, 10*, 151–175.

Lane, J. (1985). An empirical estimate of the effects of labor-market distortions on the factor content of U.S. trade. *Journal of International Economics, 18*, 187–193.

Leamer, E. E. (1980). The Leontief paradox, reconsidered. *Journal of Political Economy, 80*, 495–503.

Leamer, E. E. (1984). *Sources of comparative advantage: Theory and evidence*. Cambridge, MA: MIT Press.

Leamer, E. E. (2000). What's the use of factor contents? *Journal of International Economics, 50*, 17–49.

Leamer, E. E., & Levinsohn, J. (1995). International trade theory: The evidence. In G. M. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. III). Amsterdam: North-Holland.

Leontief, W. W. (1953). Domestic production and foreign trade: The American position reexamined. *Proceedings of the American Philosophical Society, 97*, 332–349. (Reprinted in *Input-output economics*, by W. Leontief, 1966, chap. 5 (pages refer to this reprint) New York: Oxford University Press; and in J. N. Bhagwati, Ed., 1969, pp. 93–139)

Leontief, W. W. (1956). Factor proportions and the structure of American trade: Further theoretical and empirical analysis. *Review of Economics and Statistics, 38*, 386–407. (Reprinted in W. Leontief, 1966, *Input-Output Economics*, op. cit., chap. 6)

Lerner, A. P. (1936). The symmetry between import and export taxes. *Economica, 3*, 306–313. (Reprinted in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 197–203)

Lerner, A. P. (1952), 1933. Factor prices and international trade. *Economica, 19*, 1–15. (This article reproduces the text of a paper prepared for a seminar at the London School of Economics in 1933, as it was originally written). Reprinted in *Essays in economic analysis*, pp. 67–84, by A. P. Lerner, 1953, London: Macmillan.

Levchenko, A. A. (2007). Institutional quality and international trade. *Review of Economic Studies, 74*, 791–819.

Lindbeck, A. (Ed.). (1979). *Inflation and employment in open economies*. Amsterdam: North-Holland.

Linder, S. B. (1961). *An essays on trade and transformation*. New York: Wiley.

Lipsey, R. G. (1960). The theory of customs unions: A general survey. *Economic Journal, 70*, 496–513. (Reprinted in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 261–278; in J. N. Bhagwati, Ed., 1969, pp. 218–241; and in J. N. Bhagwati, Ed., 1981, pp. 265–282)

Lipsey, R. G. (1970). *The theory of customs unions: A general equilibrium analysis*. London: Weidenfeld and Nicolson.

Lipsey, R. G., & Lancaster, K. (1956). The general theory of second best. *Review of Economic Studies, 24*, 11–32.

Lloyd, P. J. (1974). A more general theory of price distortions in open economies. *Journal of International Economics, 4*, 365–386.

Lloyd, P. J. (1982). 3×3 theory of customs unions. *Journal of International Economics, 12*, 41–63.

Lloyd, P. J. (1985). The Ricardian and Heckscher-Ohlin explanations of trade: A comment on a general theorem which is not general. *Oxford Economic Papers, 37*, 138–141.

Long, N. V. (1983). The effects of a booming export industry on the rest of the economy. *Economic Record, 59*, 57–60.

Long, N. V., & Wong, K.-Y. (1998). Endogenous growth and international trade: A survey. In B. S. Jensen & K.-Y. Wong (Eds.), *Dynamics, economic growth, and international trade*. Ann Arbor, MI: University of Michigan Press.

Lopez, R. A., & Pagoulatos, E. (1994). Rent seeking and the welfare costs of trade barriers. *Public Choice, 79*, 149–160.

Lösch, A. (1954), 1940. *Economics of location*. New Haven, CT: Yale University Press (English translation of 1940 German original).

Ludema, R. D., & Wooton, I. (2000). Economic geography and the fiscal effects of regional integration. *Journal of International Economics, 52*, 331–357.

Magee, S. P. (1980). *International trade*. Reading, MA: Addison-Wesley.

Mainwaring, L. (1984). *Value and distribution in capitalist economies*. Cambridge, UK: Cambridge University Press.

Mainwaring, L. (1988). A neo-ricardian analysis of international trade. In I. Steedman (Ed.), *Sraffian economics* (Vol. 2, pp. 201–217). Aldershot, UK: Elgar.

Mainwaring, L. (1991). *Dynamics of uneven development* (pp. 201–217). Aldershot, UK: Elgar.

Manasse, P., & Turrini, A. (2001). Trade, wages, and 'Superstars'. *Journal of International Economics, 54*, 97–117.

Markusen, J. R. (1981). Trade and the gains from trade with imperfect competition. *Journal of International Economics, 11*, 531–551.

Markusen, J. R., & Melvin, J. R. (1979). Tariffs, capital mobility, and foreign ownership. *Journal of International Economics, 9*, 395–409.

Markusen, J. R., Melvin, J. R., Kaempfer, W. H., & Maskus, K. E. (1995). *International trade: Theory and evidence*. New York: McGraw-Hill.

Markusen, J. R., & Svensson, L. E. O. (1985). Trade in goods and factors with international differences in technology. *International Economic Review, 26*, 175–192.

Markusen, J. R., & Venables, A. J. (1988). Trade policy with increasing returns and imperfect competition: Contradictory results from competing assumptions. *Journal of International Economics, 24*, 299–316.

Marshall, A. (1879, 1930). *The pure theory of foreign trade*. (Privately circulated in 1879; reprinted in 1930 and subsequently by London School of Economics, Reprints of Scarce Tracts on Political Economy, London, Lund Humphries)

Marshall, A. (1923). *Money, credit and commerce*. London: Macmillan.

Martin, J. P., & Neary, J. P. (1980). Variable labour supply and the pure theory of international trade: An empirical note. *Journal of International Economics, 10*, 549–559.

Martin, J. P., & Rogers, C. A. (1995). Industrial location and public infrastructure. *Journal of International Economics, 39*, 335–351.

Martin, L., & Panagariya, A. (1984). Smuggling, trade and price disparity: A crime-theoretic approach. *Journal of International Economics, 17*, 201–217.

Martin, P., & Ottaviano, G. I. P. (1999). Growing locations: Industry location in a model of endogenous growth. *European Economic Review, 43*, 281–302.

Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review, 42*, 947–968.

Martin, S. (1993). *Advanced industrial economics*. Oxford, UK: Blackwell.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford, UK/New York: Oxford University Press.

Maskus, K. E. (1985). A test of the Heckscher-Ohlin-Vanek theorem: The Leontief commonplace. *Journal of International Economics, 19*, 201–212.

Mayer, W. (1993). Lobbying for tariff policies. *Review of International Economics, 1*, 221–233.

McDermott, J. (1985). The Rybczynski theorem: Short-run adjustment. *Southern Economic Journal, 51*, 241–249.

McKenzie, L. (1954a). On equilibrium in Graham's model of world trade and other competitive systems. *Econometrica, 22*, 147–161.

McKenzie, L. (1954b). Specialization and efficiency in world production. *Review of Economic Studies, 21*, 165–180.

McKenzie, L. (1955). Specialization in production and the production possibility locus. *Review of Economic Studies, 23*, 56–64.

McMillan, J. (1989). *Game theory in international economics*. Chur, Switzerland/New York: Harwood Academic Publishers.

Meade, J. E. (1952). *A geometry of international trade*. London: Allen & Unwin.

Meade, J. E. (1955). *The theory of international economic policy* (Trade and welfare, Vol. 2). London/New York: Oxford University Press.

Melitz, M. (2003). The impact of trade on intra-industry reallocation and aggregate industry productivity. *Econometrica, 71*, 1695–1725.

Messerlin, P. A. (1981). The political economy of protectionism: The bureaucratic case. *Weltwirtschaftliches Archiv, 117*, 469–496.

Metcalfe, J. S., & Steedman, I. (1981). On the transformation of theorems. *Journal of International Economics, 11*, 267–271.

Metzler, L. A. (1949). Tariffs, the terms of trade, and the distribution of national income. *Journal of Political Economy, 57*, 1–29. (Reprinted in *Collected papers*, pp. 159–197, by L. A. Metzler, 1973, Harvard University Press; and in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 24–57)

Mill, J. S. (1848). *Principles of political economy* (1st ed.). London: Parker. (Reprint of the 9th ed. by W. J. Ashley Ed., 1909, London: Longmans Green; reproduced in facsimile 1961, New York: Kelley)

Minhas, B. S. (1962). The homohypallagic production function, factor intensity reversals, and the Heckscher-Ohlin theorem. *Journal of Political Economy, 70*, 138–156. (Reprinted in J. N. Bhagwati, Ed., 1969, pp. 140–168)

Moore, M. O., & Suranovic, S. M. (1992). Lobbying vs. administered protection: Endogenous industry choice and national welfare. *Journal of International Economics, 32*, 289–303.

Morey, E. S. (1984). Confuser surplus. *American economic review, 74*, 163–173.

Morkre, M. E., & Kelly, K. H. (1994). *Effects of unfair imports on domestic industries: US antidumping and countervailing duties cases, 1980 to 1988*. Washington, DC: Bureau of Economics, Federal Trade Commission.

Mosak, J. L. (1944). *General-equilibrium theory in international trade* (Cowles Commission Monograph No.7). Bloomington, IN: Principia Press.

Motta, M. (1992). Sunk costs and trade liberalisation. *Economic Journal, 102*, 578–587.

Mundell, R. A. (1957a). A geometry of transport costs in international trade theory. *Canadian Journal of Economics and Political Science, 23*, 331–348. (Reprinted in R. A. Mundell, 1968, chap. 5)

Mundell, R. A. (1957b). International trade and factor mobility. *American Economic Review, 47*, 321–335. (Reprinted in R. A. Mundell, 1968, chap. 6; in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 101–114; and in J. N. Bhagwati, Ed., 1981, pp. 17–20)

Mundell, R. A. (1968). *International economics*. New York: Macmillan.

Murata, Y. (2003). Product diversity, taste heterogeneity, and geographic distribution of economic activities: Market vs. non-market interactions. *Journal of Urban Economics, 53*, 126–144.

Murray, T., Schmidt, W., & Walter, I. (1978). Alternative forms of protection against market disruption. *Kyklos, 31*, 624–637.

Myint, H. (1958). The "Classical Theory" of international trade and the underdeveloped countries. *Economic Journal, 68*, 317–337.

Neary, J. P. (1984). The Heckscher-Ohlin model as an aggregate. In A. Ingham & A. M. Ulph (Eds.), *Demand, equilibrium and trade: Essays in honour of Ivor F. Pearce* (pp. 57–76). London: Macmillan.

Neary, J. P. (1985a). The observational equivalence of the Ricardian and Heckscher-Ohlin explanations of trade patterns. *Oxford Economic Papers, 37*, 142–147.

Neary, J. P. (1985b). Two-by-two international trade theory with many goods and factors. *Econometrica, 53*, 1233–1247.

Neary, J. P. (1991). Export subsidies and price competition. In E. Helpman & A. Razin (Eds.), *International trade and trade policy* (chap. 5). Cambridge, MA: MIT Press.

Neary, J. P. (1995). Factor mobility and international trade. *Canadian Journal of Economics, 28*(Special Issue), S4–S23.

Neary, J. P. (2001). Of hype and hyperbolas: Introducing the new economic geography. *Journal of Economic Literature, 39*, 536–561.

Neary, J. P. (2002). Foreign competition and wage inequality. *Review of International Economics, 10*, 680–693.

Neary, J. P. (2009). *International trade in general oligopolistic equilibrium*. Oxford University, mimeo.

Neary, J. P., & Tharakan, J. (2012). International trade with endogenous mode of competition in general equilibrium. *Journal of International Economics, 86*, 118–132.

Negishi, T. (1972). *General equilibrium theory and international trade*. Amsterdam: North-Holland.

Negishi, T. (1982). The labour theory of value in the Ricardian theory of international trade. *History of Political Economy, 14*, 199–210.

Newman, P. K. (Ed.). (1968). *Readings in mathematical economics* (Vol. I). Baltimore: Johns Hopkins University Press.

Nikaidô, H. (1956). On the classical multilateral exchange problem. *Metroeconomica, 8*, 135–145. (Reprinted in P. K. Newman, Ed., 1968, pp. 116–126)

Nikaidô, H. (1957). A supplementary note to "On the Classical Multilateral Exchange Problem". *Metroeconomica, 9*, 209–210. (Reprinted in P. K. Newman, Ed., 1968, pp. 127–128)

Nocco, A. (2005). The rise and fall of regional inequalities with technological differences and knowledge spillovers. *Regional Science and Urban Economics, 35*, 542–569.

Norman, V. D., & Venables, A. J. (1995). International trade, factor mobility, and trade costs. *Economic Journal, 105*, 1488–1504.

Norton, D. A. G. (1988). On the economic theory of smuggling. *Economica, 55*, 107–118.

Nunn, N. (2007). Relation-specificity, incomplete contracts and the pattern of trade. *Quarterly Journal of Economics, 122*, 569–600.

Ohlin, B. (1933). *Interregional and international trade*. Cambridge, MA: Harvard University Press (Rev. ed., 1967). (Excerpts are contained in W. R. Allen, Ed., 1965, chap. 7)

Ohlin, B., Hesselborn, P., & Wijkman, P. M. (Eds.). (1977). *The international allocation of economic activity*. London: Macmillan.

Olson, M. (1964). *The logic of collective action*. Cambridge, MA: Harvard University Press.

Oniki, H., & Uzawa, H. (1965). Patterns of trade and investment in a dynamic model of international trade. *Review of Economic Studies, 32*, 15–38.

Ottaviano, G. I. P. (1999). Integration, geography, and the burden of history. *Regional Science and Urban Economics, 29*, 245–256.

Ottaviano, G. I. P. (2001). Monopolistic competition, trade, and endogenous spatial fluctuations. *Regional science and urban economics, 31*, 51–77.

Ottaviano, G. I. P., & Robert-Nicoud, F. L. (2006). The 'Genome' of NEG models with vertical linkages: A positive and normative synthesis. *Journal of Economic Geography, 6*, 113–139.

Ottaviano, G. I. P., Tabuchi, T., & Thisse, J.-F. (2002). Agglomeration and trade revisited. *International Economic Review, 43*, 409–435.

Ottaviano, G. I. P., & van Ypersele, T. (2005). Market size and tax competition. *Journal of International Economics, 67*, 25–46.

Padoan, P. C. (1977). Una nota sulla distinzione fra beni "commerciabili" e beni "non commerciabili". *Note Economiche, 3*, 90–102.

Panagariya, A. (2000). Evaluating the factor-content approach to measuring the effect of trade on wage inequality. *Journal of International Economics, 50*, 91–116.

Pareto, V. (1906). *Manuale d'economia politica*. Milano: Societa' editrice libraria. (Reprinted, with translation of the mathematical Appendix of the French edition, 1965, Roma, Bizzarri; French translation (with new mathematical Appendix): V. Pareto, 1909, *Manuel d'économie politique*, Paris, Giard et Brière; English translation from the French edition: V. Pareto, 1971, *Manual of Political Economy*, New York: A.M. Kelley)

Parrinello, S. (1970). Introduzione a una teoria neoricardiana del commercio internazionale. *Studi Economici, 25*, 267–321.

Parrinello, S. (1979). Distribution, growth, and international trade. In I. Steedman (Ed.) (pp. 159–187).

Parrinello, S. (1988). On foreign trade and the Ricardian model of trade. *Journal of Post Keynesian Economics, 10*, 585–601.

Pasinetti, L. (1977). *Lectures on the theory of production*. New York: Columbia University Press.

Pasinetti, L. (1981). *Structural change and economic growth—A theoretical essays on the dynamics of the wealth of nations*. Cambridge, UK: Cambridge University Press.

Pearce, I. F. (1952). The factor price equalization myth. *Review of Economic Studies, 19*, 111–120. (With mathematical appendix in collaboration with S. F. James)

Pearce, I. F. (1970). *International trade*. London: Macmillan.

Petith, H. C. (1974). The pattern of trade in a neoclassical growth model with a general saving function. *Journal of International Economics, 4*, 387–393.

Philpot, G. (1970). Labor quality, returns to scale and the elasticity of factor substitution. *Review of Economics and Statistics, 52*, 194–199.

Pinto, B. (1986). Repeated games and the "Reciprocal Dumping" model of trade. *Journal of International Economics, 20*, 357–366.

Pitt, M. A. (1981). Smuggling and price disparity. *Journal of International Economics, 11*, 447–458.

Pollak, R. A. (1971). Additive utility function and linear engel curves. *Review of Economic Studies, 38*, 221–230.

Pomery, J. (1979). Uncertainty and international trade. In R. Dornbusch & J. Frenkel (Eds.), *International economic policy: Theory and evidence* (chap. 4). Baltimore: Johns Hopkins University Press.

Pomery, J. (1984). Uncertainty in trade models. In R. W. Jones & P. B. Kenen (Eds.) (chap. 9).

Pomfret, R. (1989). The economics of voluntary export restraint agreements. *Journal of Economic Surveys, 3*, 199–211.

Pomfret, R. (1992). *International trade policy with imperfect competition* (Special paper in *International economics*, Vol. 17). Princeton, NJ: Princeton University, International Finance Section.

Posner, M. V. (1961). International trade and technical change. *Oxford Economic Papers, 13*, 323–341.

Prusa, T. J. (1992). Why are so many antidumping petitions withdrawn? *Journal of International Economics, 33*, 1–20.

Puga, D. (1999). The rise and fall of regional inequalities. *European Economic Review, 43*, 303–334.

Puga, D., & Venables, A. J. (1997). Preferential trading arrangements and industrial location. *Journal of International Economics, 43*, 347–368.

Qiu, L. D. (1995). Why can't countervailing duties deter export subsidization? *Journal of International Economics, 39*, 249–272.

Razawi, H. (1984). An economic model of OPEC coalition. *Southern Economic Journal, 51*, 419–428.

Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics, 62*, 53–82.

Redding, S. J. (2010). The empirics of new economic geography. *Journal of Regional Science, 50*, 297–311.

Redding, S. J. (2011). Theories of heterogeneous firms and trade. *Annual Review of Economics, 3*, 77–105.

Redding, S. J., & Sturm, D. (2008). The costs of remoteness: Evidence from German division and reunification. *American Economic Review, 98*, 1766–1797.

Ricardo, D. (1817). *On the principles of political economy and taxation* (1st ed.). London: J. Murray. (Reprinted as Vol. 1 of *The works and correspondence of david Ricardo*, by P. Sraffa, Ed., 1951, Cambridge University Press)

Ricci, L. A. (1999). Economic geography and comparative advantage: Agglomeration versus specialization. *European Economic Review, 43*, 357–377.

Riedel, J. (1976). Intermediate products and the theory of international trade: A generalization of the pure intermediate goods case. *American Economic Review, 66*, 441–447.

Robert-Nicoud, F. L. (2005). The structure of simple ‘New Economic Geography’ models (or, on identical twins). *Journal of Economic Geography, 5*, 201–234.
Robinson, J. (1954). The production function and the theory of capital. *Review of Economic Studies, 21*, 81–106. (Reprinted (in part) in *Collected economic papers*, Vol. 11, pp. 114–131, by J. Robinson, 1960, Oxford, UK: Blackwell)
Robson, P. (1998). *The economics of international integration* (4th ed.). London: Allen & Unwin.
Romalis, J. (2004). Factor proportions and the structure of commodity trade. *American Economic Review, 91*, 67–97.
Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives, 8*, 3–22.
Ruffin, R. J. (1984). International factor movements. In R. W. Jones & P. B.Kenen (Eds.) (chap. 5).
Rybczynski, T. M. (1955). Factor endowment and relative commodity prices. *Economica, 22*, 336–341. (Reprinted in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 72–77)
Salop, S. C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics, 10*, 141–156.
Salter, W. E. G. (1959). Internal and external balance: The role of price and expenditure effects. *Economic Record, 35*, 226–238.
Salvatore, D. (2010). *International economics: Trade and finance* (10th ed.). New York: Wiley.
Samuelson, P. A. (1947). *Foundations of economic analysis*. Cambridge, MA: Harvard University Press (enlarged edition, 1983).
Samuelson, P. A. (1948). International trade and the equalisation of factor prices. *Economic Journal, 58*, 163–184. (Reprinted in P. A. Samuelson, 1966b, chap. 67)
Samuelson, P. A. (1949). International factor price equalisation once again. *Economic Journal, 59*, 181–197. (Reprinted in P. A. Samuelson, 1966b, chap. 68; in R. E. Caves & H. G. Johnson, Eds., 1968, pp. 58–71; and in J. N. Bhagwati, Ed., 1981, pp. 3–16)
Samuelson, P. A. (1953). Prices of factors and goods in general equilibrium. *Review of Economic Studies, 21*, 1–20. (Reprinted (with postscript) in P. A. Samuelson, 1966b, chap. 70)
Samuelson, P. A. (1954). The transfer problem and transport costs, II: Analysis of effects of trade impediments. *Economic Journal, 63*, 264–289. (Reprinted in P. A. Samuelson, 1966b, pp. 1012–1037)
Samuelson, P. A. (1962). Parable and realism in capital theory: The surrogate production function. *Review of Economic Studies, 29*, 193–206. (Reprinted in P. A. Samuelson, 1966a, chap. 28)
Samuelson, P. A. (1965). Equalization by trade of the interest rate along with the real wage. In Various Authors (Eds.), *Trade, growth and the balance of payments: Essays in honour of G. Haberler.* (pp. 35–52). Chicago: Rand McNally. (Reprinted in P. A. Samuelson, 1966b, chap. 71)
Samuelson, P. A. (1966a). Edited by J. E. Stiglitz, *Collected scientific papers* (Vol. 1). Cambridge, MA: MIT Press.
Samuelson, P. A. (1966b). Edited by J. E. Stiglitz, *Collected scientific papers* (Vol. II). Cambridge, MA: MIT Press.
Samuelson, P. A. (1967). Summary on factor price equalization. *International Economic Review, 8*, 286–295. (Reprinted in P. A. Samuelson, 1972, chap. 161)
Samuelson, P. A. (1971). Ohlin was right. *Swedish Journal of Economics, 73*, 365–384. (Reprinted in P. A. Samuelson, 1977, pp. 667–686)
Samuelson, P. A. (1972). Edited by R. Merton, *Collected scientific papers* (Vol. III). Cambridge, MA: MIT Press.
Samuelson, P. A. (1977). Edited by H. Nagatani & K. Crowley, *Collected scientific papers* (Vol. IV). Cambridge, MA: MIT Press.
Sanyal, K. K., & Jones, R. W. (1982). The theory of trade in middle products. *American Economic Review, 72*, 16–31.
Sarkar, A. (1984). A dynamic assignment problem. *International Economic Review, 25*, 663–670.
Sarkar, A. (1985). A model of trade in intermediate goods. *Journal of International Economics, 19*, 85–98.

Savosnick, K. M. (1958). The box diagram and the production possibility curve. *Ekonomisk Tidskrift, 60*, 183–197.

Schefold, B. (1985). Cambridge price theory: Special model or general theory of value? *American Economic Review, 75*, 140–145 (Papers and Proceedings).

Schucknecht, L. (1992). *Trade protection in the European community*. Reading, UK: Harwood Academic Publishers.

Sgro, P. M. (1986). *The theory of duality and international trade*. New York: New York University Press.

Shaked, A., & Sutton, J. (1982). Relaxing price competition through product differentiation. *Review of Economic Studies, 49*, 3–13.

Shaked, A., & Sutton, J. (1983). Natural oligopolies. *Econometrica, 51*, 1469–1483.

Shaked, A., & Sutton, J. (1984). Natural oligopolies and international trade. In H. Kierzkowski (Ed.) (pp. 34–50).

Sheikh, M. A. (1974). Smuggling, production and welfare. *Journal of International Economics, 4*, 335–364. (Reprinted in J. N. Bhagwati, Ed., 1981, pp. 233–242)

Smith, A. (1984). Capital theory and trade theory. In R. W. Jones & P. B. Kenen (Eds.) (chap. 6).

Smith, M. A. M. (1977). Capital accumulation in the open two-sector economy. *Economic Journal, 87*, 273–282. (Reprinted in J. N. Bhagwati, Ed., 1981, pp. 329–341)

Snape, R. H. (1977). Effects of mineral development on the economy. *Australian Journal of Agricultural Economics, 21*, 147–156.

Södersten, B., & Reed, G. (1994). *International economics* (3rd ed.). London: Macmillan.

Solow, R. M. (1992). *Siena lectures on endogenous growth theory*. University of Siena, Department of Economics Monographs series No. 6. (A revised and enlarged edition has been published in Italian as R. M. Solow, 1994, *Lezioni sulla teoria della crescita endogena* (with appendices by Serena Sordi). Roma: La Nuova Italia Scientifica).

Spaventa, L. (1968). Realism without parables in capital theory. In *Récherches récentes sur la fonction de production* (pp. 15–45). Namur, Belgium: Centre d'Etudes et de Récherches Universitaires de Namur.

Spence, A. M. (1976). Product selection, fixed costs, and monopolistic competition. *Review of Economic Studies, 43*, 215–235.

Sraffa, P. (1960). *Production of commodities by means of commodities*. Cambridge, UK: Cambridge University Press.

Steedman, I. (Ed.). (1979). *Fundamental issues in trade theory*. London: Macmillan.

Steedman, I. (1979). *Trade amongst growing economies*. Cambridge, UK: Cambridge University Press.

Stern, R. M. (1973). Tariffs and other measures of trade control: A survey of recent developments. *Journal of Economic Literature, 11*, 857–888.

Stern, R. M. (1975). Testing trade theories. In P. B. Kenen (Ed.), *International trade and finance*. Cambridge, UK: Cambridge University Press.

Stern, R. M. (2003). Labor standards and trade agreements. *Revue d'Economie du Developpement, 17*, 125–150.

Stern, R. M., & Maskus, K. E. (1981). Determinants of the structure of U.S. foreign trade, 1958–1976. *Journal of International Economics, 11*, 207–224.

Stolper, W. F., & Samuelson, P. A. (1941). Protection and real wages. *Review of Economic Studies, 9*, 50–73. (Reprinted in J. N. Bhagwati, Ed., 1969, pp. 245–268; and in P. A. Samuelson, 1972, pp. 831–846)

Svensson, L. E. O. (1984). Factor trade and goods trade. *Journal of International Economics, 16*, 365–378.

Swann, D. (2000). *The economics of Europe: From common market to European union* (9th ed.). London: Penguin.

Tabuchi, T., & Thisse, J.-F. (2002). Taste heterogeneity, labor mobility and economic geography. *Journal of Development Economics, 69*, 155–177.

Takacs, W. E. (1978). The nonequivalence of tariffs, import quotas, and voluntary export restraints. *Journal of International Economics, 8*, 565–573.

Takayama, A. (1972). *International trade: An approach to the theory*. New York: Holt, Rinehart & Winston.

Tharakan, P. K. M. (Ed.). (1983). *Intra-industry trade: Empirical and methodological aspects*. Amsterdam: North-Holland.

Torrens, R. (1815). *An essay on the external corn trade* (1st ed.). London: J. Hatchard.

Travis, W. P. (1964). *The theory of trade and protection*. Cambridge, MA: Harvard University Press.

Trefler, D. (1993). International factor price differences: Leontief was right! *Journal of Political Economy, 101*, 961–987.

Trefler, D. (1995). The case of the missing trade and other mysteries. *American Economic Review, 85*, 1029–1046.

Trionfetti, F. (2001). Public procurement, economic integration, and income inequality. *Review of International Economics, 9*, 29–41.

Trionfetti, F. (2012). Public debt and economic geography. *International Regional Science Review*. Published online 9 November 2012 as doi:10.1177/0160017612462873.

Tsoulakis, L. (1997). *The new European economy revisited* (3rd ed.). Oxford, UK: Oxford University Press.

Unel, B. (2010). Firm heterogeneity, trade, and wage inequality. *Journal of Economic Dynamics & Control, 34*, 1369–1379.

United Nations. (1975). *Standard international trade classification* (Revision 2). New York: UN (Statistical Office, Department of Economic and Social Affairs).

Vanek, J. (1959). The natural resource content of foreign trade, 1870–1955, and the relative abundance of natural resources in the United States. *Review of Economics and Statistics, 41*, 146–153.

Vanek, J. (1962). *International trade: Theory and economic policy*. Homewood, IL: Irwin.

Vanek, J. (1963). Variable factor proportions and interindustry flows in the theory of international trade. *Quarterly Journal of Economics, 77*, 129–142.

Vanek, J. (1966). Towards a more general theory of growth with technological change. *Economic Journal, 76*, 841–854.

Vanek, J. (1968). The factor proportions theory: The n-factor case. *Kyklos 21*, 749–754.

van Wijnbergen, S. (1984). The 'Dutch Disease': A disease after all? *Economic Journal, 94*, 41–55.

Varian, H. R. (1992). *Microeconomic analysis* (3rd ed.). New York: Norton.

Various Authors (Eds.). (1965). *Trade, growth and the balance of payments: Essays in honour of G. Haberler*. Chicago: Rand McNally.

Various Authors. (1966). Paradoxes in capital theory: A symposium. *Quarterly Journal of Economics, 80*, 503–583.

Various Authors. (1973). The theory of effective protection in general equilibrium: A symposium. *Journal of International Economics, 3*(3), 205–281.

Various Authors. (1983). International factor mobility: A symposium. *Journal of International Economics, 14*(3/4), 209–412.

Venables, A. J. (1985). Trade and trade policy with imperfect competition: The case of identical products and free entry. *Journal of International Economics, 19*, 1–19.

Venables, A. J. (1987). Trade and trade policy with differentiated products: A Chamberlinian-Ricardian model. *Economic Journal, 97*, 700–717.

Venables, A. J. (1990). International capacity choice and national market games. *Journal of International Economics, 29*, 23–42.

Vernon, R. (1966). International investment and international trade in the product cycle. *Quarterly Journal of Economics, 80*, 190–207.

Vernon, R. (1979). The product cycle hypothesis in a new international environment. *Oxford Bulletin of Economics and Statistics, 41*, 255–267.

Vernon, R. (Ed.). (1970). The technology factor in international trade. New York: NBER.

Viner, J. (1937). *Studies in the theory of international trade*. New York: Harper & Brothers.

Viner, J. (1950). *The customs union issue*. New York: Carnegie Endowment for International Peace.

Vona, S. (1990). Intra-industry trade: A statistical artifact or a real phenomenon? *Banca Nazionale del Lavoro Quarterly Review, 175*, 383–412.

von Thünen, J. H. (1966), 1826. *Von Thünen's isolated state*. (English translation of 1826 German original, P. Hall, Ed., Oxford: Pergamon Press)

Vori, S. (1984). L'ipotesi della "varietà" nelle nuove teorie del commercio internazionale: il modello e la sua verifica empirica. *Quaderni Sardi di Economia, 14*, 163–190.

Vousden, N. (1990). *The economics of trade protection*. Cambridge, UK: Cambridge University Press.

Wan, H. Y., Jr. (1971). *Economic growth*. New York: Harcourt Brace Jovanovich.

Whitin, T. M. (1953). Classical theory, Graham's theory, and linear programming. *Quarterly Journal of Economics, 47*, 521–544.

Williams, J. R. (1970). The resource content in international trade. *Canadian Journal of Economics, 3*, 111–122.

Williamson, G. J. (2006). *Globalization and the poor periphery before 1950*. Cambridge, MA: MIT Press.

Wilson, C. (1980). On the general structure of Ricardian models with a continuum of goods: Applications to growth, tariffs and technical change. *Econometrica, 48*, 1675–1702.

Winters, L. A. (1989). Britain in Europe: A survey of quantitative trade studies. In A. Jacquemin & A. Sapir (Eds.), *The European internal market: Trade and competition*. Oxford, UK: Oxford University Press.

Winters, L. A. (1992). The welfare and policy implications of the international trade consequences of "1992". *American Economic Review, 82*, 104–108 (Papers and Proceedings).

Wong, K.-Y. (1995). *International trade in goods and factor mobility*. Cambridge, MA: MIT Press.

Wood, A. (1994). Give Heckscher and Ohlin a chance! *Weltwirtschaftliches Archiv, 130*, 20–49.

Woodland, A. D. (1982). *International trade and resource allocation*. Amsterdam: North-Holland.

WTO. (1995). *Trading into the future: WTO-the World trade organization*. Geneva, Switzerland: World Trade Organization.

Yabuuchi, S., & Tanaka, K. (1981). Non-traded inputs, interindustry flows, resource allocation and the ERP theory. *Journal of International Economics, 11*, 99–115.

Yeaple. S. R. (2005). A simple model of firm heterogeneity and international trade, and wages. *Journal of International Economics, 65*, 1–20.

Yeats, A. J. (1974). Effective tariff protection in the United States, the European economic community, and Japan. *The Quarterly Review of Economics and Business, 14*, 41–50.

Yeung, P., & Tsang, H. (1972). Generalized production function and factor-intensity crossovers: An empirical analysis. *Economic Record, 48*, 387–399.

Yntema, T. O. (1932). *A mathematical reformulation of the general theory of international trade*. Chicago: University of Chicago Press.

Yoo, J.-H. (1993). The political economy of protection structure in Korea. In A. O. Krueger & T. Ito (Eds.), *Trade and protectionism*. Chicago: University of Chicago Press.

Index

G. Gandolfo, *International Trade Theory and Policy*, Springer Texts in Business and Economics, DOI 10.1007/978-3-642-37314-5,