Second Edition

ADVANCED INTERNATIONAL TRADE

Theory and Evidence

ROBERT C. FEENSTRA

ADVANCED INTERNATIONAL TRADE

Second Edition

Advanced International Trade

THEORY AND EVIDENCE

Second Edition

Robert C. Feenstra

Princeton University Press
Princeton and Oxford

Copyright © 2016 by Princeton University Press
Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press, 6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

press.princeton.edu

Jacket image: *Wards Island Bridge*, Bascove © 2001, private collection

All Rights Reserved

ISBN 978-0-691-16164-8

Library of Congress Control Number 2015946714

British Library Cataloging-in-Publication Data is available

This book has been composed in Minion Pro and Myriad

Printed on acid-free paper. ∞

Printed in the United States of America

1 3 5 7 9 10 8 6 4 2

To Heather and Evan

CONTENTS

ACKNOWLEDGMENTS

A book like this would not be possible without the assistance of many people; indeed, without a career full of teachers and colleagues who have shaped the field and my own views of it. I am fortunate to have had many of these who were all generous with their time and insights. As an undergraduate at the University of British Columbia, I first learned international trade from the original edition of Richard Caves and Ronald Jones, *World Trade and Payments* (Little, Brown, 1973), and then from the "dual" point of view with Alan Woodland. His continuing influence will be clear in this book. I also took several graduate courses in duality theory from Erwin Diewert, and while it took longer for these ideas to affect my research, their impact has been quite profound.

In 1977 I began my graduate work at MIT, where I learned international trade from Jagdish Bhagwati and T. N. Srinivasan, both leaders and collaborators in the field. These courses greatly expanded my knowledge of all aspects of international trade and stimulated me to become a trade economist. The book I have written probably does not do justice to the great breadth of research topics I learned there. These topics are well presented in the textbook by Jagdish Bhagwati, Arvind Panagariya, and T. N. Srinivasan, *Lectures in International Trade* (MIT Press, 1998), to which the reader is referred.

In 1981, I joined Jagdish Bhagwati as a colleague at Columbia University, and for the next five years enjoyed the company of Ronald Findlay, Guillermo Calvo, Robert Mundell, Maury Obstfeld, Stanislaw Wellisz, and other regular attendees at the Wednesday afternoon international seminar. Richard Brecher was also a visitor at Columbia during this period. I am much indebted to these colleagues for the stimulating seminars and conversations. My early work in trade did not stray too far from the familiar two-sector model, but in 1982 I was invited by Robert Baldwin to participate in a conference of the National Bureau of Economic Research (NBER) focusing on the empirical assessment of U.S. trade policies. For this conference I prepared a paper on the voluntary export restraint with Japan in automobiles, and thus my empirically oriented research in trade began.

Following this, I participated in more conferences of the NBER, and since 1992 have directed the International Trade and Investment group there. These conferences have been enormously influential to my research. During the 1980s, there was a large amount of research on trade and trade policy under imperfect competition. Some of this work was done for NBER conferences, and I was able to witness all of it. In the 1990s the focus shifted very substantially to "endogenous" growth models, and this research was again supported by a working group at the NBER and many other researchers. I was fortunate to spend a sabbatical leave in 1989 at the Institute for Advanced Study of the Hebrew University of Jerusalem, along with Wilfred Ethier, Gene Grossman, Paul Krugman, and James Markusen, and under the organization of Elhanan Helpman and Assaf Razin, during which time these growth models were being developed. Furthermore, throughout the 1980s and 1990s there has been an increased awareness of empirical issues in international trade, with fine work by many new colleagues. And during the past decade a great deal of attention has been given to the role of heterogeneous firms in shaping trade patterns and in the gains from trade. So in the span of three decades since beginning my career, I have seen four great waves of research in international trade, and I attempt to summarize all of them in this second edition of the book.

My own research, much of it included here, would not have possible without generous funding from the National Science Foundation, the Ford Foundation, and the Sloan Foundation, as well as the assistance of many graduate students over the years. Some of their research appears in this book, and many others have labored faithfully to construct datasets that have been widely distributed. Let me thank in particular Dorsati Madani, Maria Yang, and Shunli Yao for their past work on datasets, along with David Yue, Roger Butters, Seungjoon Lee, Songhua Lin, and Alyson Ma for their help with the first edition of this textbook. My colleagues and students at UC Davis have been very helpful in discussing this second edition, and I especially thank Vladimir Tyazhelnikov, who provided invaluable research assistance. Finally, I would like to thank a number of other colleagues who contributed to this book: Lee Branstetter provided a number of the empirical exercises, with the STATA programs written by Kaoru Nabeshima; Bin Xu read nearly all the chapters at an early stage and provided detailed comments; and Bruce Blonigen, Donald Davis, Earl Grinols, Doug Irwin, Jiandong Ju, James Levinsohn, Nina Pavcnik, Larry Qiu, James Rauch, and Daniel Trefler provided comments or datasets for specific chapters. I am grateful to these individuals and many others whose input has improved this book.

FOREWORD TO THE SECOND EDITION

This book is intended for a graduate course in international trade. I assume that all readers have completed graduate courses in microeconomics and econometrics. My goal is to bring the reader from that common point up to the most recent research in international trade, in both theory and empirical work. This book is not intended to be difficult, and the mathematics used should be accessible to any graduate student. The material covered will give the reader the skills needed to understand the latest articles and working papers in the field.

The first edition of this textbook was published in 2004, and I am gratified that it has been used in many graduate courses across the United States and around the world. The goals for this second edition are the same as for the first: to provide an accessible treatment of the theory and empirical applications in international trade, thereby bringing the reader up to the frontier of research. But this second edition has another goal, too, and that is to incorporate the most important new material that has appeared in the ten years since the first edition. Foremost among that research are the monopolistic competition model of international trade with heterogeneous firms, due to Melitz (2003), and the Ricardian model of international trade with heterogeneous productivities across countries, due to Eaton and Kortum (2002). While these articles were published prior to the 2004 first edition of this book, they did not appear there other than as references for further reading. That shortcoming is corrected in this second edition.

Because the work of Eaton and Kortum (2002) builds on the Ricardian model with a continuum of goods, we present that model at the end of chapter 3 (drawing upon the presentation in Matsuyama 2008). Then the new chapter 6 has been added, dealing with monopolistic competition and the gravity equation with heterogeneous firms. The presentation in that chapter owes a great deal to articles by Melitz and Redding (2014a) and by Head and Mayer (2014) in the new *Handbook of International Economics, Volume 4.* Melitz and Redding introduce some simplifications to the original Melitz (2003) model and a method of solving it that I rely on in chapter 6, while Head and Mayer show how the gravity equation can be obtained quite generally from a wide range of models, including the Eaton-Kortum model, which is also introduced in chapter 6.

Besides the new chapter 6, other chapters have been rewritten extensively. Chapter 4, dealing with trade in intermediate inputs and wages, now draws on the presentation in my Ohlin Lecture (Feenstra 2010a) and incorporates the work by Grossman and Rossi-Hansberg (2008) on trade in tasks. Chapter 5, focusing on monopolistic competition under homogeneous firms, now includes material on measuring the gains from product variety drawing on Feenstra (1994, 2010b). Chapter 8, which discusses import tariffs and dumping, includes new material on the pass-through of tariffs (or exchange rates) using functional forms that lead to exact expressions and also recent empirical work in this area. Chapter 9, an inquiry into import quotas and export subsidies, naturally lends itself to a fuller treatment of the quality of traded goods, drawing on Feenstra and Romalis (2014). Chapter

10, on the political economy of trade policy, now includes the work of Ossa (2011). Chapter 11, dealing with endogenous growth, includes a brief exposition of growth with heterogeneous firms due to Sampson (2016). And chapter 12, an exploration of multinational corporations and the organization of the firm, has also been rewritten extensively, drawing upon Antràs (2003) and Antràs and Helpman (2004). Much of that material is covered in more depth in the volume by Antràs (2015). Throughout this second edition, I have benefited greatly from researchers providing the surveys and articles that I have incorporated here, for which I offer my thanks.

As before, an instructor's manual that accompanies this book provides solutions to the problems at the end of the chapters.[1] In addition, I have included empirical exercises that replicate the results in some chapters. Completing all of these could be the topic for a second course, but even in a first course there will be a payoff to trying some of the exercises. The data and programs for these can be found on my home page at www.robertfeenstra.info.

The notational conventions from the first edition have been retained in the second edition. I consistently use *subscripts* to refer to goods or factors, whereas *superscripts* refer to consumers or countries. In general, then, *subscripts refer to commodities and superscripts refer to agents*. The index used (h, i, j, k, ℓ, m, or n) will depend on the context. The symbol "c" is used for both costs and consumption, though in some chapters I instead use "$d(p)$" for consumption to avoid confusion. The output of firms is consistently denoted by "y" and exports are denoted by "x", though in some cases "x" denotes inputs. Uppercase letters are used in some cases to denote vectors or matrixes, and in other cases to denote the number of goods (N), factors (M), households (H), or countries (C), and sporadically elsewhere. The symbols α and β are used generically for intercept and slope coefficients, including fixed and marginal labor costs, except that with heterogeneous firms we use the symbol φ for the output produced by one unit of labor and $1/\varphi$ for marginal labor costs.

The contents of several chapters included here have been previously published. Chapter 4 draws on material from my book *Offshoring in the Global Economy: Theory and Evidence* (MIT Press, 2010a) and chapter 5 draws upon the *Scottish Journal of Political Economy* (Feenstra 2002). Some material from chapters 8–10 has appeared in articles published in the *Journal of International Economics* and the *Quarterly Journal of Economics,* and material from chapter 11 has appeared in the *Journal of Development Economics* and in the *American Economic Review.*

[1] Faculty wishing to obtain the instructors manual should contact Princeton University Press.

ADVANCED INTERNATIONAL TRADE

1

Preliminaries: Two-Sector Models

We begin our study of international trade with the classic Ricardian model, which has two goods and one factor (labor). The Ricardian model introduces us to the idea that technological differences across countries matter. In comparison, the Heckscher-Ohlin model dispenses with the notion of technological differences and instead shows how *factor endowments* form the basis for trade. While this may be fine in theory, the model performs very poorly in practice: as we show in the next chapter, the Heckscher-Ohlin model is hopelessly inadequate as an explanation for historical or modern trade patterns unless we allow for technological differences across countries. For this reason, the Ricardian model is as relevant today as it has always been. Our treatment of it in this chapter is a simple review of undergraduate material, but we will present a more sophisticated version of the Ricardian model (with a continuum of goods) in chapter 3.

After reviewing the Ricardian model, we turn to the two-good, two-factor model that occupies most of this chapter and forms the basis of the Heckscher-Ohlin model. We shall suppose that the two goods are traded on international markets, but do not allow for any movements of factors across borders. This reflects the fact that the movement of labor and capital across countries is often subject to controls at the border and is generally much less free than the movement of goods. Our goal in the next chapter will be to determine the pattern of international trade between countries. In this chapter, we simplify things by focusing primarily on *one* country, treating world prices as given, and examine the properties of this two-by-two model. The student who understands all the properties of this model has already come a long way in his or her study of international trade.

RICARDIAN MODEL

Indexing goods by the subscript i, let a_i denote the labor needed per unit of production of each good at home, while a_i^* is the labor need per unit of production in the foreign country, $i=1,2$. The total labor force at home is L and abroad is L^*. Labor is perfectly mobile between the industries in each country, but immobile across countries. This means that both goods are produced in the home country only if the wages earned in the two industries are the same. Since the marginal product of labor in each industry is $1/a_i$, and workers are paid the value of their marginal

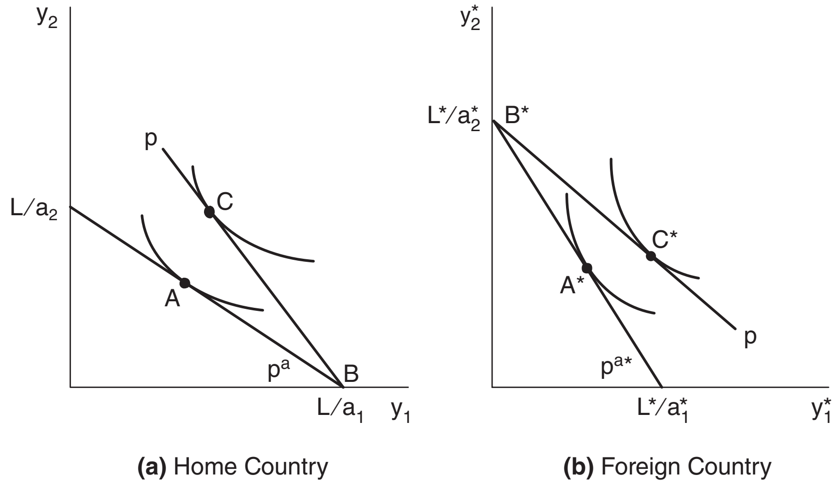

Figure 1.1

products, wages are equalized across industries if and only if $p_1/a_1 = p_2/a_2$, where p_i is the price in each industry. Letting $p = p_1/p_2$ denote the *relative* price of good 1 (using good 2 as the numeraire), this condition is $p = a_1/a_2$.

These results are illustrated in figure 1.1(a) and (b), where we graph the production possibility frontiers (PPFs) for the home and foreign countries. With all labor devoted to good i at home, it can produce L/a_i units, $i = 1, 2$, so this establishes the intercepts of the PPF, and similarly for the foreign country. The slope of the PPF in each country (ignoring the negative sign) is then a_1/a_2 and a_1^*/a_2^*. Under autarky (i.e., no international trade), the equilibrium relative prices p^a and p^{a*} must equal these slopes in order to have both goods produced in both countries, as argued above. Thus, the autarky equilibrium at home and abroad might occur at points A and A^*. Suppose that the home country has a *comparative advantage* in producing good 1, meaning that $a_1/a_2 < a_1^*/a_2^*$. This implies that the home autarky relative price of good 1 is *lower* than that abroad.

Now letting the two countries engage in international trade, what is the equilibrium price p at which world demand equals world supply? To answer this, it is helpful to graph the world relative supply and demand curves, as illustrated in figure 1.2. For the relative price satisfying $p < p^a = a_1/a_2$ and $p < p^{a*} = a_1^*/a_2^*$ both countries are fully specialized in good 2 (since wages earned in that sector are higher), so the world relative supply of good 1 is zero. For $p^a < p < p^{a*}$, the home country is fully specialized in good 1 whereas the foreign country is still specialized in good 2, so that the world relative supply is $(L/a_1)/(L^*/a_2^*)$, as labeled in figure 1.2. Finally, for $p > p^a$ and $p > p^{a*}$, both countries are specialized in good 1. So we see that the world relative supply curve has a "stair-step" shape, which reflects the linearity of the PPFs.

To obtain world relative demand, let us make the simplifying assumption that tastes are identical and homothetic across the countries. Then demand will be independent of the distribution of income *across* the countries. Demand being homothetic means that *relative* demand d_1/d_2 in either country is a downward-sloping function of the relative price p, as illustrated in figure 1.2. In the case we have shown, relative demand intersects relative supply at the world price p that lies *between* p^a and p^{a*}, but this does

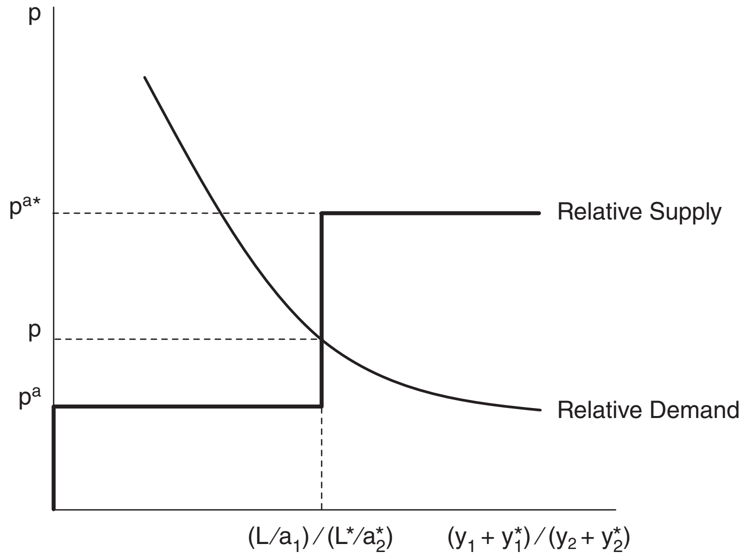

Figure 1.2

not need to occur: instead, we can have relative demand intersect one of the flat segments of relative supply, so that the equilibrium price with trade *equals* the autarky price in one country.[1]

Focusing on the case where $p^a < p < p^{a*}$, we can go back to the PPF of each country and graph the production and consumption points with free trade. Since $p > p^a$, the home country is fully specialized in good 1 at point *B*, as illustrated in figure 1.1(a), and then trades at the relative price p to obtain consumption at point *C*. Conversely, since $p < p^{a*}$, the foreign country is fully specialized in the production of good 2 at point B^* in figure 1.1(b), and then trades at the relative price p to obtain consumption at point C^*. Clearly, *both* countries are better off under free trade than they were in autarky: trade has allowed them to obtain a consumption point that is above the PPF.

Notice that the home country exports good 1, which is in keeping with its comparative advantage in the production of that good, $a_1/a_2 < a_1^*/a_2^*$. Thus, *trade patterns are determined by comparative advantage,* which is a deep insight from the Ricardian model. This occurs even if one country has an *absolute disadvantage* in both goods, such as $a_1 > a_1^*$ and $a_2 > a_2^*$, so that more labor is needed per unit of production of *either* good at home than abroad. The reason that it is still possible for the home country to export is that its *wages* will adjust to reflect its productivities: under free trade, its wages are lower than those abroad.[2] Thus, while trade patterns in the Ricardian model are determined by *comparative advantage*, the level of wages across countries is determined by *absolute advantage*.

[1] This occurs if one country is very large. Use figures 1.1 and 1.2 to show that if the home country is very large, then $p = p^a$ and the home country does not gain from trade.

[2] The home country exports good 1, so wages earned with free trade are $w = p/a_1$. Conversely, the foreign country exports good 2 (the numeraire), and so wages earned there are $w^* = 1/a_2^* > p/a_1^*$, where the inequality follow since $p < a_1^*/a_2^*$ in the equilibrium being considered. Then using $a_1 > a_1^*$, we obtain $w = p/a_1 < p/a_1^* < w^*$.

TWO-GOOD, TWO-FACTOR MODEL

While the Ricardian model focuses on technology, the Heckscher-Ohlin model, which we study in the next chapter, focuses on factors of production. So we now assume that there are two factor inputs—labor and capital. Restricting our attention to a single country, we will suppose that it produces two goods with the production functions $y_i = f_i(L_i, K_i)$, $i = 1, 2$, where y_i is the output produced using labor L_i and capital K_i. These production functions are assumed to be increasing, concave, and homogeneous of degree one in the inputs (L_i, K_i).[3] The last assumption means that there are *constant returns to scale* in the production of each good. This will be a maintained assumption for the next several chapters, but we should be point out that it is rather restrictive. It has long been thought that *increasing returns to scale* might be an important reason to have trade between countries: if a firm with increasing returns is able to sell in a foreign market, this expansion of output will bring a reduction in its average costs of production, which is an indication of greater efficiency. Indeed, this was a principal reason why Canada entered into a free-trade agreement with the United States in 1989: to give its firms free access to the large American market. We will return to these interesting issues in chapter 5, but for now, ignore increasing returns to scale.

We will assume that labor and capital are *fully mobile* between the two industries, so we are taking a "long run" point of view. Of course, the amount of factors employed in each industry is constrained by the endowments found in the economy. These resource constraints are stated as

$$\begin{aligned} L_1 + L_2 &\leq L, \\ K_1 + K_2 &\leq K, \end{aligned} \tag{1.1}$$

where the endowments L and K are fixed. Maximizing the amount of good 2, $y_2 = f_2(L_2, K_2)$, subject to a given amount of good 1, $y_1 = f_1(L_1, K_1)$, and the resource constraints in (1.1) give us $y_2 = h(y_1, L, K)$. The graph of y_2 as a function of y_1 is shown as the PPF in figure 1.3. As drawn, y_2 is a *concave* function of y_1, $\partial^2 h(y_1, L, K)/\partial y_1^2 < 0$. This familiar result follows from the fact that the production functions $f_i(L_i, K_i)$ are assumed to be concave. Another way to express this is to consider all points $S = (y_1, y_2)$ that are feasible to produce given the resource constraints in (1.1). This production possibilities set S is *convex*, meaning that if $y^a = (y_1^a, y_2^a)$and $y^b = (y_1^b, y_2^b)$ are both elements of S, then any point between them $\lambda y^a + (1-\lambda) y^b$ is also in S, for $0 \leq \lambda \leq 1$.[4]

The production possibilities frontier summarizes the technology of the economy, but in order to determine where the economy produces on the PPF we need to add some assumptions about the market structure. We will assume perfect competition in the product markets and factor markets. Furthermore, we will suppose that product prices are given *exogenously*: we can think of these prices as established on world markets, and outside the control of the "small" country being considered.

[3] Students not familiar with these terms are referred to problems 1.1 and 1.2.

[4] See problems 1.1 and 1.3 to prove the convexity of the production possibilities set, and to establish its slope.

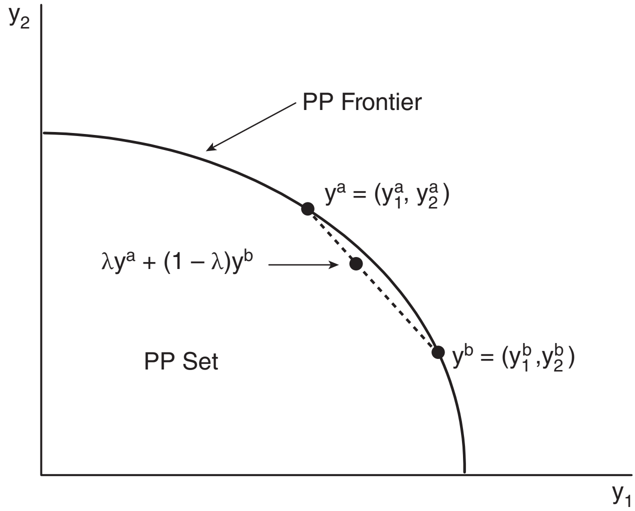

Figure 1.3

GDP FUNCTION

With the assumption of perfect competition, the amounts produced in each industry will maximize gross domestic product (GDP) for the economy: this is Adam Smith's "invisible hand" in action. That is, the industry outputs of the competitive economy will be chosen to maximize GDP:

$$G(p_1,p_2,L,K) = \max_{y_1,y_2} p_1 y_1 + p_2 y_2 \quad \text{s.t.} \quad y_2 = h(y_1,L,K). \tag{1.2}$$

To solve this problem, we can substitute the constraint into the objective function and write it as choosing y_1 to maximize $p_1 y_1 + p_2 h(y_1,L,K)$. The first-order condition for this problem is $p_1 + p_2(\partial h/\partial y_1) = 0$, or,

$$p = \frac{p_1}{p_2} = -\frac{\partial h}{\partial y_1} = -\frac{\partial y_2}{\partial y_1}. \tag{1.3}$$

Thus, the economy will produce where the relative price of good 1, $p = p_1/p_2$, is equal to the slope of the production possibilities frontier.[5] This is illustrated by the point A in figure 1.4, where the line tangent through point A has the slope of (negative) p. An increase in this price will *raise* the slope of this line, leading to a new tangency at point B. As illustrated, then, the economy will produce more of good 1 and less of good 2.

The GDP function introduced in (1.2) has many convenient properties, and we will make use of it throughout this book. To show just one property, suppose that we differentiate the GDP function with respect to the price of good i, obtaining

[5]Notice that the slope of the price line tangent to the PPF (in absolute value) equals the relative price of the good on the *horizontal* axis, or good 1 in figure 1.4.

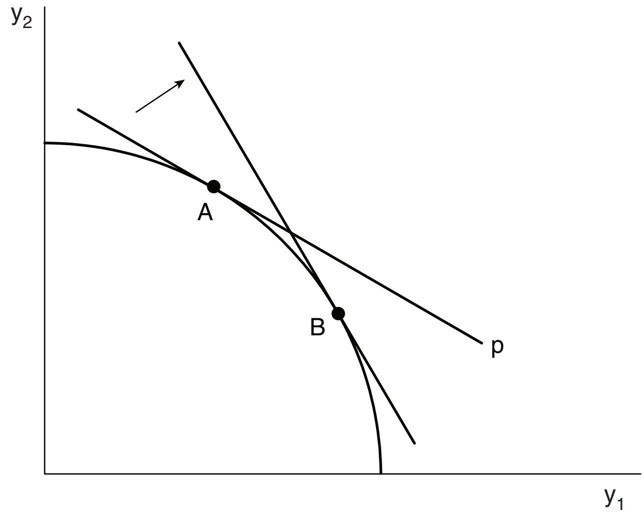

Figure 1.4

$$\frac{\partial G}{\partial p_i} = y_i + \left(p_1 \frac{\partial y_1}{\partial p_i} + p_2 \frac{\partial y_2}{\partial p_i} \right). \tag{1.4}$$

It turns out that the terms in parentheses on the right of (1.4) sum to *zero,* so that $\partial G/\partial p_i = y_i$. In other words, the derivative of the GDP function with respect to *prices* equals the *outputs* of the economy.

The fact that the terms in parentheses sum to zero is an application of the "envelope theorem," which states that when we differentiate a function that has been maximized (such as GDP) with respect to an exogenous variable (such as p_i), then we can *ignore* the changes in the endogenous variables (y_1 and y_2) in this derivative. To prove that these terms sum to zero, totally differentiate $y_2 = h(y_1, L, K)$ with respect to y_1 and y_2 and use (1.3) to obtain $p_1 dy_1 = -p_2 dy_2$, or $p_1 dy_1 + p_2 dy_2 = 0$. This equality must hold for any small movement in y_1 and y_2 around the PPF, and in particular, for the small movement in outputs *induced* by the change in p_i. In other words, $p_1(\partial y_1/\partial p_i) + p_2(\partial y_2/\partial p_i) = 0$, so the terms in parentheses on the right of (1.4) vanish and it follows that $\partial G/\partial p_i = y_i$.[6]

EQUILIBRIUM CONDITIONS

We now want to state succinctly the equilibrium conditions to determine factor prices and outputs. It will be convenient to work with the *unit-cost functions* that are dual to the production functions $f_i(L_i, K_i)$. These are defined by

$$c_i(w, r) = \min_{L_i, K_i \geq 0} \{ wL_i + rK_i \,|\, f_i(L_i, K_i) \geq 1 \}. \tag{1.5}$$

[6]Other convenient properties of the GDP function are explored in problem 1.4.

In words, $c_i(w,r)$ is the minimum cost to produce one unit of output. Because of our assumption of constant returns to scale, these unit-costs are equal to both marginal costs and average costs. It is easily demonstrated that the unit-cost functions $c_i(w,r)$ are non-decreasing and concave in (w,r). We will write the *solution* to the minimization in (1.5) as $c_i(w,r) = wa_{iL} + ra_{iK}$, where a_{iL} is optimal choice for L_i, and a_{iK} is optimal choice for K_i. It should be stressed that these optimal choices for labor and capital *depend* on the factor prices, so that they should be written in full as $a_{iL}(w, r)$ and $a_{iK}(w, r)$. However, we will usually not make these arguments explicit.

Differentiating the unit-cost function with respect to the wage, we obtain

$$\frac{\partial c_i}{\partial w} = a_{iL} + \left(w\frac{\partial a_{iL}}{\partial w} + r\frac{\partial a_{iK}}{\partial w} \right). \tag{1.6}$$

As we found with differentiating the GDP function, it turns out that the terms in parentheses on the right of (1.6) sum to zero, which is again an application of the "envelope theorem." It follows that the derivative of the unit-costs with respect to the wage equals the labor needed for one unit of production, $\partial c_i / \partial w = a_{iL}$. Similarly, $\partial c_i / \partial r = a_{iK}$.

To prove this result, notice that the constraint in the cost-minimization problem can be written as the isoquant $f_i(a_{iL}, a_{iK}) = 1$. Totally differentiate this to obtain $f_{iL}\, da_{iL} + f_{iK}\, da_{iK} = 0$, where $f_{iL} \equiv \partial f_i / \partial L_i$ and $f_{iK} \equiv \partial f_i / \partial K_i$. This equality must hold for any small movement of labor da_{iL} and capital da_{iK} around the isoquant, and in particular, for the change in labor and capital *induced* by a change in wages. Therefore, $f_{iL}(\partial a_{iL} / \partial w) + f_{iK}(\partial a_{iK} / \partial w) = 0$. Now multiply this through by the product price p_i, noting that $p_i f_{iL} = w$ and $p_i f_{iK} = r$ from the profit-maximization conditions for a competitive firm. Then we see that the terms in parentheses on the right of (1.6) sum to zero.

The first set of equilibrium conditions for the two-by-two economy is that *profits equal zero.* This follows from free entry under perfect competition. The zero-profit conditions are stated as

$$\begin{aligned} p_1 &= c_1(w,r), \\ p_2 &= c_2(w,r). \end{aligned} \tag{1.7}$$

The second set of equilibrium conditions is full employment of both resources. These are the same as the resource constraints (1.1), except that now we express them as equalities. In addition, we will rewrite the labor and capital used in each industry in terms of the derivatives of the unit-cost function. Since $\partial c_i / \partial w = a_{iL}$ is the labor used for *one unit* of production, it follows that the total labor used in $L_i = y_i a_{iL}$, and similarly the total capital used is $K_i = y_i a_{iK}$. Substituting these into (1.1), the full-employment conditions for the economy are written as

$$\begin{aligned} \underbrace{a_{1L}y_1}_{L_1} + \underbrace{a_{2L}y_2}_{L_2} &= L, \\ \underbrace{a_{1K}y_1}_{K_1} + \underbrace{a_{2K}y_2}_{K_2} &= K. \end{aligned} \tag{1.8}$$

Notice that (1.7) and (1.8) together are *four* equations in *four* unknowns, namely, (w,r) and (y_1, y_2). The parameters of these equations, p_1, p_2, L, and K, are given exogenously. Because the unit-cost functions are nonlinear, however, it is not enough to just count equations and unknowns: we need to study these equations in detail to understand whether the solutions are unique and strictly positive, or not. Our task for the rest of this chapter will be to understand the properties of these equations and their solutions.

To guide us in this investigation, there are three key questions that we can ask: (1) what is the solution for factor prices? (2) if prices change, how do factor prices change? (3) if endowments change, how do outputs change? Each of these questions is taken up in the sections that follow. The methods we shall use follow the "dual" approach of Woodland (1977, 1982), Mussa (1979), and Dixit and Norman (1980).

DETERMINATION OF FACTOR PRICES

Notice that our four-equation system above can be decomposed into the zero-profit conditions as *two* equations in *two* unknowns—the wage and rental—and then the full-empoyment conditions, which involve both the factor prices (which affect a_{iL} and a_{iK}) and the outputs. It would be especially convenient if we could uniquely solve for the factor prices from the zero-profit conditions, and then just substitute these into the full-employment conditions. This will be possible when the hypotheses of the following lemma, are satisfied.

LEMMA (FACTOR PRICE INSENSITIVITY)

So long as both goods are produced, and factor intensity reversals (FIRs) do not occur, then each price vector (p_1, p_2) corresponds to unique factor prices (w, r).

This is a remarkable result, because it says that the factor endowments (L, K) do not matter for the determination of (w, r). We can contrast this result with a one-sector economy, with production of $y = f(L, K)$, wages of $w = pf_L$, and diminishing marginal product $f_{LL} < 0$. In this case, any increase in the labor endowments would certainly reduce wages, so that countries with higher labor/capital endowments (L/K) would have lower wages. This is the result we normally expect. In contrast, the above lemma says that in a two-by-two economy, with a fixed product price p, it is possible for the labor force or capital stock to grow *without* affecting their factor prices! Thus, Leamer (1995) refers to this result as "factor price insensitivity." Our goal in this section is to prove the result and also develop the intuition for why it holds.

Two conditions must hold to obtain this result: first, that both goods are produced; and second, that factor intensity reversals (FIRs) do not occur. To understand FIRs, consider figures 1.5 and 1.6. In the first case, presented in figure 1.5, we have graphed the two zero-profit conditions, and the unit-cost lines intersect only *once,* at point A. This illustates the lemma: given (p_1, p_2), there is a *unique* solution for (w, r). But another case is illustrated in figure 1.6, where the unit-cost lines interesect *twice,* at points A and B. Then there are two possible solutions for (w, r), and the result stated in the lemma no longer holds.

The case where the unit-cost lines intersect more than once corresponds to "factor intensity reversals." To see where this name comes from, let us compute the labor and capital requirements in the two industries. We have already shown that a_{iL} and a_{iK} are the derivatives of the unit-cost function with respect to factor prices, so it follows that the vectors (a_{iL}, a_{iK}) are the *gradient vectors* to the iso-cost curves for the two industries

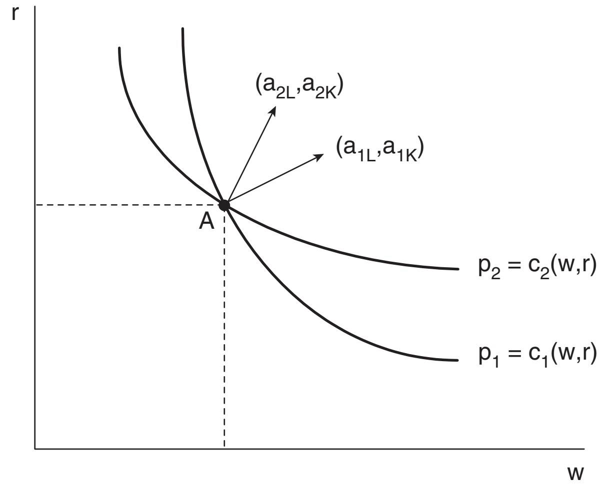

Figure 1.5

in figure 1.5. Recall from calculus that gradient vectors point in the direction of the maximum increase of the function in question. This means that they are *orthogonal* to their respective iso-cost curves, as shown by (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}) at point *A*. Each of these vectors has slope (a_{iK}/a_{iL}), or the capital-labor ratio. It is clear from figure 1.5 that (a_{1L}, a_{1K}) has a smaller slope than (a_{2L}, a_{2K}), which means that *industry 2 is capital intensive*, or equivalently, *industry 1 is labor intensive.*[7]

In figure 1.6, however, the situation is more complicated. Now there are two sets of gradient vectors, which we label by (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}) at point *A* and by (b_{1L}, b_{1K}) and (b_{2L}, b_{2K}) at point *B*. A close inspection of the figure will reveal that industry 1 is *labor intensive* $(a_{1K}/a_{1L} < a_{2K}/a_{2L})$ at point *A*, but is *capital intensive* $(b_{1K}/b_{1L} > b_{2K}/b_{2L})$ at point *B*. This illustrates a *factor intensity reversal*, whereby the comparison of factor intensities changes at different factor prices.

While FIRs might seem like a theoretical curiosum, they are actually quite realistic. Consider the footwear industry, for example. While much of the footwear in the world is produced in developing nations, the United States retains a small number of plants. For sneakers, New Balance has a plant in Norridgewock, Maine, where employers earn about \$14 per hour.[8] Some operate computerized equipment with up to twenty sewing machine heads running at once, while others operate automated stitchers guided by cameras, which allow one person to do the work of six. This is a far cry from the plants in Asia that produce shoes for Nike, Reebok, and other U.S. producers, using century-old technology and paying less than \$1 per hour. The technology used to make sneakers in Asia is like that of industry 1 at point *A* in figure 1.5, using labor-intensive

[7] Alternatively, we can totally differentiate the zero-profit conditions, holding prices fixed, to obtain $0 = a_{iL}dw + a_{iK}dr$. It follows that the slope of the iso-cost curve equals $dr/dw = -a_{iL}/a_{iK} = -L_i/K_i$. Thus, the slope of each iso-cost curve equals the relative demand for the factor on the horizontal axis, whereas the slope of the gradient vector (which is orthogonal to the iso-cost curve) equals the relative demand for the factor on the vertical axis.

[8] The material that follows is drawn from Aaron Bernstein, "Low-Skilled Jobs: Do They Have to Move?" *Business Week,* February 26, 2001, pp. 94–95.

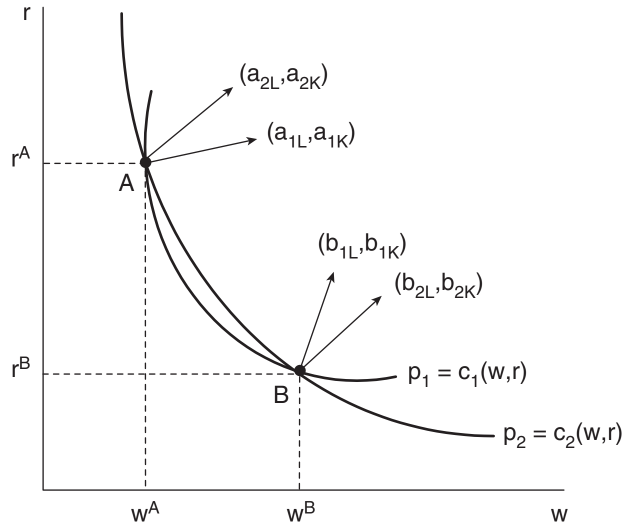

Figure 1.6

technology and paying low wages w^A, while industry 1 in the United States is at point *B*, paying higher wages w^B and using a capital-intensive technology.

As suggested by this discussion, when there are two possible solutions for the factor prices such as points *A* and *B* in figure 1.6, then some countries can be at one equilibrium and others countries at the other. How do we know which country is where? This is a question that we will answer at the end of the chapter, where we will argue that a *labor-abundant* country will likely be at equilibrium *A* of figure 1.6, with a low wage and high rental on capital, whereas a *capital-abundant* country will be at equilibrium *B*, with a high wage and low rental. Generally, to determine the factor prices in each country we will need to examine its full-employment conditions in addition to the zero-profit conditions.

Let us conclude this section by returning to the simple case of no FIR, in which the lemma stated above applies. What are the implications of this result for the determination of factor prices under free trade? To answer this question, let us sketch out some of the assumptions of the Heckscher-Ohlin model, which we will study in more detail in the next chapter. We assume that there are two countries, with identical technologies but different factor endowments. We continue to assume that labor and capital are the two factors of production, so that under free trade the equilibrium conditions (1.7) and (1.8) apply in *each* country with the *same* product prices (p_1, p_2). We can draw figure 1.5 for each country, and in the absence of FIR, this *uniquely* determines the factor prices in each countries. In other words, the wage and rental determined by figure 1.5 are *identical* across the two countries. We have therefore proved the factor price equalization (FPE) theorem, which is stated as follows.

FACTOR PRICE EQUALIZATION THEOREM (SAMUELSON 1949)

Suppose that two countries are engaged in free trade, having identical technologies but different factor endowments. If both countries produce both goods and FIRs do not occur, then the factor prices (w, r) are equalized across the countries.

The FPE theorem is a remarkable result because it says that *trade in goods* has the ability to equalize factor prices: in this sense, trade in goods is a "perfect substitute" for trade in factors. We can again contrast this result with that obtained from a one-sector economy in both countries. In that case, equalization of the product price through trade would certainly not equalize factor prices: the labor-abundant country would be paying a lower wage. Why does this outcome *not occur* when there are two sectors? The answer is that the labor-abundant country can *produce more of, and export,* the labor-intensive good. In that way it can fully employ its labor while still paying the same wages as a capital-abundant country. In the two-by-two model, the opportunity to disproportionately produce more of one good than the other, while exporting the amounts not consumed at home, is what allows factor price equalization to occur. This intuition will become even clearer as we continue to study the Heckscher-Ohlin model in the next chapter.

CHANGE IN PRODUCT PRICES

Let us move on now to the second of our key questions of the two-by-two model: if the product prices change, how will the factor prices change? To answer this, we perform comparative statics on the zero-profit conditions (1.7). Totally differentiating these conditions, we obtain

$$dp_i = a_{iL}dw + a_{iK}dr \Rightarrow \frac{dp_i}{p_i} = \frac{wa_{iL}}{c_i}\frac{dw}{w} + \frac{ra_{iK}}{c_i}\frac{dr}{r}, i = 1,2. \tag{1.9}$$

The second equation is obtained by multiplying and dividing like terms, and noting that $p_i = c_i(w,r)$. The advantage of this approach is that it allows us to express the variables in terms of *percentage changes,* such as $d\ln w = dw/w$, as well as *cost-shares.* Specifically, let $\theta_{iL} = wa_{iL}/c_i$ denote the cost-share of labor in industry *i*, while $\theta_{iK} = ra_{iK}/c_i$ denotes the cost-share of capital. The fact that costs equal $c_i = wa_{iL} + ra_{iK}$ ensures that the shares sum to unity, $\theta_{iL} + \theta_{iK} = 1$. In addition, denote the percentage changes by $dw/w = \hat{w}$ and $dr/r = \hat{r}$. Then (1.9) can be re-written as

$$\hat{p}_i = \theta_{iL}\hat{w} + \theta_{iK}\hat{r}, i = 1,2. \tag{1.9'}$$

Expressing the equations using these cost-shares and percentage changes follows Jones (1965) and is referred to as the "Jones algebra." This system of equations can be written in matrix form and solved as

$$\begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix} = \begin{pmatrix} \theta_{1L} & \theta_{1K} \\ \theta_{2L} & \theta_{2K} \end{pmatrix} \begin{pmatrix} \hat{w} \\ \hat{r} \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{w} \\ \hat{r} \end{pmatrix} = \frac{1}{|\theta|} \begin{pmatrix} \theta_{2K} & -\theta_{1K} \\ -\theta_{2L} & \theta_{1L} \end{pmatrix} \begin{pmatrix} \hat{p}_1 \\ \hat{p}_2 \end{pmatrix}, \tag{1.10}$$

where $|\theta|$ denotes the determinant of the two-by-two matrix on the left. This determinant can be expressed as

$$\begin{aligned} |\theta| &= \theta_{1L}\theta_{2K} - \theta_{1K}\theta_{2L} \\ &= \theta_{1L}(1 - \theta_{2L}) - (1 - \theta_{1L})\theta_{2L} \\ &= \theta_{1L} - \theta_{2L} = \theta_{2K} - \theta_{1K} \end{aligned} \tag{1.11}$$

where we have repeatedly made use of the fact that $\theta_{iL} + \theta_{iK} = 1$.

In order to fix ideas, let us assume henceforth that *industry 1 is labor intensive.* This implies that its labor cost-share in industry 1 exceeds that in industry 2, $\theta_{1L} - \theta_{2L} > 0$, so that $|\theta| > 0$ in (1.11).[9] Furthermore, suppose that the relative price of good 1 *increases,* so that $\hat{p} = \hat{p}_1 - \hat{p}_2 > 0$. Then we can solve for the change in factor prices from (1.10) and (1.11) as

$$\hat{w} = \frac{\theta_{2K}\hat{p}_1 - \theta_{1K}\hat{p}_2}{|\theta|} = \frac{(\theta_{2K} - \theta_{1K})\hat{p}_1 + \theta_{1K}(\hat{p}_1 - \hat{p}_2)}{(\theta_{2K} - \theta_{1K})} > \hat{p}_1, \tag{1.12a}$$

since $\hat{p}_1 - \hat{p}_2 > 0$, and,

$$\hat{r} = \frac{\theta_{1L}\hat{p}_2 - \theta_{2L}\hat{p}_1}{|\theta|} = \frac{(\theta_{1L} - \theta_{2L})\hat{p}_2 - \theta_{2L}(\hat{p}_1 - \hat{p}_2)}{(\theta_{1L} - \theta_{2L})} < \hat{p}_2, \tag{1.12b}$$

since $\hat{p}_1 - \hat{p}_2 > 0$.

From the result in (1.12a), we see that the wage increases *by more* than the price of good 1, $\hat{w} > \hat{p}_1 > \hat{p}_2$. This means that workers can afford to buy more of good 1 (w/p_1 has gone up), as well as more of good 2 (w/p_2 has gone up). When labor can buy more of *both goods* in this fashion, we say that the *real wage* has increased. Looking at the rental on capital in (1.12b), we see that the rental r changes by *less than* the price of good 2. It follows that capital-owner can afford less of good 2 (r/p_2 has gone down), and also less of good 1 (r/p_1 has gone down). Thus the *real return to capital* has fallen. We can summarize these results with the following theorem.

STOLPER-SAMUELSON (1941) THEOREM

An increase in the relative price of a good will increase the real return to the factor used intensively in that good, and reduce the real return to the other factor.

To develop the intuition for this result, let us go back to the differentiated zero-profit conditions in (1.9′). Since the cost-shares add up to unity in each industry, we see from equation (1.9′) that $\hat{p}_i$ is a *weighted average* of the factor price changes $\hat{w}$ and $\hat{r}$. This implies that $\hat{p}_i$ necessarily lies in between $\hat{w}$ and $\hat{r}$. Putting these together with our assumption that $\hat{p}_1 - \hat{p}_2 > 0$, it is therefore clear that

$$\hat{w} > \hat{p}_1 > \hat{p}_2 > \hat{r}. \tag{1.13}$$

Jones (1965) has called this set of inequalities the "magnification effect": they show that any change in the product price has a *magnified effect* on the factor prices. This is an extremely important result. Whether we think of the product price change as due to export opportunities for a country (the export price goes up), or due to lowering import tariffs (so the import price goes down), the magnification effect says that there will be both gainers and losers due to this change. Even though we will argue in chapter 6 that there are gains from trade in some overall sense, it is still the case that trade opportunities have strong *distributional* consequences, making some people worse off and some better off!

[9] As an exercise, show that $L_1/K_1 > L_2/K_2 \Leftrightarrow \theta_{1L} > \theta_{2L}$. This is done by multipying the numerator and denominator on both sides of the first inequality by like terms, so as to convert it into cost-shares.

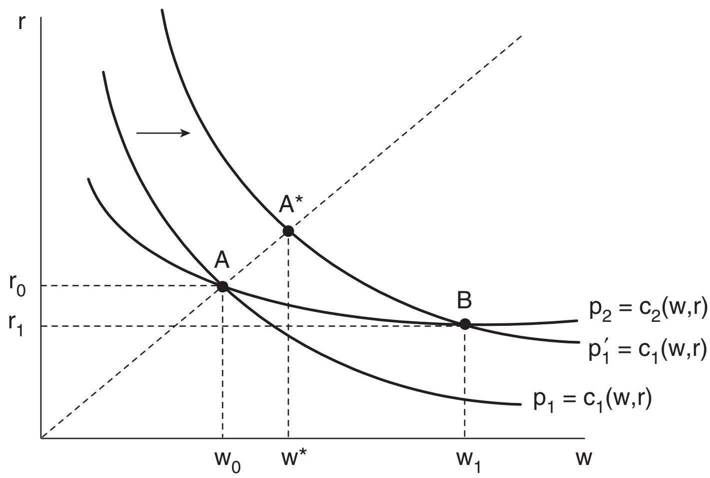

Figure 1.7

We conclude this section by illustrating the Stolper-Samuelson theorem in figure 1.7. We begin with an initial factor price equilibrium given by point A, where industry 1 is labor intensive. An increase in the price of that industry will shift out the iso-cost curve, and as illustrated, move the equilibrium to point B. It is clear that the wage has gone up, from w_0 to w_1, and the rental has declined, from r_0 to r_1. Can we be sure that the wage has increased in percentage terms *by more* than the relative price of good 1? The answer is yes, as can be seen by drawing a ray from the origin through the point A. Because the unit-cost functions are homogeneous of degree one in factor prices, moving along this ray increases p and (w, r) in the same proportion. Thus, at the point A^*, the increase in the wage exactly matched the percentage change in the price p_1. But it is clear that the equilibrium wage increases by *more*, $w_1 > w^*$, so the percentage increase in the wage *exceeds* that of the product price, which is the Stolper-Samuelson result.

CHANGES IN ENDOWMENTS

We turn now to the third key question: if endowments change, how do the industry outputs change? To answer this, we hold the product prices *fixed* and totally differentiate the full-employment conditions (1.8) to obtain

$$\begin{aligned} a_{1L}dy_1 + a_{2L}dy_2 &= dL, \\ a_{1K}dy_1 + a_{2K}dy_2 &= dK. \end{aligned} \tag{1.14}$$

Notice that the a_{ij} coefficients *do not* change, despite the fact that they are functions of the factor prices (w, r). These coefficients are fixed because p_1 and p_2 do not change, so from our earlier lemma, the factor prices are also fixed.

By rewriting the equations in (1.14) using the "Jones algebra," we obtain

$$\begin{aligned} \frac{y_1 a_{1L}}{L}\frac{dy_1}{y_1} + \frac{y_2 a_{2L}}{L}\frac{dy_2}{y_2} &= \frac{dL}{L} \\ \frac{y_1 a_{1K}}{K}\frac{dy_1}{y_1} + \frac{y_2 a_{2K}}{K}\frac{dy_2}{y_2} &= \frac{dK}{K} \end{aligned} \;\Rightarrow\; \begin{aligned} \lambda_{1L}\hat{y}_1 + \lambda_{2L}\hat{y}_2 &= \hat{L} \\ \lambda_{1K}\hat{y}_1 + \lambda_{2K}\hat{y}_2 &= \hat{K}. \end{aligned} \tag{1.14′}$$

To move from the first set of equations to the second, we denote the percentage changes $dy_1/y_1 = \hat{y}_1$, and likewise for all the other variables. In addition, we define $\lambda_{iL} \equiv (y_i a_{iL}/L) = (L_i/L)$, which measures the *fraction of the labor force employed in industry i,* where $\lambda_{1L} + \lambda_{2L} = 1$. We define λ_{iK} analogously as the fraction of the capital stock employed in industry *i*.

This system of equations is written in matrix form and solved as

$$\begin{bmatrix} \lambda_{1L} & \lambda_{2L} \\ \lambda_{1K} & \lambda_{2K} \end{bmatrix} \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} = \begin{pmatrix} \hat{L} \\ \hat{K} \end{pmatrix} \Rightarrow \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \end{pmatrix} = \frac{1}{|\lambda|} \begin{bmatrix} \lambda_{2K} & -\lambda_{2L} \\ -\lambda_{1K} & \lambda_{1L} \end{bmatrix} \begin{pmatrix} \hat{L} \\ \hat{K} \end{pmatrix}, \tag{1.15}$$

where $|\lambda|$ denotes the determinant of the two-by-two matrix on the left, which is simplified as

$$\begin{aligned} |\lambda| &= \lambda_{1L}\lambda_{2K} - \lambda_{2L}\lambda_{1K} \\ &= \lambda_{1L}(1 - \lambda_{1K}) - (1 - \lambda_{1L})\lambda_{1K} \\ &= \lambda_{1L} - \lambda_{1K} = \lambda_{2K} - \lambda_{2L}, \end{aligned} \tag{1.16}$$

where we have repeatedly made use of the fact that $\lambda_{1L} + \lambda_{2L} = 1$ and $\lambda_{1K} + \lambda_{2K} = 1$.

Recall that we assumed *industry 1 to be labor intensive.* This implies that the share of the labor force employed in industry 1 exceeds the share of the capital stock used there, $\lambda_{1L} - \lambda_{1K} > 0$, so that $|\lambda| > 0$ in (1.16).[10] Suppose further that the endowments of labor is increasing, while the endowments of capital remains fixed such that $\hat{L} > 0$, and $\hat{K} = 0$. Then we can solve for the change in outputs from (1.15)–(1.16) as

$$\hat{y}_1 = \frac{\lambda_{2K}}{(\lambda_{2K} - \lambda_{2L})} \hat{L} > \hat{L} > 0 \quad \text{and} \quad \hat{y}_2 = \frac{-\lambda_{1K}}{|\lambda|} \hat{L} < 0. \tag{1.17}$$

From (1.17), we see that the output of the labor-intensive industry 1 expands, whereas the output of industry 2 contracts. We have therefore established the Rybczynski theorem.

RYBCZYNSKI (1955) THEOREM

An increase in a factor endowment will increase the output of the industry using it intensively, and decrease the output of the other industry.

To develop the intuition for this result, let us write the full-employment conditions in vector notation as:

$$\begin{pmatrix} a_{1L} \\ a_{1K} \end{pmatrix} y_1 + \begin{pmatrix} a_{2L} \\ a_{2K} \end{pmatrix} y_2 = \begin{pmatrix} L \\ K \end{pmatrix}. \tag{1.8$'$}$$

We have already illustrated the gradient vectors (a_{iL}, a_{iK}) to the iso-cost curves in figure 1.5 (with not FIR). Now let us take these vectors and regraph them, in figure

[10] As an exercise, show that $L_1/K_1 > L/K > L_2/K_2 \Leftrightarrow \lambda_{1L} > \lambda_{1K}$ and $\lambda_{2K} > \lambda_{2L}$.

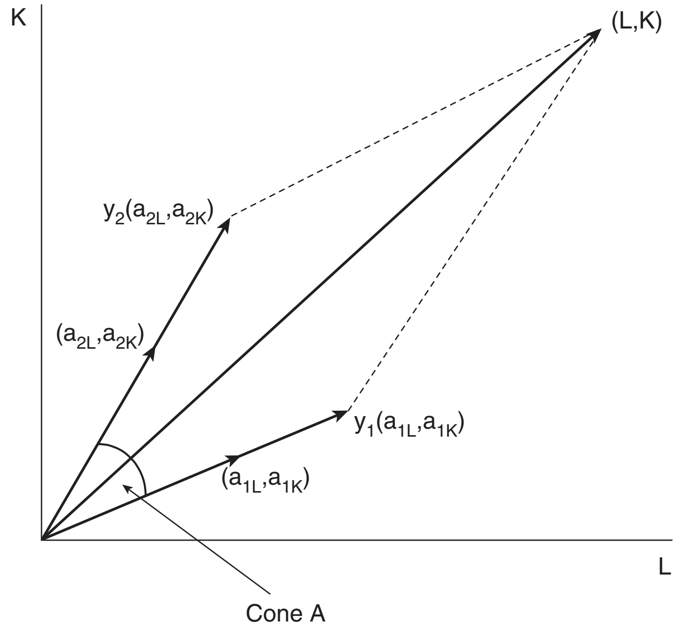

Figure 1.8

1.8. Multiplying each of these by the output of their respective industries, we obtain the total labor and capital demands $y_1(a_{1L}, a_{1K})$ and $y_2(a_{2L}, a_{2K})$. Summing these as in (1.8′) we obtain the labor and capital endowments (L, K). But this exercise can also be performed in reverse: for any endowment vector (L, K), there will be a *unique* value for the outputs (y_1, y_2) such that when (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}) are multiplied by these amounts, they will sum to the endowments.

How can we be sure that the outputs obtained from (1.8′) are positive? It is clear from figure 1.8 that the outputs in both industries will be positive if and only if the endowment vector (L, K) lies *in between* the factor requirement vectors (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}). For this reason, the space spanned by these two vectors is called a "cone of diversification," which we label by cone A in figure 1.8. In contrast, if the endowment vector (L, K) lies *outside* of this cone, then it is *impossible* to add together any positive multiples of the vectors (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}) and arrive at the endowment vector. So if (L, K) lies outside of the cone of diversification, then it must be that only *one good* is produced. At the end of the chapter, we will show how to determine which good it is.[11] For now, we should just recognize that when only one good is produced, the factor prices are determined by the marginal products of labor and capital as in the one-sector model, and will certainly depend on the factor endowments.

Now suppose that the labor endowment increases to $L' > L$, with no change in the capital endowment, as shown in figure 1.9. Starting from the endowments (L', K), the *only* way to add up multiples of (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}) and obtain the endowments is to *reduce* the output of industry 2 to y_2', and *increase* the output of industry 1 to y_1'. This means that not only does industry 1 absorb the entire amount of the extra labor

[11] See problem 1.5.

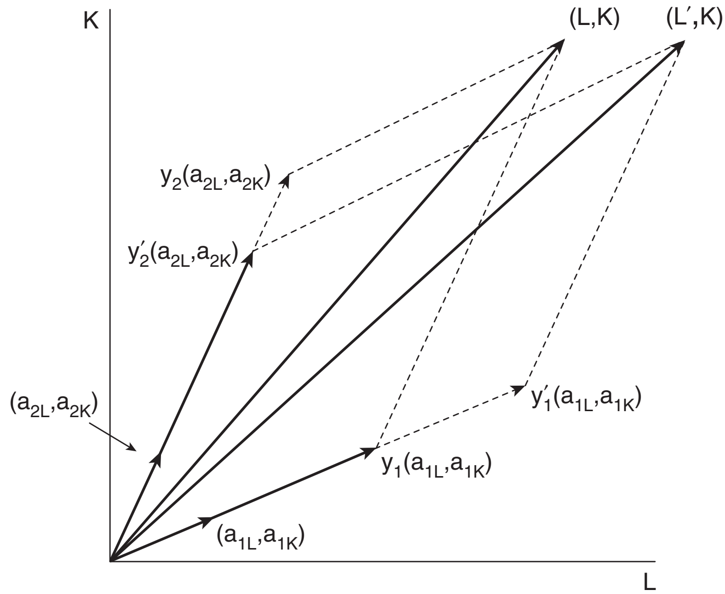

Figure 1.9

endowment, but it also absorbs further labor and capital from industry 2 so that its ultimate labor/capital ratio is unchanged from before. The labor/capital ratio in industry 2 is also unchanged, and this is what permits both industries to pay exactly the same factor prices as they did before the change in endowments.

There are many examples of the Rybczynski theorem in practice, but perhaps the most commonly cited is what is called the "Dutch Disease."[12] This refers to the discovery of oil off the coast of the Netherlands, which led to an increase in industries making use of this resource. (Shell Oil, one of the world's largest producers of petroleum products, is a Dutch company.) At the same time, however, other "traditional" export industries of the Netherlands contracted. This occurred because resources were attracted away from these industries and into those that were intensive in oil, as the Rybczynski theorem would predict.

We have now answered the three questions raised earlier in the chapter: how are factor prices determined; how do changes in product prices affect factor prices; and how do changes in endowments affect outputs? But in answering all of these, we have relied on the assumptions that *both goods are produced*, and also that factor intensity reversals do not occur, as was stated explicitly in the FPE theorem. In the remainder of this chapter we need to investigate both of these assumptions, to understand either when they will hold or the consequences of their not holding.

We begin by tracing through the changes in the outputs induced by changes in endowments, along the equilibrium of the production possibility frontier. As the labor endowment grows in figure 1.9, the PPF will shift out. This is shown in figure 1.10, where the outputs will shift from point A to point A' with an increase of good 1 and reduction of good 2, at the unchanged price p. As the endowment of labor rises, we

[12] See, for example, Corden and Neary (1982) and Jones, Neary, and Ruane (1987).

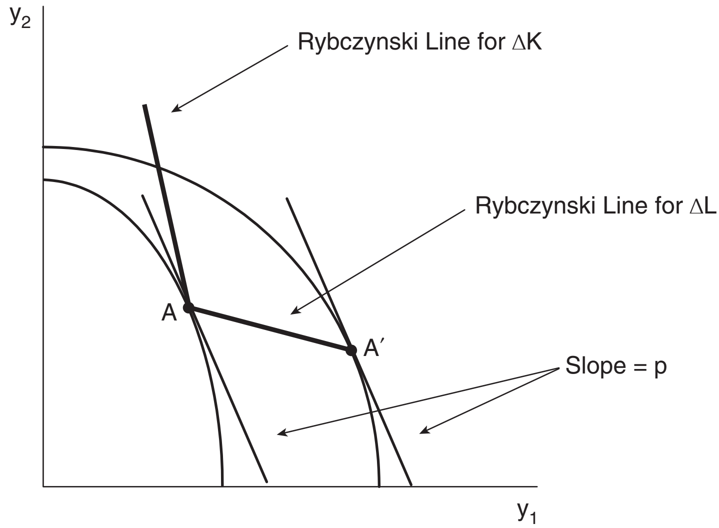

Figure 1.10

can join up all points such as A and A' where the slopes of the PPFs are equal. These form a downward-sloping line, which we will call the Rybczynski line for changes in labor (ΔL). The Rybczynski line for ΔL indicates how outputs change as labor endowment expands.

Of course, there is also a Rybczynski line for ΔK, which indicates how the outputs change as the capital endowment grows: this would lead to an increase in the output of good 2, and reduction in the output of good 1. As drawn, both of the Rybczynski lines are illustrated as *straight* lines: can we be sure that this is the case? The answer is yes: the fact that the product prices are fixed along a Rybczynski line, implying that factor prices are also fixed, ensures that these are straight lines. To see this, we can easily calculate their slopes by differentiating the full-employment conditions (1.8). To compute the slope of the Rybczynski line for ΔL, it is convenient to work with the full-employment condition for *capital,* since that endowment does not change. Total differentiating (1.8) for capital gives

$$a_{1K}y_1 + a_{2K}y_2 = K \Rightarrow a_{1K}dy_1 + a_{2K}dy_2 = 0 \Rightarrow \frac{dy_2}{dy_1} = -\frac{a_{1K}}{a_{2K}}. \tag{1.18}$$

Thus, the slope of the Rybczynski line for ΔL is the negative of the ratio of capital/output in the two industries, which is constant for fixed prices. This proves that the Rybczynski lines are indeed straight.

If we continue to increase the labor endowment, outputs will move downward on the Rybczynski line for ΔL in figure 1.10, until this line hits the y_1 axis. At this point the economy is fully specialized in good 1. In terms of figure 1.9, the vector of endowments (L, K) is coincident with the vector of factor requirements (a_{1L}, a_{1K}) in industry 1. For further increases in the labor endowment, the Rybczynski line for ΔL then *moves right along the* y_1 *axis* in figure 1.10, indicating that the economy remains specialized in good 1. This corresponds to the vector of endowments (L, K)

lying outside and below the cone of diversification in figure 1.9. With the economy fully specialized in good 1, factor prices are determined by the marginal products of labor and capital in that good, and the earlier "factor price insensitivity" lemma no longer applies.

FACTOR PRICE EQUALIZATION REVISITED

Our finding that the economy produces both goods whenever the factor endowments remain *inside* the cone of diversification allows us to investigate the FPE theorem more carefully. Let us continue to assume that there are no FIRs, but now rather than *assuming* that both goods are produced in both countries, we will instead derive this as an *outcome* from the factor endowments in each country. To do so, we engage in a thought experiment posed by Samuelson (1949) and further developed by Dixit and Norman (1980).

Initially, suppose that labor and capital are *free to move* between the two countries until their factor prices are equalized. Then all that matters for factor prices are the *world* endowments of labor and capital, and these are shown as the length of the horizontal and vertical axis in figure 1.11. The amounts of labor and capital choosing to reside at home are measured relative to the origin 0, while the amounts choosing to reside in the foreign country are measured relative to the origin 0^*; suppose that this allocation is at point *B*. Given the world endowments, we establish equilibrium prices for goods and factors in this "integrated world equilibrium." The factor prices determine the demand for labor and capital in each industry (assuming no FIR), and using

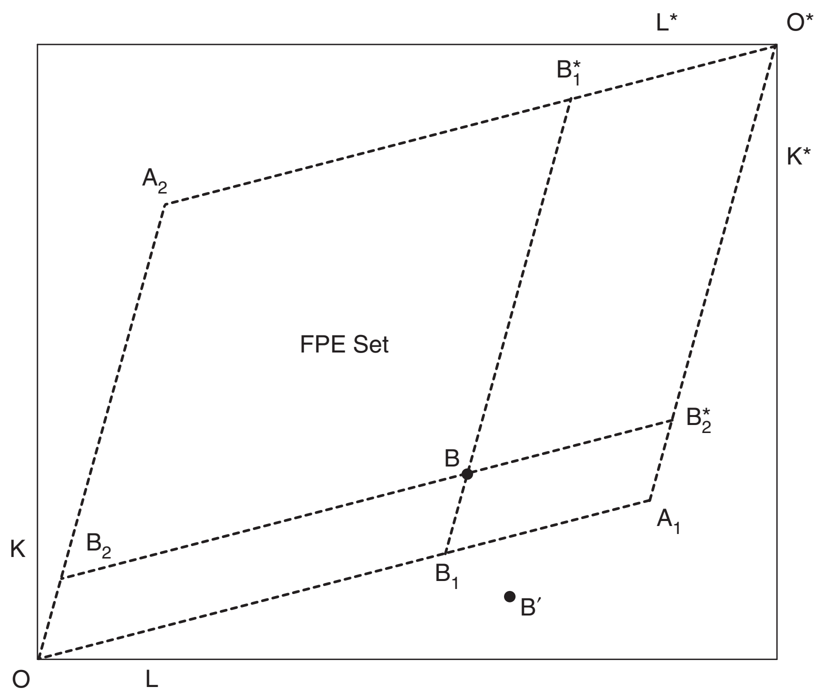

Figure 1.11

these, we can construct the diversification cone (since factor prices are the same across countries, then the diversification cone is also the same). Let us plot the diversification cone relative to the home origin 0, and again relative to the foreign origin 0^*. These cones form the parallelogram $0A_1 0^* A_2$.

For later purposes, it is useful to identify precisely the points A_1 and A_2 on the vertices of this parallelogram. The vectors $0A_i$ and $0^* A_i$ are proportional to (a_{iL}, a_{iK}), the amount of labor and capital used to produce one unit of good i in each country. Multiplying (a_{iL}, a_{iK}) by world demand for good i, D_i^w, we then obtain the *total* labor and capital used to produce that good, so that $A_i = (a_{iL}, a_{iK}) D_i^w$. Summing these gives the total labor and capital used in world demand, which equals the labor and capital used in world production, or world endowments.

Now we ask whether we can achieve exactly the same world production and equilibrium prices as in this "integrated world equilibrium," but *without* labor and capital mobility. Suppose there is some allocation of labor and capital endowments across the countries, such as point B. Then can we produce the same amount of each good as in the "integrated world equilibrium"? The answer is clearly yes: with labor and capital in each country at point B, we could devote $0B_1$ of resources to good 1 and $0B_2$ to good 2 at home, while devoting $0^*B_1^*$ to good 1 and $0^*B_2^*$ toward good 2 abroad. This will ensure that the same amount of labor and capital worldwide is devoted to each good as in the "integrated world equilibrium," so that production and equilibrium prices must be the same as before. Thus, we have achieved the same equilibrium but without factor mobility. It will become clear in the next chapter that there is still *trade in goods* going on to satisfy the demands in each country.

More generally, for *any allocation* of labor and capital within the parallelogram $0A_1 0^* A_2$ both countries remain diversified (producing both goods), and we can achieve the same equilibrium prices as in the "integrated world economy." It follows that factor prices *remain equalized across countries* for allocations of labor and capital within the parallelogram $0A_1 0^* A_2$, which is referred to as the *factor price equalization (FPE) set*. The FPE set illustrates the range of labor and capital endowments between countries over which both goods are produced in both countries, so that factor price equalization is obtained. In contrast, for endowments *outside* of the FPE set such as point B', then at least one country would have to be fully specialized in one good and FPE no longer holds.

FACTOR INTENSITY REVERSALS

We conclude this chapter by returning to a question raised earlier: when there are "factor intensity reversals" giving multiple solutions to the zero-profit conditions, how do we know which solution will prevail in each country? To answer this, it is necessary to combine the zero-profit with the full-employment conditions, as follows.

Consider the case in figure 1.6, where the zero-profit conditions allows for two solutions to the factor prices. Each of these determine the labor and capital demands shown orthogonal to the iso-cost curves, labeled as (a_{1L}, a_{1K}) and (a_{2L}, a_{2K}), and (b_{1L}, b_{1K}) and (b_{2L}, b_{2K}). We have redrawn these in figure 1.12, after multiplying each of them by the outputs of their respective industries. These vectors create *two* cones of diversification, labeled as cones A and B. Initially, suppose that the factor endowments for each country lie within one cone or the other (then we will consider the case where the endowments are outside both cones).

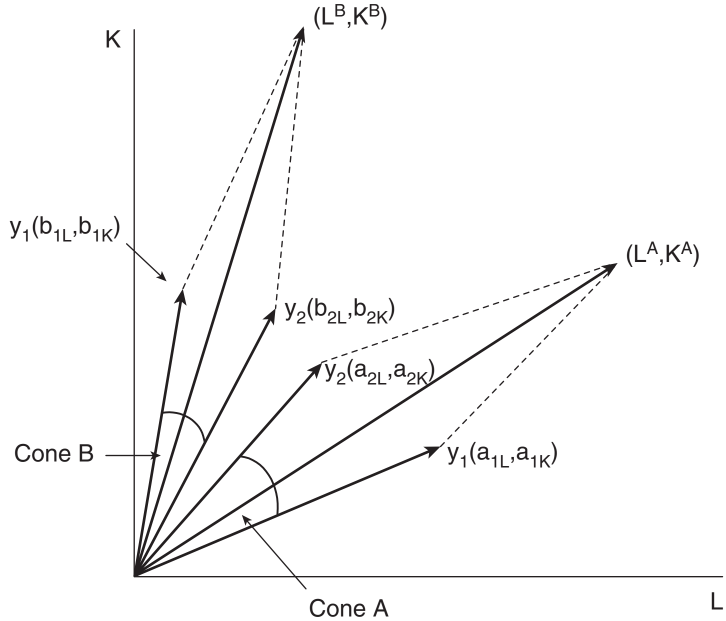

Figure 1.12

Now we can answer the question of which factor prices will apply in each country: a *labor-abundant* economy, with a high ratio of labor/capital endowments, such as (L^A, K^A) in cone A of figure 1.12, will have factor prices given by (w^A, r^A) in figure 1.6, with low wages; whereas a *capital-abundant* economy with a high ratio of capital/labor endowments such as shown by (L^B, K^B) in cone B of figure 1.12, will have factor prices given by (w^B, r^B) in figure 1.6, with high wages. Thus, factor prices depend on the endowments of the economy. A labor-abundant country such as China will pay low wages and a high rental (as in cone A), while a capital-abundant country such as the United States will have high wages and a low rental (as in cone B). Notice that we have now reintroduced a link between factor endowments and factor prices, as we argued earlier in the one-sector model: when there are FIR in the two-by-two model, factor prices vary systematically with endowments *across* the cones of diversification, even though factor prices are independent of endowments *within* each cone.

What if the endowment vector of a country does not lie in either cone? Then the country will be fully specialized in one good or the other. Generally, we can determine which good it is by tracing through how the outputs change as we move through the cones of diversification, and it turns out that outputs depend *non-monotonically* on the factor endowments.[13] For example, textiles in South Korea or Taiwan expanded during the 1960s and 1970s, but contracted later as capital continued to grow. Despite the complexity involved, many trade economists feel that countries do in fact produce in different cones of diversification, and taking this possibility into account is a topic of research.[14]

[13] See problem 1.5.

[14] Empirical evidence on whether developed countries fit into the same cone is presented by Debaere and Demiroğlu (2003), and the presence of multiple cones is explored by Leamer (1987), Harrigan and Zakrajšek (2000), and Schott (2003).

CONCLUSIONS

In this chapter we have reviewed several two-sector models: the Ricardian model, with just one factor, and the two-by-two model, with two factors, both of which are fully mobile between industries. There are other two-sector models, of course: if we add a third factor, treating capital as specific to each sector but labor as mobile, then we obtain the Ricardo-Viner or "specific-factors" model, as will be discussed in chapter 3. We will have an opportunity to make use of the two-by-two model throughout this book, and a thorough understanding of its properties— both the equations and the diagrams *labor-abundant* economy—is essential for all the material that follows.

One special feature of this chapter is the dual determination of factor prices, using the unit-cost function in the two industries. This follows the dual approach of Woodland (1977, 1982), Mussa (1979), and Dixit and Norman (1980). Samuelson (1949) uses a quite different diagramatic approach to prove the FPE theorem. Another method that is quite commonly used is the so-called Lerner (1952) diagram, which relies on the production rather than cost functions.[15] We will not use the Lerner diagram in this book, but it will be useful to understand some articles, for example, Findlay and Grubert (1959) and Deardorff (1979), so we include a discussion of it in the appendix to this chapter.

This is the only chapter where we do not present any accompanying empirical evidence. The reader should not infer from this that the two-by-two model is unrealistic: while it is usually necessary to add more goods or factors to this model before confronting it with data, the relationships between prices, outputs, and endowments that we have identified in this chapter will carry over in some form to more general settings. Evidence on the pattern of trade is presented in the next chapter, where we extend the two-by-two model by adding another country, and then many countries, trading with each other. We also allow for many goods and factors, but for the most part restrict attention to situations where factor price equalization holds. In chapter 3, we examine the case of many goods and factors in greater detail, to determine whether the Stolper-Samuelson and Rybczynski theorems generalize and also how to estimate these effects. In chapter 4, evidence on the relationship between product prices and wages is examined in detail, using a model that allows for trade in intermediate inputs. The reader is already well prepared for the chapters that follow, based on the tools and intuition we have developed from the two-by-two model. Before moving on, you are encouraged to complete the problems at the end of this chapter.

APPENDIX: THE LERNER DIAGRAM AND FACTOR PRICES

The Lerner (1952) diagram for the two-by-two model can be explained as follows: With perfect competition and constant returns to scale, we have that revenue = costs in both industries. So let us choose a special isoquant in each industry such that revenue = 1. In each industry, we therefore choose the isoquant $p_i y_i = 1$, or

$$Y = f_i(L_i, K_i) = 1/p_i \Rightarrow wL_i + rK_i = 1.$$

[15]This diagram was used in a seminar presented by Abba Lerner at the London School of Economics in 1933, but not published until 1952. The history of this diagram is described at the "Origins of Terms in International Economics," maintained by Alan Deardorff at http://www-personal.umich.edu/~alandear/glossary/orig.html. See also Samuelson (1949, 181 n.1).

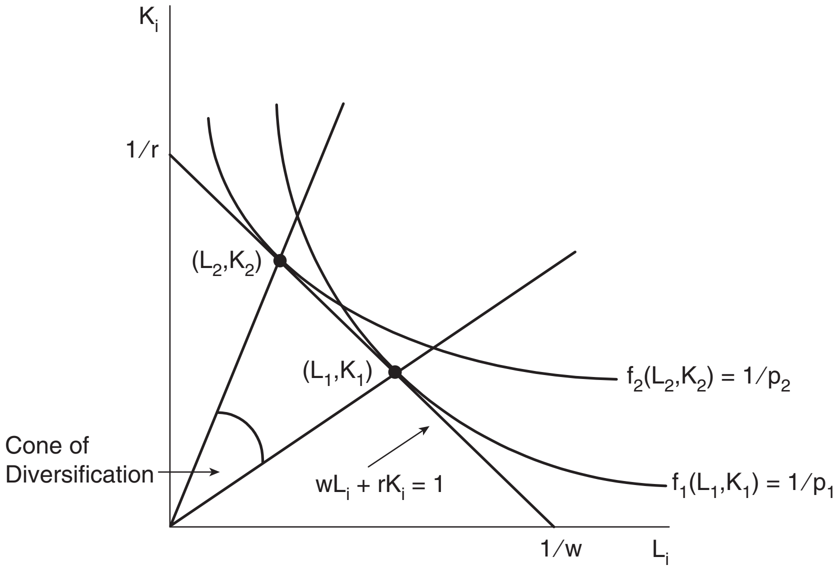

Figure 1.13

Therefore, from cost minimization, the $1/p_i$ isoquant in each industry will be *tangent* to the line $wL_i + rK_i = 1$. This is the same line for both industries, as shown in figure 1.13.

Drawing the rays from the origin through the points of tangency, we obtain the cone of diversification, as labeled in figure 1.13. Furthermore, we can determine the factor prices by computing where $wL_i + rK_i = 1$ intersects the two axis: $L_i = 0 \Rightarrow K_i = 1/r$, and $K_i = 0 \Rightarrow L_i = 1/w$. Therefore, given the prices p, we determine the two isoquants in figure 1.13, and drawing the (unique) line tangent to both of these, we determine the factor prices as the intercepts of this line. Notice that these equilibrium factor prices do not depend on the factor endowments, provided that the endowment vector lies within the cone of diversification (so that both goods are produced). We have thus obtained an alternative proof of the "factor price insensitivity" lemma, using a primal rather than dual approach. Furthermore, with two countries having the same prices (through free trade) and technologies, then figure 1.13 holds in both of them. Therefore, their factor prices will be equalized.

Lerner (1952) also showed how figure 1.13 can be extended to the case of factor intensity reversals, in which case the isoquants intersect twice. In that case there will be *two* lines $wL_i + rK_i = 1$ that are tangent to both isoquants, and there are two cones of diversification. This is shown in figure 1.14. To determine which factor prices apply in a particular country, we plot its endowments vector and note which cone of diversification it lies in: the factor prices in this country are those applying to that cone. For example, the endowments (L^A, K^A) will have the factor prices (w^A, r^A), and the endowments (L^B, K^B) will have the factor prices (w^B, r^B). Notice that the labor-abundant country with endowments (L^A, K^A) has the low wage and high rental, whereas the capital-abundant country with endowments (L^B, K^B) has the high wage and low rental.

How likely is it that the isoquants of industries 1 and 2 intersect twice, as in figure 1.14? Lerner (1952, 11) correctly suggested that it depends on the elasticity of substitution between labor and capital in each industry. For simplicity, suppose that each

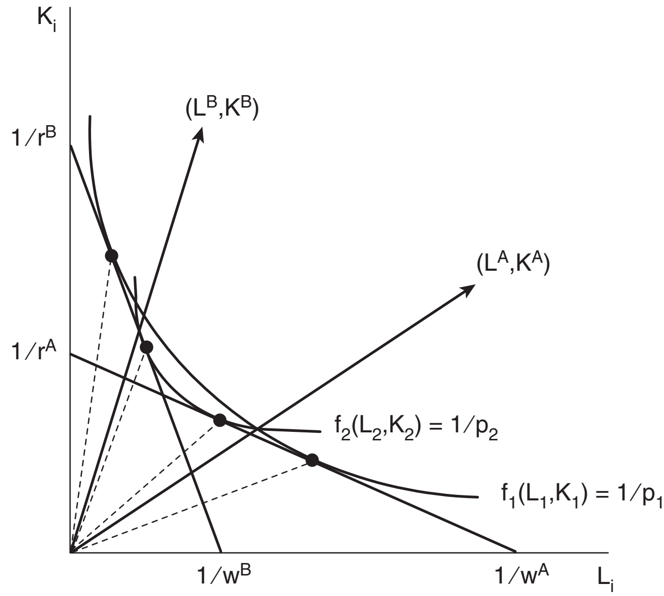

Figure 1.14

industry has a constant elasticity of substitution production function. If the elasticities are the same across industries, then it is impossible for the isoquants to intersect twice. If the elasticities of substitution differ across industries, however, and we choose prices p_i, $i=1,2$, such that the $1/p_i$ isoquants intersect at least once, then it is *guaranteed* that they intersect twice. Under exactly the same conditions, the iso-cost lines in figure 1.6 intersect twice. Thus, the occurrence of FIR is very likely once we allow elasticities of substitution to differ across industries. Minhas (1962) confirmed that this was the case empirically, and discussed the implications of FIR for factor prices and trade patterns. This line of empirical research was dropped thereafter, perhaps because FIR seemed too complex to deal with, and has been picked up again more recently (see note 14 of this chapter).

PROBLEMS

1.1 Rewrite the production function $y_1 = f_1(L_1, K_1)$ as $y_1 = f_1(v_1)$, and similarly, $y_2 = f_2(v_2)$. Concavity means that given two points $y_1^a = f_1(v_1^a)$ and $y_1^b = f_1(v_1^b)$, and $0 \le \lambda \le 1$, then $f_1(\lambda v_1^a + (1-\lambda)v_1^b) \ge \lambda y_1^a + (1-\lambda)y_1^b$. Similarly for the production function $y_2 = f_2(v_2)$. Consider two points $y^a = (y_1^a, y_2^a)$ and $y^b = (y_1^b, y_2^b)$, both of which can be produced while satisfying the full-employment conditions $v_1^a + v_2^a \le V$ and $v_1^b + v_2^b \le V$, where V represents the endowments. Consider a production point midway between these, $\lambda y^a + (1-\lambda)y^b$. Then use the concavity of the production functions to show that this point can *also* be produced while satisfying the full-employment conditions. This proves that the production possibilities set is *convex*. (Hint: Rather than showing that $\lambda y^a + (1-\lambda)y^b$ can be produced while satisfying the full-employment conditions, consider instead allocating $\lambda v_1^a + (1-\lambda) v_1^b$ of the resources to industry 1, and $\lambda v_2^a + (1-\lambda) v_2^b$ of the resources to industry 2.)

1.2 Any function $y = f(v)$ is homogeneous of degree α if for all $\lambda > 0$, $f(\lambda v) = \lambda^\alpha f(v)$. Consider the production function $y = f(L, K)$, which we assume is homogeneous of degree one, so that $f(\lambda L, \lambda K) = \lambda f(L, K)$. Now differentiate this expression with respect to L, and answer the following: Is the marginal product $f_L(L, K)$ homogeneous, and of what degree? Use the expression you have obtained to show that $f_L(L/K, 1) = f_L(L, K)$.

1.3 Consider the problem of maximizing $y_1 = f_1(L_1, K_1)$, subject to the full-employment conditions $L_1 + L_2 \le L$ and $K_1 + K_2 \le K$, and the constraint $y_2 = f_2(L_2, K_2)$. Set this up as a Lagrangian, and obtain the first-order conditions. Then use the Lagrangian to solve for dy_1/dy_2, which is the slope of the production possibilities frontier. How is this slope related to the marginal product of labor and capital?

1.4 Consider the problem of maximizing $p_1 f_1(L_1, K_1) + p_2 f_2(L_2, K_2)$, subject to the full-employment constraints $L_1 + L_2 \le L$ and $K_1 + K_2 \le K$. Call the result the GDP function $G(p, L, K)$, where $p = (p_1, p_2)$ is the price vector. Then answer the following:
   (a) What is $\partial G/\partial p_i$? (Hint: we solved for this in the chapter.)
   (b) Give an economic interpretation to $\partial G/\partial L$ and $\partial G/\partial K$.
   (c) Give an economic interpretation to $\partial^2 G/\partial p_i \partial L = \partial^2 G/\partial L \partial p_i$, and $\partial^2 G/\partial p_i \partial K = \partial^2 G/\partial K \partial p_i$.

1.5 Trace through changes in outputs when there are factor intensity reversals. That is, construct a graph with the capital endowment on the horizontal axis, and the output of goods 1 and 2 on the vertical axis. Starting at a point of diversification (where both goods are produced) in cone A of figure 1.12, draw the changes in output of goods 1 and 2 as the capital endowment grows outside of cone A, into cone B, and beyond this.

2

The Heckscher-Ohlin Model

We begin this chapter by describing the Heckscher-Ohlin (HO) model with two countries, two goods, and two factors (or the two-by-two-by-two model). This formulation is often called the Heckscher-Ohlin-Samuelson (HOS) model, which is based on the work of Paul Samuelson, who developed a mathematical model from the original insights of Eli Heckscher and Bertil Ohlin.[1] The goal of that model is to predict the pattern of trade in goods between the two countries, based on their differences in factor endowments. Following this, we present the multigood, multifactor extension that is associated with the work of Vanek (1968), and is often called the Heckscher-Ohlin-Vanek (HOV) model. As we shall see, in this latter formulation we do not attempt to keep track of the trade pattern in individual goods but instead compute the "factor content" of trade, that is, the amounts of labor, capital, land, and so on embodied in the exports and imports of a country.

The factor-content formulation of the HOV model has led to a great deal of empirical research, beginning with Leontief (1953) and continuing with Leamer (1980), Bowen, Leamer, and Sveikauskas (1987), Trefler (1993a, 1995), Davis and Weinstein (2001), with many other writers in between. We will explain the twists and turns in this chain of empirical research. The bottom line is that the HOV model performs quite poorly empirically unless we are willing to dispense with the assumption of identical technologies across countries. This brings us back to the earlier tradition of the Ricardian model of allowing for technological differences, which also implies differences in factor prices across countries. We will show several ways that technological differences can be incorporated into an "extended" HO model, with their empirical results, and this remains a question of ongoing research.

HECKSCHER-OHLIN-SAMUELSON (HOS) MODEL

The basic assumptions of the HOS model were already introduced in the previous chapter: identical technologies across countries; identical and homothetic tastes across countries; differing factor endowments; and free trade in goods (but not factors). For the most part, we will also assume away the possibility of factor intensity

[1] The original 1919 article of Heckscher and the 1924 dissertation of Ohlin have been translated from Swedish and edited by Harry Flam and June Flanders and published as Heckscher and Ohlin (1991).

reversals. Provided that all countries have their endowments *within* their "cone of diversification," this means that factor prices are equalized across countries

We begin by supposing that there are just two countries, two sectors and two factors, exactly like the two-by-two model we introduced in chapter 1. We shall assume that the home country is labor abundant, so that $L/K > L^*/K^*$. We will also assume that good 1 is labor intensive. The countries engage in free trade, and we also suppose that trade is balanced (value of exports = value of imports). Then the question is, What is the pattern of trade in goods between the countries? This is answered by the following theorem.

HECKSCHER-OHLIN THEOREM

Each country will export the good that uses its abundant factor intensively.

Thus, under our assumptions the home country will export good 1 and the foreign country will export good 2. To prove this, let us take a particular case of the factor endowment differences $L/K > L^*/K^*$, and assume that the labor endowments are identical in the two countries, $L^* = L$, while the foreign capital endowments exceed that at home, $K^* > K$.[2] In order to derive the pattern of trade between the countries, we proceed by first establishing what the relative product price is in each country *without* any trade, or *in autarky.* As we shall see, the pattern of autarky prices can then be used to predict the pattern of trade: a country will export the good whose free trade price is higher than its autarky price, and import the other.

Let us begin by illustrating the *home autarky* equilibrium, at point A in figure 2.1. We have assumed a representative consumer with homothetic tastes, so we can use indifference curves to reflect demand. The autarky equilibrium is established where an indifference curve is tangent to the home production possibility frontier (PPF, at point A. The price line drawn tangent to the PPF and indifference curve has a slope of (negative) the autarky relative price of good 1, $p^a \equiv p_1^a/p_2^a$. Let us now consider the foreign PPF, which is drawn outside the home PPF in the figure 2.1. In order to determine where the foreign autarky equilibrium lies, let us initially suppose that p^a is *also* the autarky equilibrium abroad, and see whether this assumption leads to a contradiction.

If p^a is also the autarky price in the foreign country, then production must occur at the tangency between the price line with slope p^a and the foreign PPF, or at point B'. Notice that from the Rybcynzski theorem, point B' must lie *above and to the left* of point A: the higher capital endowment abroad leads to more of good 2 and less of good 1. The price line through point B' acts like a budget constraint for the representative consumer in the foreign country, so that the consumer chooses the highest indifference curve on this price line. Since tastes are homothetic, the foreign representative consumer will demand the two goods in exactly the same proportion as does the home representative consumer. In other words, the foreign consumption point must lie on the budget constraint through point B', and also on a ray from the origin through point A. Thus, foreign consumption must occur at point C', which is *above and to the right* of point A. Since points B' and C' do not coincide, we have arrived at a contradiction: the relative price p^a at home *cannot* equal the autarky price abroad, and on the

[2]Because of the assumptions of identical homothetic tastes and constant returns to scale, the result we are establishing remains valid if the labor endowments also differ across countries.

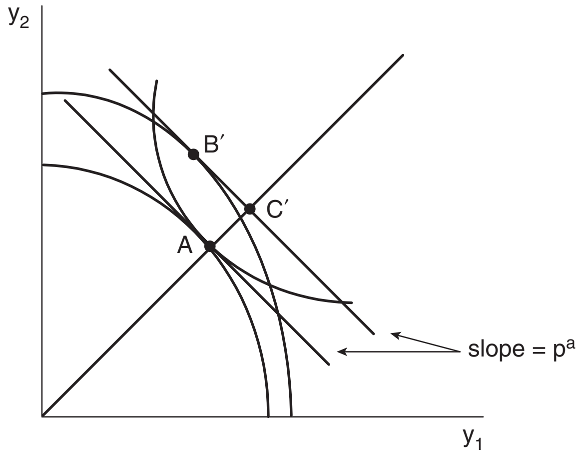

Figure 2.1

contrary, at this price there is an excess demand for good 1 in the foreign country. This excess demand will bid up the relative price of good 1, so that the foreign autarky price must be *higher* than at home, $p^{a*} > p^a$.

To establish the free trade equilibrium price, let $z(p)$ denote the excess demand for good 1 at any prevailing price p at home, while $z^*(p^*)$ denotes the excess demand for good 1 abroad. World excess demand at a common price is therefore $z(p) + z^*(p)$, and a free-trade equilibrium occurs when world excess demand is equal to zero. The home autarky equilibrium satisfies $z(p^a) = 0$, and we have shown above that $z^*(p^a) > 0$. It follows that $z(p^a) + z^*(p^a) > 0$. If instead we reversed the argument in figure 2.1, and started with the foreign autarky price satisfying $z^*(p^{a*}) = 0$, then we could readily prove that $z(p^{a*}) < 0$, so at the foreign autarky price there is excess supply of good 1 at home. It follows that world excess demand would satisfy $z(p^{a*}) + z^*(p^{a*}) < 0$. Then by continuity of the excess demand functions, there must be a price p, with $p^{a*} > p > p^a$, such that $z(p) + z^*(p) = 0$. This is the equilibrium price with free trade.

Let us illustrate the free trade equilibrium, in figure 2.2. In panel (a) we show the equilibrium at home, and in panel (b) we show the equilibrium in the foreign country. Beginning at the home autarky point A, the relative price of good 1 *rises* at home, $p > p^a$. It follows that production will occur at a point like B, where the price line through point B has the slope p. Once again, this price line acts as a budget constraint for the representative consumer, and utility is maximized at point C. The difference between production at point B and consumption at point C is made up through exporting good 1 and importing good 2, as illustrated by the "trade triangle" drawn. This trade pattern at home establishes the HO theorem stated above. In the foreign country, the reverse pattern occurs: the relative price of good 1 *falls*, $p^{a*} > p$, and production moves from the autarky equilibrium point A^* to production at point B^* and consumption at point C^*, where good 1 is imported and good 2 is exported. Notice that the trade triangles drawn at home and abroad are *identical* in size: the exports of one country must be imports of the other.[3]

[3] Also, notice that the slope of the hypotenuse of the trade triangle for the home country is (import$_2$/export$_1$) = p. It follows that ($p \cdot$ export$_1$) = import$_2$, so that trade is balanced, and this also holds for the foreign country.

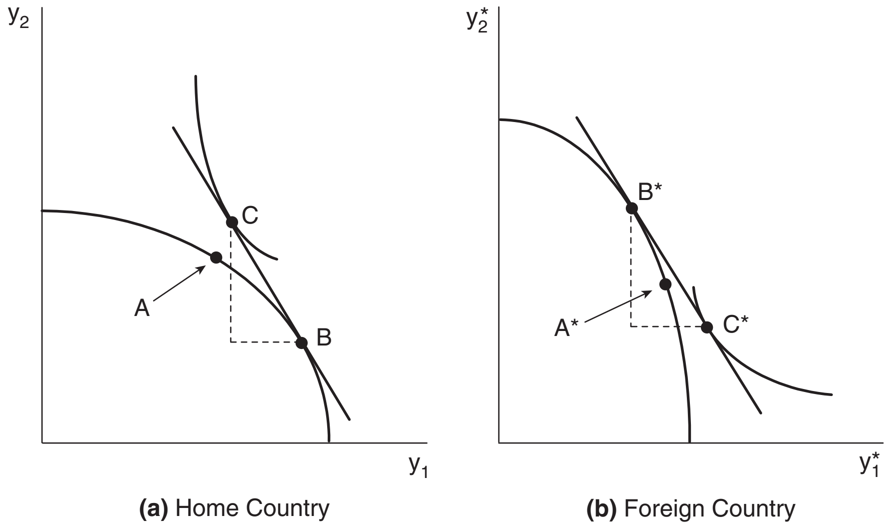

Figure 2.2

In addition to establishing the trade pattern, the HO model has precise implications about who gains and who loses from trade: the *abundant* factor in each country gains from trade, and the *scarce* factor loses. This result follows from the pattern of price changes ($p^{a*} > p > p^a$) and the Stolper-Samuelson theorem. With the relative price of good 1 rising at home, the factor used intensively in that good (labor) will gain in real terms, and the other factor (capital) will lose. Notice that labor is the abundant factor at home. The fact that $L/K > L^*/K^*$ means that labor would have been earning less in the home autarky equilibrium than in the foreign autarky equilibrium: its marginal product at home would have been lower (in both goods) than abroad. However, with free trade the home country can shift production toward the labor-intensive good, and export it, thereby absorbing the abundant factor without lowering its wage. Indeed, factor prices are *equalized* in the two countries after trade, as we argued in the previous chapter. Thus, the abundant factor, whose factor price was bid down in autarky, will gain from the opening of trade, while the scarce factor in each country loses.

Our presentation of the HO model above is about as far as most discussion of this model goes at the undergraduate level. After showing something like figure 2.2, it would be common to provide some rough data or anecdotes to illustrate the HO theorem (e.g., the United States is abundant in scientists, so it exports high-tech goods; Canada is abundant in land, so it exports natural resources, etc.). As plausible as these illustrations are, it turns out that the HO model is a *rather poor* predictor of actual trade patterns, indicating that its assumptions are not realistic. It has taken many years, however, to understand why this is the case, and we begin this exploration by considering the earliest results of Leontief (1953).

LEONTIEF'S PARADOX

Leontief (1953) was the first to confront the HO model with data. He had developed the set of input-output accounts for the U.S. economy, which allowed him to compute

TABLE 2.1: Leontief's (1953) Test

	Exports	Imports
Capital ($ million)	$2.5	$3.1
Labor (person-years)	182	170
Capital/Labor ($/person)	$13,700	$18,200

Note: Each column shows the amount of capital or labor needed per $1 million worth of exports or imports into the United States, for 1947.

the amounts of labor and capital used in each industry for 1947. In addition, he utilized U.S. trade data for the same year to compute the amounts of labor and capital used in the production of $1 million of U.S. exports and imports. His results are shown in table 2.1.

Leontief first measured the amount of capital and labor required for $1 million worth of U.S. exports. This calculation requires that we measure the labor and capital used *directly,* that is, in each exporting industry, and also those factors used *indirectly,* that is, in the industries that produce intermediate inputs that are used in producing exports. From the first row of table 2.1, we see that $2.5 million worth of capital was used in $1 million of exports. This amount of capital seems much too high, until we recognize that what is being measured is the *capital stock,* so that only the annual depreciation on this stock is actually used. For labor, 182 person-years were used to produce the exports. Taking the ratio of these, we find that each person employed in producing exports (directly or indirectly) is working with $13,700 worth of capital.

Turning to the import side of the calculation, we immediately run into a problem: it is not possible to measure the amount of labor and capital used in producing imports unless we have knowledge of the *foreign* technologies, which Leontief certainly did not know in 1953! Indeed, it is only recently that researchers have begun to use data on foreign technologies to test the HO model, as we will describe later in the chapter. So Leontief did what many researchers have done since: he simply used the *U.S. technology* to calculate the amount of labor and capital used in imports. Does this invalidate the test of the HO model? Not really, because recall that an assumption of the HO model is that technologies are the same across countries. Thus, under the null hypothesis that the HO model is true, it would be valid to use the U.S. technology to measure the labor and capital used in imports. If we find that this null hypothesis is rejected, then one explanation would be that the assumption of identical technologies is false.

Using the U.S. technology to measure the labor and capital used in imports, both directly and indirectly, we arrive at the estimates in the last column of table 2.1: $3.1 million of capital, 170 person-years, and so a capital/labor ratio in imports of $18,200. Remarkably, this is *higher* than the capital/labor ratio found for U.S. exports! Under the presumption that the United States was capital-abundant in 1947, this appears to contradict the HO theorem. Thus, this finding came to be called "Leontief's paradox."

A wide range of explanations have been offered for this paradox:

- U.S. and foreign technologies are not the same;
- By focusing only on labor and capital, Leontief ignored land;
- Labor should have been disaggregated by skill (since it would not be surprising to find that U.S. exports are intensive in skilled labor);

- The data for 1947 may by unusual, since World War II had just ended;
- The U.S. was not engaged in free trade, as the HO model assumes.

These reasons are all quite valid criticisms of the test that Leontief performed, and research in the years following his test aimed to redo the analysis while taking into account land, skilled versus unskilled labor, other years, and so on. This research is well summarized by Deardorff (1984a), and the general conclusion is that the paradox continued to occur in some cases. It was not until two decades later, however, that Leamer (1980) provided the definitive critique of the Leontief paradox: it turned out that Leontief had performed the wrong test! That is, even if the HO model is true, it turns out the capital/labor ratios in export and imports, as reported in table 2.1, should not be compared. Instead, an alternative test should be performed. The test that Leamer proposed relies on the "factor content" version of the Heckscher-Ohlin model, developed by Vanek (1968), which we turn to next.

HECKSCHER-OHLIN-VANEK (HOV) MODEL

Let us now consider many countries, indexed by $i=1,\dots,C$; many industries, indexed by $j=1,\dots,N$; and many factors, indexed by k or $\ell=1,\dots,M$. We will continue to assume that technologies are identical across countries, and that factor price equalization prevails under free trade. In addition, we assume that tastes are identical and homothetic across countries.

Let the $(M\times N)$ matrix $A=[a_{jk}]'$ denote the amounts of labor, capital, land, and other primary factors needed for one unit of production in each industry.[4] Notice that this matrix applies in any country. The *rows* measure the different factors k, $\ell=1,\dots,M$, while the *columns* of this matrix measure the different industries $j=1,\dots,N$. For example, with just two industries using only labor and capital, this matrix would be $A=\begin{bmatrix} a_{1L} & a_{2L} \\ a_{1K} & a_{2K}\end{bmatrix}$.

Next, let Y^i denote the $(N\times 1)$ vector of outputs in each industry for country i, and let D^i denote the $(N\times 1)$ vector of demands of each good, so that $T^i=Y^i-D^i$ equals the vector of *net exports* for country i. The *factor content of trade* is then defined as $F^i\equiv AT^i$, which is an $(M\times 1)$ vector. We will denote individual components of this vector as F_k^i, where a positive value indicates that the factor is exported, while a negative value indicates that the factor is imported. For example, with just labor and capital, the factor content of trade is

$$\begin{pmatrix} F_\ell^i \\ F_k^i \end{pmatrix} \equiv AT^i.$$

The goal of the HOV model is to relate the factor content of trade AT^i to the underlying endowments of country i. To do so, we can proceed by computing AY^i and AD^i. The term AY^i equals the demand for factors in country i. Analogous to the full-employment conditions studied in chapter 1, AY^i equals the endowments of country i, which we write as $AY^i=V^i$. Turning to AD^i, this term is simplified by using our

[4]This matrix should include both the *direct* primary factors used in the production of each good, and the *indirect* primary factors used through the intermediate inputs. In practice, the indirect factors are measured using the input-output matrix for the economy. That is, denoting the $(M\times N)$ *direct* factor requirements by $\tilde{A}$ and the $(N\times N)$ input-output matrix by B, we compute the *total* factor requirements as $A=\tilde{A}(I-B)^{-1}$.

assumption of identical and homothetic tastes. Since product prices are equalized across countries by free trade, it follows that the consumption vectors of all countries must be *proportional* to each other. We shall write this as $D^i = s^i D^w$, where D^w denotes the *world* consumption vector and s^i is the share of country i in world consumption.[5] It follows that $AD^i = s^i AD^w$. Note that if trade is balanced, then s^i also equals country i's share of world GDP.[6] Since world consumption must equal world production, we therefore obtain $AD^i = s^i AD^w = s^i AY^w = s^i V^w$, where the last equality is the full-employment condition at the world level.

Making use of these expressions for AY^i and AD^i, we have therefore proved

$$F^i \equiv AT^i = V^i - s^i V^w, \tag{2.1}$$

which is a statement of the Heckscher-Ohlin-Vanek (HOV) theorem. In terms of individual factors, this is written as $F_k^i = V_k^i - s^i V_k^w$. If country i's endowment of factor k relative to the world endowment *exceeds* country i's share of world GDP ($V_k^i / V_k^w > s^i$), then we say that country i is *abundant* in that factor. In that case, (2.1) says that the factor-content of trade in factor k should also be positive ($F_k^i > 0$), and conversely if country i is scarce in factor k ($V_k^i / V_k^w < s^i$).

What does the HOV theorem tell us about the Leontief test? To anwer this, let us focus on just two elements of the factor-content vector, for labor and capital. These are written as

$$F_k^i = K^i - s^i K^w, \tag{2.2a}$$

$$F_\ell^i = L^i - s^i L^w, \tag{2.2b}$$

where F_k^i and F_ℓ^i are the computed factor contents of trade, and K^i and L^i are the capital and labor endowments for country i. Following Leamer (1980), we define capital to be abundant relative to labor in country i if $K^i/K^w > L^i/L^w$. Then using (2.2), the implications of capital abundance are:

THEOREM (LEAMER 1980)

If capital is abundant relative to labor in country i, then the HOV theorem (2.1) implies that the capital/labor ratio embodied in *production* for country i exceeds the capital/labor ratio embodied in *consumption:*

$$K^i/L^i > (K^i - F_k^i)/(L^i - F_\ell^i). \tag{2.3}$$

Proof:

From equation (2.2), we have $K^w = (K^i - F_k^i)/s^i$ and $L^w = (L^i - F_\ell^i)/s^i$. It follows that $K^i/K^w = s^i K^i/(K^i - F_k^i)$ and $L^i/L^w = s^i L^i/(L^i - F_\ell^i)$. Then $K^i/K^w > L^i/L^w$ implies that $K^i/(K^i - F_k^i) > L^i/(L^i - F_\ell^i)$, which is rewritten as (2.3). QED

[5] Letting p denote the vector of prices, then $p'D^i = s^i p'D^w$ so that $s^i = p'D^i / p'D^w$.

[6] Continuing from note 5, if trade is balanced so that expenditure = income in each country, then $p'D^i = p'Y^i$ and so $s^i = p'Y^i / p'Y^w = GDP^i / GDP^w$.

To interpet this result, note that K^i and L^i are simply the endowments of capital and labor, or alternatively, the capital and labor embodied in *production*. If we subtract the content of these factors embodied in trade, then what we end up with can be defined as the factor-content of *consumption,* or $K^i - F_k^i$ and $L^i - F_\ell^i$. Then equation (2.3) states that the capital/labor ratio embodied in *production* (on the left) must exceed the capital/labor ratio embodied in *consumption* (on the right).

The results from making this comparison for the United States in 1947 are shown in table 2.2. In the first column we list the capital and labor endowments for the United States, and in the second column we show the capital and labor embodied in consumption. Taking the ratio of these, it is indeed the case that the capital/labor ratio embodied in production exceeds that in consumption. This is the precise application of the HOV theorem, and it turns out to be satisfied for the United States in 1947, contrary to what Leontief concluded. Thus, there was no paradox after all!

It is useful to see the HOV theorem and Leamer's result in a diagram. In figure 2.3, the length of the horizontal axis is the world labor endowment $L^w = L^1 + L^2$, and the length of the vertical axis is the world capital endowment $K^w = K^1 + K^2$. The origin for country 1 is in the lower-left corner, and for country 2 is in the upper-right corner. Thus, any point in the world endowment box measures the endowments (L^i, K^i) of the two countries. Suppose that the endowments are at the point V^i where country 1 is capital-abundant, $K^1/L^1 > K^w/L^w > K^2/L^2$.

Under the assumptions of the HOV model, the consumption of each country D^i is proportional to world consumption D^w, which means that the factor content of consumption AD^i is proportional to $AD^w = V^w$. In other words, the factor content of consumption must lie along the *diagonal* in the world endowment box, as illustrated by point AD^i. Therefore, a line from point V^i to point AD^i measures the factor content of trade. In figure 2.3, country 1 exports F_k^1 of capital services and imports F_ℓ^1 of labor services. With balanced trade, the slope of the line between V^i and AD^i measures the ratio of factor prices.

The theorem of Leamer (1980) states that if country 1 is capital abundant, as illustrated, then the capital/labor ratio embodied in *production* must exceed the capital/labor ratio embodied in *consumption.* That is, since the consumption point AD^i must lie on the diagonal, it is necessarily to the right of and below the endowment point V^i. While this is graphically obvious, note that it does not depend in any way on whether trade is balanced or not. For example, if country 1 is running a trade surplus (with the value of production exceeding consumption), then we should move the consumption point AD^i *to the left* down the diagonal. This would have no effect whatsoever on the capital/labor *ratio* embodied in consumption as compared to the capital/labor *ratio* embodied in production. So Leamer's test of the HOV theorem in (2.3) is completely robust to having nonbalanced trade. Indeed, Leamer (1980) argues that this was the key problem with the way that Leontief did the original test: Leontief's method of testing

Table 2.2: Leamer's (1980) Reformulation of the Leontief Test

	Production	Consumption
Capital ($ billion)	$327	$305
Labor (person-years)	47 million	45 million
Capital/Labor ($/person)	$6,949	$6,737

Note: Each column shows the amount of capital or labor embodied in production or consumption in the United States, for 1947.

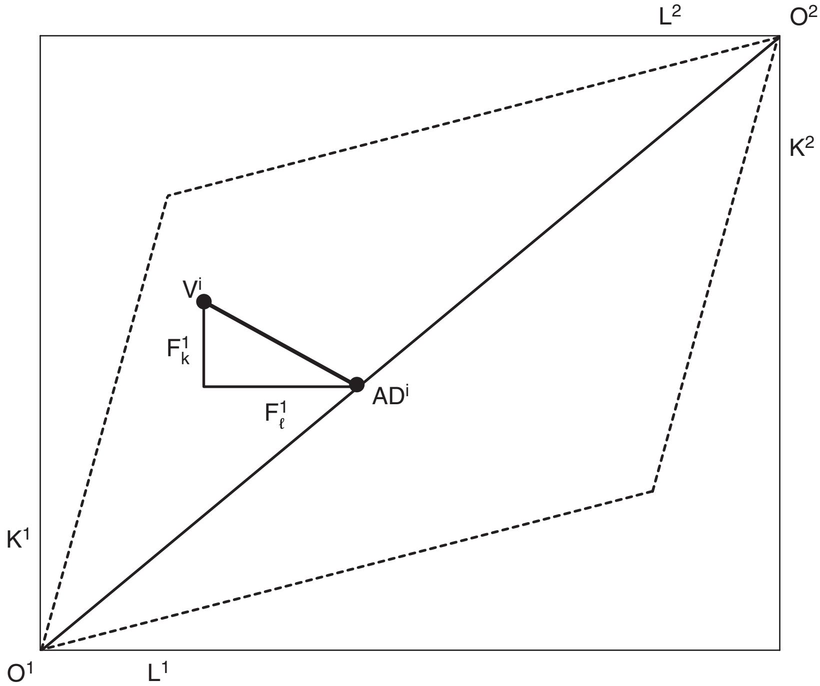

Figure 2.3

the HOV theorem was not valid with nonbalanced trade, and in 1947, the United States had a trade surplus and was exporting both labor and capital as embodied in trade.[7]

PARTIAL TESTS OF THE HOV THEOREM

The statement of the HOV theorem in (2.1) tells us immediately how a complete test of the theory should by performed: simply compute the left-hand side (using data on trade T^i and technology A), compute the right-hand side (using data on endowments V^i and V^w), and compare them. Depending on how well they match up, we can judge whether the theorem is an empirical success or not. This complete test requires both trade and endowments data for many countries, and technology data for at least one country. While such data are readily available today,[8] this was not the case two or three decades ago. Accordingly, many researchers, including Leontief himself, performed what we can call "partial tests" of the HOV model, using only two rather than all three types of data. Before considering the complete test, we will review several other partial tests that were performed.

[7] This situation is shown in figure 2.3 by moving the consumption AD^i so far to the left and down the diagonal that both capital services F_k^i and labor services F_ℓ^i are exported. Brecher and Choudhri (1982a) point out that the United States being a net exporter of labor is itself a paradox: from (2.2b), $F_\ell^i > 0$ if and only if $L^i > s^i L^w$, which implies $p'D^i L^i < p'D^w L^w$ using the expression for s^i from note 5. In other words, the United States will be exporting labor if and only if its per capita income is less than the world average, which is plainly false!

[8] Input-output tables as well as bilateral trade data for many countries are available from the Organization for Economic Cooperation and Development (OECD) STAN and I-O datasets, as well as from the World Input-Output Database, or WIOD (see Timmer et al. 2014).

Let us assume that the number of goods equals the number of factors, so that A is a square matrix, which we assume is invertible. Then we can rewrite (2.1) as

$$T^i = A^{-1}(V^i - s^i V^w). \tag{2.4}$$

This equation could be tested in several ways. First, if we think about the matrix A^{-1} as data, then we could run a regression of T^i on A^{-1}, and coefficients obtained would serve as an estimate of the relative abundance $(V^i - s^i V^w)$ of each factor. Baldwin (1971) performed a test similar to this, but instead of regressing T^i on A^{-1}, he actually regressed T^i on A'. With two factors and three goods, for example, A' is the matrix

$$\begin{bmatrix} a_{1L} & a_{1K} \\ a_{2L} & a_{2K} \\ a_{3L} & a_{3K} \end{bmatrix}.$$

Thus, Baldwin regressed the *adjusted net exports* of each industry[9] on its *labor and capital requirements* for one unit of production.

Using data for sixty U.S. industries for years around 1960, and disaggregating workers by their types, Baldwin (1971) obtained the following result:

$$\begin{gathered}
\text{Adjusted net exports} = -1.37^{*}\ \text{(physical capital/worker)} + 7011^{*}\ \text{(scientists/worker)} \\
-1473\ \text{(managers/worker)} + 71\ \text{(clerical staff/worker)} \\
+1578^{*}\ \text{(craftsmen \& foremen/worker)} + 248\ \text{(operatives/worker)} \\
+761\ \text{(unskilled employees/worker)} + 845^{*}\ \text{(farmers/worker)} \\
+\ \text{(other variables included for industry scale and unionization),} \\
N = 60,\ R^2 = 0.44,\ ^{*} = \text{significant at 95\% level}
\end{gathered}$$

Thus, looking across the U.S. industries, Baldwin finds that those industries using more scientists, craftsmen and foremen, or farmers relative to total workers will tend to have *higher* exports. The importance of scientists and farmers in predicting U.S. exports is not surprising at all, since the United States is abundant in skilled labor and land; and the importance of craftsmen and foremen is perhaps reasonable, too. What is surprising, however, is the *negative* coefficient found on the very first variable, physical capital/worker. Taken literally, this coefficient says that U.S. industries using *more* capital per worker will tend to export *less*. This is exactly the opposite of what we would expect if the United States were capital abundant. Thus, this result appears to be similar to the "paradox" found by Leontief.

Various writers after Baldwin have redone the type of regression shown above, with mixed results: sometimes the capital coefficient is positive, but other times it is again negative (see the survey by Deardorff 1984a). What are we to make of these results? Well, as we argued for Leontief's original paradox, it can be questioned whether Baldwin's approach is a *valid test* of the HOV model. From (2.4), a valid test would be to regress T^i on A^{-1}, but instead Baldwin regressed T^i on A'. To see the consequences of this, use the ordinary least squares (OLS) formula for the coefficients β that would be obtained from this regression:

$$T^i = A'\beta \Rightarrow \hat{\beta} = (AA')^{-1}AT^i = (AA')^{-1}(V^i - s^i V^w), \tag{2.5}$$

where the final equality follows from using the HOV theorem (2.1).

[9] Adjusted net exports are defined as industry exports per million dollars of total exports minus industry imports per million dollars of total imports.

Thus, we see that $\hat{\beta}$ is a *contaminated estimate* of the vector of relative factor endowments $(V^i - s^i V^w)$. Namely, rather than equaling the relative factor endowments vector, $\hat{\beta}$ equals the positive definite matrix $(AA')^{-1}$ times $(V^i - s^i V^w)$. Because $(AA')^{-1}$ is certainly not the identity matrix, it is entirely possible that the *sign pattern* of the elements in $\hat{\beta}$ will differ from the sign pattern of $(V^i - s^i V^w)$.[10] In other words, even if the United States was capital abundant and the HOV theorem held, it would still be possible for the Baldwin regression to find a negative coefficient on capital: this does not contradict the HOV theorem, because like the original Leontief approach, it is the wrong test.

The careful reader might point out, however, that possible differences in the sign pattern of $\hat{\beta}$ and $(V^i - s^i V^w)$ can presumably be *checked for* using the actual data on $(AA')^{-1}$. This is exactly the approach taken by Bowen and Sveikaukus (1992), who argue that such sign reversals are very unlikely to occur in practice. Thus, the Baldwin regression may be acceptable in practice, and in fact, it has been used in research by Romalis (2004).[11] As a descriptive tool to show how trade is related to industry factor requirements this regression makes good sense, but as a definitive *test* of the HOV theorem it is inadequate for the reasons we have described.

A second "partial test" of the HOV theorem in equation (2.4) has been undertaken by Leamer (1984). In contrast to Baldwin, Leamer tested (2.4) by treating factor endowments $(V^i - s^i V^w)$ as data, while estimating the elements of A^{-1}. To follow this test, notice that (2.4) applies across all countries i. Focusing on a single industry j, and letting the elements of A^{-1} be written as β_{jk}, we can write (2.4) in scalar form as

$$T_j^i = \sum_{k=1}^{M} \beta_{jk}(V_k^i - s^i V_k^w), \; i = 1, \ldots, C. \tag{2.6}$$

Notice that the summation in (2.6) is across factors, while the observations are across countries. Thus, to estimate this regression we would combine the observation for a *single* industry j across multiple countries $i = 1, \ldots, C$, where the coefficients β_{jk} are estimated. These coefficients should be interpreted as Rybczynski effects and, as shown in chapter 1, can be positive or negative.

To run this regression, we first need to choose some aggregation scheme for net exports and factors of production. Leamer (1984) works with the trade data for sixty countries in two years, 1958 and 1975. The trade data are organized according to the Standard Industrial Trade Classification (SITC). Because these are *traded* goods, they include only merchandise, that is, agriculture, mining, and manufacturing. Leamer organizes these goods into ten aggregates, as shown in the first column of table 2.3: two primary products (petroleum and raw materials); four crops (forest products, tropical agricultural products, animal products, and cereals); and four manufactured products (labor-intensive manufactures, capital-intensive manufactures, machinery, and chemicals).[12] Implicitly there is an eleventh product in the economy, making up all the nontraded goods, so Leamer includes total GNP as an aggregate to reflect this. There are also eleven factors of production, listed along the top row of table 2.3: capital,

[10] See Leamer and Bowen (1981) and Aw (1983).

[11] Romalis's model relies on the monopolistic competition framework, which we do not cover until chapter 5. It is recommended for further reading after that.

[12] The method used to form these aggregates involves looking at the cross-country correlations of the disaggregate products in each group with the aggregate itself. That is, "cereals" is an aggregate because countries that export this product in total tend to have high exports of all the products within this group. It turns out that the ratios of capital to worker and professional to all workers are also reasonably similar within the aggregates. Feenstra and Hanson (2000) consider how the aggregation scheme might lead to bias in calculations of the factor content of trade.