www
.press
.umich
.edu

MICHIGAN

Dynamics, Economic Growth, and International Trade

SIE *Studies in International Economics* includes works dealing with the theory; empirical analysis; political, economic, and legal aspects; and evaluation of international policies and institutions.

Keith E. Maskus, Peter M. Hooper, Edward E. Leamer, and
J. David Richardson, Editors.
Quiet Pioneering: Robert M. Stern and His International Economic Legacy

Bjarne S. Jensen and Kar-yiu Wong, Editors.
Dynamics, Economic Growth, and International Trade

Dynamics, Economic Growth, and International Trade

Edited by Bjarne S. Jensen and Kar-yiu Wong

Ann Arbor

THE UNIVERSITY OF MICHIGAN PRESS

Published in the United States of America by
The University of Michigan Press
Manufactured in the United States of America
♾ Printed on acid-free paper

2000 1999 1998 1997 4 3 2 1

*A CIP catalog record for this book is available
from the British Library.*

Library of Congress Cataloging-in-Publication Data
applied for
ISBN 0-472-10918-9

ISBN13 978-0-472-10918-0 (cloth)
ISBN13 978-0-472-02641-8 (electronic)

Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

Chapter 9

Chapter 10

Chapter 11

List of Tables

Contributors

Eric W. Bond, Pennsylvania State University, U.S.A.

Theo S. Eicher, University of Washington, U.S.A.

Rolf Färe, Southern Illinois University, U.S.A.

Oded Galor, Brown University, U.S.A.

Shawna Grosskopf, Southern Illinois University, U.S.A.

Bjarne S. Jensen, Copenhagen Business School, Denmark

Pantelis Kalaitzidakis, University of Cyprus, Cyprus

Shoukang Lin, Brown University, U.S.A.

Ngo Van Long, McGill University, Canada

Kazuo Nishimura, Kyoto University, Japan

Koji Shimomura, Kobe University, Japan

Kathleen Trask, Pennsylvania State University, U.S.A.

Stephen J. Turnovsky, University of Washington, U.S.A.

Pham Hoang Van, Cornell University, U.S.A.

Henry Y. Wan, Jr., Cornell University, U.S.A.

Chunyan Wang, Copenhagen Business School, Denmark

Kar-yiu Wong, University of Washington, U.S.A.

Part I

Introduction

CHAPTER 1

Introduction

Bjarne S. Jensen and Kar-yiu Wong

This volume presents the new contributions of a number of economists on dynamics, economic growth, and international trade. It includes one survey on endogenous growth and international trade and nine chapters that provide new analysis and new results about many important topics in this area.

The chapters were written amidst the growing interest of economists in the sources and effects of economic growth. The recent development of endogenous growth literature has introduced many new approaches to analyzing economic growth. The interest in endogenous growth is sparked by some observations about the growth rates of different countries, and it has led to important analyses that suggest various new ways of investigating theoretical and empirical aspects of economic growth.

It has been recognized that a substantial part of the recent growth literature focuses on closed and isolated economies, thus ignoring the common notion that trade is an engine of growth, and the fact that many countries showing impressive growth are open economies. Fortunately, this shortcoming of the literature is well understood, and many efforts have been made to analyze different issues related to the growth of open economies.

The contributions to this volume attempt to go beyond the present literature by focusing on economies that are linked to each other through movements of goods or factors of production. They examine the relationship between accumulation of factors, technological progress, efficiency, economic growth, international trade in goods, international factor movement, income distribution, and welfare. New approaches to analyzing these issues are suggested, and new results are obtained.

This book has three distinctive features:

1. A survey on endogenous growth and international trade gives the readers a critical review of recent developments in the literature of growth and trade. A unified model is developed to explain the main features of different models of endogenous growth and to show how they are related. Some results concerning the relationship between growth and trade are also explained. The possibility of convergence of growth rates of countries with or without international trade in goods or international factor mobility is discussed.

2. Some of the chapters in this volume analyze some "traditional" issues in a new context. For example, the possibility of diversification and sustained growth in the neoclassical framework with different rules of saving and with or without endogenous population/labor growth is investigated, and the patterns of trade with overlapping generations or human capital accumulation are derived.

3. Other chapters of the volume examine some newer issues, such as the relationship between the accumulation of different types of capital in growing economies, the interdependence between growth and international factor movement, and the dynamics of international factor movement. Furthermore, one of the chapters measures the changes in technological progress and efficiency of many countries, and another suggests a new theory of growth based on trade and technology transfer through learning by doing.

There are five parts in this volume. Part I contains this introduction (chapter 1) and a survey on endogenous growth and international trade (chapter 2). This survey, by Long and Wong, provides a systematic examination and presentation of major developments in the theory of endogenous growth and the relationship between economic growth and international trade. Those readers who find the present literature voluminous, confusing, and difficult to follow would find this survey helpful in sorting out different approaches to endogenizing economic growth of closed and open economies.

The survey is divided into two major sections. The first covers the theory of endogenous growth for closed economies. Using a unified framework, which reduces to several models of endogenous growth in special cases, it discusses several important factors of growth that have been proposed in the literature. It emphasizes the major features of each theory and shows how it is different from the neoclassical theory of growth. Growth due to human capital accumulation or R&D activities is thoroughly discussed.

While the majority of articles on endogenous growth focuses exclusively on closed economies, there has been a growing interest in the growth of open economies linked to each other through movement of goods, factors, and/or knowledge. The second section of the survey covers the major works in the literature on growth and trade. Some of the models are direct extensions of those for closed economies, but some are new. It is shown that trade has an important impact on growth, and with open economies, many new issues arise. However, many of the results are generally quite sensitive to the models used. The survey also covers recent work on economic growth and international factor mobility, with special attention to why international factor movement may or may not affect the convergence of the growth rates of economies.

Part II consists of three chapters that focus mainly on the dynamics of basic trade models. The chapter by Jensen and Wang (chapter 3) provides an extensive examination of two central issues concerning the general equilibrium dynamics of factor allocation in trading economies: diversification of production in a steady state and perpetual growth in the long run. These two issues are related to the traditional questions of whether a trading economy with the usual neoclassical settings can remain diversified and grow perpetually without relying on exogenous technological progress.

Jensen and Wang suggest a unified framework to analyze these two issues. This framework covers different rules of factor accumulation. In terms of physical capital accumulation, savings can be provided as a fixed proportion of national income or as a fixed proportion of capitalists' income (the classical assumption), or can be determined optimally according to the Ramsey rule. The accumulation of the labor force, on the other hand, can be exogenous or endogenous. Jensen and Wang obtain some new, interesting results. For example, they argue that if labor grows endogenously, then the perpetual growth of a small open economy is not possible with stationary technologies, irrespective of the savings rule. Diversification is possible with proportional saving, but is not possible under classical or Ramsey saving, unless the labor supply is endogenous. They also derive necessary and sufficient conditions for diversification in two large trading economies.

While most of the chapters on growth and trade consider models characterized by constant-returns technologies, the one by Long, Nishimura, and Shimomura (chapter 4) provides an interesting analysis of a growing, open economy under variable returns to scale. The economy has an infinite horizon, with optimal saving determined by a representative agent, and there are two production sectors, each of which has a technology that shows increasing returns for small scale of operations. The chapter examines the adjustment of such an economy under free trade with no external monopoly power, and determines whether growth can be sustained in the long run.

This model suggests a new theory to explain the phenomenon of a poverty trap. It shows that if the initial capital stock is below a threshold value, the capital stock in the economy will gradually run down to zero, as the interest rate approaches zero, but starting with an initial capital stock above the threshold level, the economy will grow perpetually in consumption per head, while the economy will sooner or later specialize in the production of the capital-intensive good. An important feature of this model is that, in a closed economy, capital will not run down to zero, because it does not depreciate and cannot be consumed, but for an open economy, capital can be traded for consumption, and depletion is possible. Thus this chapter points out the possibility of a negative impact of trade on growth.

While the dynamics of growth and trade are analyzed in the previous two chapters using infinite-horizon models, the chapter by Galor and Lin (chapter 5) considers instead a two-country model characterized by overlapping generations. In this model, the production side of each economy in each period is the traditional two-sector, two-factor framework with perfect competition and constant returns to scale, but with two important features to distinguish it from the infinite-horizon models. First, saving is done by workers. Second, contrary to the Uzawa condition of factor-intensity ranking of the sectors, it is assumed that the investment sector is capital intensive. The latter feature implies that a steady-state equilibrium, if it exists, is saddle-path stable, and it is the saddle-path stability of the steady state that makes the dynamic equilibrium of this model determinate.

Galor and Lin apply the model to derive several interesting results. For example, they show that the low time preference country exports the capital-intensive good in a free-trade steady state. The impacts of trade on factor prices are also derived. Furthermore, they argue that diversification in production is possible in both countries in a steady state, thus implying factor price equalization.

In part III, there are two chapters that analyze endogenously growing open economies with explicit consideration of accumulation of several types of capital and their effects on growth. The chapter by Turnovsky (chapter 6) analyzes the growth of a small, open economy that is due to the accumulation of two types of capital, public and private. The externality of public capital is the source of sustained growth. However, unlike many other endogenous growth articles that treat government expenditures as a flow, this chapter assumes explicitly that output depends on the stock, not the flow, of public capital (like its dependence on the stock of private capital). As this chapter argues, this approach is more appropriate to government expenditures for building public infrastructures. Furthermore, this chapter incorporates the adjustment cost and the congestion that public capital is subject to.

Turnovsky first examines a centrally planned small open economy, in which the government controls all quantities directly, and shows that government expenditures maximizing welfare may not coincide with those maximizing growth. The chapter then examines a decentralized economy and derives the optimal, time-varying taxes that will generate a transitional adjustment path and a balanced growth path identical to those of the centrally planned economy.

The chapter by Bond and Trask (chapter 7) analyzes trade and growth in the presence of endogenous accumulation of physical and human capital. As in the Uzawa-Lucas model, the human capital stock in an economy can grow perpetually through education and knowledge spillover, and such growth supports perpetual accumulation of physical capital.

Bond and Trask then consider a model of a small open economy, in which both investment and consumption goods are tradable, but human capital is not. They show the existence, uniqueness, and saddle path stability of a balanced growth path (BGP) for the small open economy, given constant world prices for the traded goods. They show that there is a unique world price at which all three goods are produced: for any other world prices the economy is specialized in one of the traded goods and the non-traded good. They examine the relationship between world prices and growth rates, factor endowments, and welfare on the balanced growth path, and show how technical progress affects growth rates and the patterns of trade.

Part IV contains two chapters on assessing and theorizing technological progress, with special emphasis on the Asian Pacific countries and the newly industrialized economies. The chapter by Färe and Grosskopf (chapter 8) estimates the performance of the countries in APEC (Asian-Pacific Economic Community) in terms of productivity, efficiency, and technological progress for the years 1975–1990. Their work consists of three sections. In the first section, the output-oriented Farrell measure of technical efficiency is computed for each of the countries in each year of the period, and these measures are then compared with the per capita income levels of the countries. Two findings were obtained: a negative correlation between efficiency and income level, and a greater dispersion of efficiency for poor countries than for rich countries. The first finding supports the hypothesis that some countries are persistently poor because of inefficiency, but the second finding apparently is not explained by any existing theory.

Färe and Grosskopf then turn from the static model to two dynamic ones. In the first one, they measure the efficiency change, technical change, and the Malmquist productivity index, which is the product of the efficiency change and the technical change, for each of the countries for each pair of consecutive years. However, they do not find any obvious signs of convergence: all three indexes do not show any strong correlation with the per capita income levels of the countries, although the dispersion of each index is greater for poor countries than for rich countries. When they compute the values of the indexes over the whole period, some countries did perform impressively; for example, in terms of productivity change, Hong Kong, Singapore, Korea, and Canada all did very well. Färe and Grosskopf then consider a dynamic activity analysis model in which the efficiency levels of countries are estimated, which can be interpreted as the potential efficiency for the countries, under the condition that investment is chosen optimally. They observe a result similar to what they got in the static mode, i.e., the efficiency levels obtained are negatively correlated with per capita income.

While Färe and Grosskopf focus on measuring the productivity and efficiency of these Pacific countries, Van and Wan (chapter 9) provide a theory of the relationship between technological progress, capital accumulation, economic growth and international trade, drawing upon the growth experience of the newly industrialized economies (NIEs: Hong Kong, Singapore, Taiwan and Korea). According to this theory, not only is technological progress an important component of the growth of these economies, which is supported by the findings of Färe and Grosskopf, but growth also depends crucially on the learning experience the economies obtain through international trade.

The theory of Van and Wan begins with three essential components of growth: capital accumulation, gain in technology, and international trade, which are assumed to possess generalized complementarity. Technology in a developing country, in the form of a surrogate production function, can be improved through learning and emulation. International trade provides an important channel owing to contagion effect. As a developing country learns to produce more sophisticated products, it gains technological competitiveness, and learning can go on without bounds. Accompanying the technological gain is physical capital accumulation. However, instead of using factor accumulation to explain growth of these economies, as some literature suggests, Van and Wan argue that it is the result of technological progress.

In part V there are two chapters on the dynamics of international factor mobility. They are interesting not only because they provide analysis of some crucial issues, but also because the literature has so far paid very little attention to the important phenomenon of international factor mobility in a dynamic context.

The chapter by Wong (chapter 10) is an attempt to address two drawbacks in the literature of international labor migration. First, most, if not all, work on the dynamics of international labor mobility is based on the neoclassical framework of growth in the sense that in steady state, an economy has zero or exogenous growth. Second, literature on international labor migration usually focuses on one type of migration at a time, assuming implicitly or explicitly that other types of migration do not exist.

Wong suggests a model in which the growth of an economy is dependent endogenously on the rate of human capital accumulation through education, and in which individuals are allowed to emigrate in the form of either permanent migration, temporary migration, or brain drain. Thus, in his model, both the type and amount of migration are determined endogenously in a dynamic context. Wong examines how each type of migration may affect growth, income distribution, and education, and how growth may affect the type of migration individuals choose over time. He also examines how workers would choose endogenously the type of migration, i.e., whether and when to migrate, and whether and when to return. It is argued that a growing emigration economy may go through different stages with different types of international labor emigration.

The chapter by Eicher and Kalaitzidakis (chapter 11) focuses on foreign direct investment (FDI). Noting that FDI is more likely to occur among countries with abundant human capital and advanced technologies, they suggest that a multinational corporation, when investing in a host country, must train workers to work with firm-specific technologies. The cost of training workers depends, among other things, on the workers' abilities, which are known to the workers but not to the firms. There is asymmetric information, in the sense that the firm does not observe an individual worker's ability, although the average ability of all workers in a country is a public information.

After constructing a model of adverse selection and efficiency wages, Eicher and Kalaitzidakis examine several issues related to trade and foreign direct investment. In terms of trade, they show that trade between two countries leads to wage convergence and relocation of workers in each country to the sector of comparative advantage, in terms of production of informational efficiency. This impact of trade is what they call informational efficiency gains from trade. In terms of FDI, firms from a developed country (DC) with its more advanced technology invest in the less developed country (LDC), paying a wage higher than that paid by local firms, but lower than what they pay in the DC. The role of asymmetric information in the investing firm's decision and the dynamics of endogenous technological change are explicitly derived.

All of the chapters in this volume were presented and discussed at a conference on "Dynamics, Economic Growth, and International Trade", which was held in Helsingør, Denmark, from August 15 to 17, 1996.

CHAPTER 2

Endogenous Growth and International Trade: A Survey

Ngo Van Long and Kar-yiu Wong

1. Introduction

Economists have long been interested in searching for the causes and effects of the growth of income and wealth of countries. Some earlier attempts to analyze economic growth with rigorous models appeared in the twenties and thirties, mainly characterized by the work of Ramsey (1928), Harrod (1939) and Domar (1946). While Ramsey is concerned about the maximization of intertemporal utility, Harrod and Domar concentrated on the equilibrium path of an economy. The work of Harrod and Domar is followed by growing interest in the theory of growth and a series of relevant papers, mainly in the fifties. One of the more influential contributions is due to Solow (1956) and Swan (1956). By introducing production substitution possibilities that exist in neoclassical production functions, Solow and Swan show that the equilibrium paths in the Harrod-Domar model could be more stable than otherwise suggested.

The work of Solow and Swan has been extended in many directions (for example, an increase in the dimension of the model and the introduction of new factors that may affect growth), and has been applied in different economic fields. In the field of international trade, different versions of the neoclassical model have been used to examine a wide range of issues; see Findlay (1984), Smith (1984) for in-depth surveys.

The interest of economists in economic growth was rekindled recently. First, there is the paper by Romer (1986) which heavily criticizes the neoclassical theory. Romer also suggests a model that endogenizes the growth rate of economies. Lucas (1988) provides some alternative, more appealing, ways to remedy a few of the shortcomings of the neoclassical theory. Since then, there has been a flood of papers and books on endogenous growth in the literature. For example, more than 50 papers published in economic journals (not including working papers) and several books can be found, all between 1990 and early 1996, to have contributed, in one way or another, to the endogenous growth and international trade literature.

The purpose of the present survey is to introduce the major contributions of this literature, focusing on what this literature has emphasized, what new ideas have been suggested, and the main features of some of the models. It is hoped that this survey, using some unified frameworks, can give the reader a brief summary and introduction to this growing literature. However, because the literature has become so voluminous, some issues and results have not been covered in this survey, and the reader is encouraged to read the original articles.

Section 2 introduces a unified growth model, which reduces to the neoclassical growth model and many of the endogenous growth models in some special cases. The unified model thus brings out the fact that the neoclassical growth model and many endogenous growth models are mathematically similar. Section 3 uses a reduced form of the unified model to examine the basic features of the neoclassical growth theory. This will help the reader understand the recent criticism of this theory. Section 4 explains the basic mechanics of the endogenous growth theory. Different models and how they endogenize the growth rates of economies are explained and compared. It will be pointed out that some of the ideas that have been used and developed in several endogenous growth papers can be traced back to several papers in the sixties and seventies. In particular, we found some "old" papers in the sixties and seventies that have already developed endogenous growth models. In section 5, we focus on two types of technological progress: horizontal innovation and vertical innovation. Section 6 introduces some of the papers on endogenous growth and international trade, while section 7 focuses on some recent work on growth and international factor mobility. Section 8 provides some concluding remarks.

2. A Unified Growth Model

Consider a closed economy, in which many competitive firms produce a homogeneous good using two factors, physical capital and labor. The good can be used for either consumption or production. The aggregate technology at time t, $t \in [0, \infty]$, can be represented by the following production function:

$$Y = AF(K, L), \tag{1}$$

where Y is the output, A a technology index, K the physical capital input, and L the labor input. Unless confusion arises, the time subindex is dropped for simplicity. We assume that the production function in (1) satisfies all neoclassical assumptions: increasing, linearly homogeneous, and concave in (K, L).

The labor input depends on two factors: the average human capital level, h, and the number of workers, M. For simplicity, we assume that the labor force is equal to $L = hM$. This formulation implies that a worker with two units of human capital is as productive as two workers, each with one unit of human capital, working together. Using this formulation, the production function reduces to

$$Y = AF(K, hM). \tag{1'}$$

Linear homogeneity of the production function (1′) implies that the per capita output, $y \equiv Y/M$, is equal to

$$y = Ahf(k), \tag{2}$$

where $f(k) \equiv F(k, 1)$, and $k \equiv K/L$ is the capital-labor ratio.

Perfect competition and cost minimization mean that factors are paid their marginal products, and perfect price flexibility implies that factors are fully employed.

The accumulation of physical capital in the economy comes from saving, or the gap between output and consumption. Denote the saving rate as a fraction of output by $s \in (0, 1)$. In equilibrium, saving is equal to investment, I. Thus $s = I/Y$. There are two common ways of determining the optimal saving: (a) It is chosen optimally by some or all individuals in a decentralized economy; and (b) It is chosen optimally by a social planner to maximize the per capita consumption under the Golden Rule, or to maximize the intertemporal utility of a representative consumer, as in the Ramsey model. In the present context, it suffices to treat s as a parameter.

Saving is converted into investment, meaning that the change of the capital stock over time is

$$\dot{K} = I - \delta K = sY - \delta K, \tag{3}$$

where a "dot" above a variable denotes the change of the variable with respect to time, and where $\delta > 0$ is the (exogenously given and stationary) depreciation rate. Suppose that the population grows at an exogenously given rate of n. Equation (3) then gives the rate of growth of the capital-labor ratio:

$$\widehat{k} = \frac{sy}{hk} - \delta - \widehat{h} - n, \tag{3'}$$

where a "hat" denotes the proportional rate of change of a variable; for example, $\widehat{k} \equiv \dot{k}/k$.

The growth of the economy is usually expressed in terms of the growth rate of its per capita output/income, which is obtained from (2):

$$\widehat{y} = \widehat{A} + \widehat{h} + \varepsilon_f \widehat{k}, \tag{4}$$

where ε_f is the elasticity of function $f(k)$. Condition (4) states that the growth of the economy depends on how technology, human capital, and physical capital grow over time. The growth rate of physical capital is given by (3′).

We now explain how the above model may reduce to the neoclassical model and some endogenous growth models.

3. Features of the Neoclassical Theory of Growth

Before turning to the endogenous growth literature, we first explain the features of the neoclassical model attributable mainly to Solow (1956) and Swan (1956). Suppose that in the economy under consideration the human capital stock is constant over time, meaning that we can write $h = 1$ and $L = M$. Furthermore, the production function is of the Cobb-Douglas type:

$$Y = AK^{\alpha}L^{1-\alpha}, \quad 0 < \alpha < 1, \tag{5}$$

which means that the per capita output is given by

$$y = Ak^{\alpha}. \tag{5′}$$

Therefore the growth rate of per capita output reduces to

$$\widehat{y} = \widehat{A} + \alpha\widehat{k}, \tag{6}$$

where in the present case $\varepsilon_f = \alpha$.

Very often, attention is paid to the steady state or long-run equilibrium of the economy. In the present model, the marginal product of capital is equal to $\alpha Ak^{\alpha-1}$ and its average product, y/k, is equal to $Ak^{\alpha-1}$. As a result, for a given A, the steady-state value of k is bounded from above and below: As $k \to \infty$, $y/k \to 0$, and by (3′), $\widehat{k} < 0$, and as $k \to 0$, $y/k \to \infty$, implying that $\widehat{k} > 0$. From (3′), we can obtain a value of k at which $\widehat{k} = 0$. This is used to define the steady state of the economy, i.e.,

$$sy/k = \delta + n. \tag{7}$$

In the steady state, factor prices are stationary. This equilibrium condition is illustrated in figure 1. The schedule representing sy/k is strictly downward sloping, and the steady state occurs when this schedule has a value of $n + \delta$, giving a steady-state capital-labor ratio equal to $\widetilde{k}$. It is clear from the above analysis and this diagram that, in the present model, the steady state exists and is unique. This result can also be obtained using any linearly homogeneous production function with the following Inada conditions: $f'(k) > 0$, $f''(k) < 0$, $f'(0) = \infty$, and $f'(\infty) = 0$.

As a result, the growth of the economy's per capita output is zero in the long run, unless there is growth in technological knowledge. Since the neoclassical model does not have an explicit theory of technology progress, the latter is either assumed to be given exogenously or treated as zero. Therefore, in a steady state an economy in the neoclassical model either does not grow or grows according to the exogenously given technology progress.

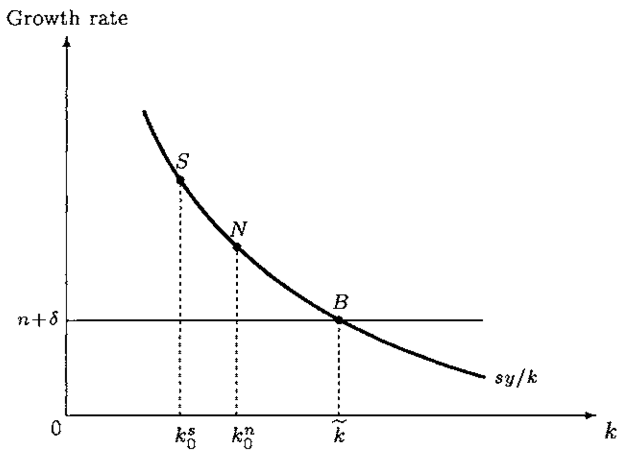


Fig. 1. Growth in the neoclassical model

The above model, however, has been under criticism recently. We now present some of the more common criticism, and explain how these have been used to motivate the endogenous growth theory.

(A) Exogeneity of the Growth Rate – As explained earlier, because k is constant in a steady state, the per capita output must grow according to the exogenous growth rate of technology. Exogeneity of the growth rate in the present model has several important implications. First, in a steady state, the per capita output of an economy with fixed technology

will not grow. Second, by condition (7), the steady-state output-capital ratio is positively related to the saving rate. This means that policies such as a saving subsidy will eventually raise the steady-state capital-labor ratio. However, policies that change the saving rate will not change the steady-state growth rate of the economy. In other words, saving has only *level* effect but no *growth* effect (Pitchford, 1960; Lucas, 1988).[1]

This result is contradictory to the usual notion that an economy can grow faster if it saves more. Similarly, a once-and-for-all change in technology or population will not have any long-run growth effect. Furthermore, government policies that do not affect the growth rate of technology or that of population will not change the steady-state growth rate of the economy. For example, trade liberalization will not have any growth effect, as long as it does not affect the growth rate of technology (Lucas, 1988).

(B) Disparities in International Growth Rates – The above model suggests that any two countries that have the same steady-state (or long-run) growth rate of technology should have the same steady-state growth rate of their per capita income, irrespective to their prevailing size or technology level. What we see instead is that countries consistently have wide disparities of growth rates. (Azariadis and Drazen, 1990). Furthermore, for some countries, this simple model is not supported by data. For example, by using (5′) and (3), (6) reduces to

$$\widehat{y} = \widehat{A} + \alpha\big[sA^{1/\alpha}y^{-(1-\alpha)/\alpha} - \delta - n\big]. \tag{6′}$$

Romer (1994) notes that between 1960 and 1985, the United States and the Philippines had about the same growth rate of per capita income, but in 1960, the Philippines' per capita income was only about ten percent of that of the United States. If the parameter α, which is the capital share, is taken to be 0.4 for both countries, equation (6′) suggests that the saving rates for the United States over this period should be 30 times larger than those of the Philippines in order to produce the same growth rate, assuming that both countries had the same technologies, population growth rate, and depreciation rate of capital, but this condition does not seem to be supported by evidence. It is possible, and likely, that these two countries have different technologies, population growth,

1. The level effect of an increase in the saving rate can also be shown in figure 1. An increase in the saving rate shifts up the schedule for sy/k, leading to an increase in the steady-state capital-labor ratio. The steady-state growth rate of per capita output remains to be the same as that of technological progress. However, it should be noted that an increase in saving could have positive effects on growth during the transitional period.

and depreciation rates, but it is not clear whether these differences are large enough to explain why these two countries grow so differently.

(C) Convergence of Growth Rates – The neoclassical growth theory has important implications for the changes in the growth rates of different countries. In terms of figure 1, the gap between the sy/k schedule and the horizontal line corresponding to $n + \delta$ gives the speed of adjustment of the capital-labor ratio (or the per capita output) for a closed economy. Since the sy/k schedule is strictly downward sloping for an economy with a Cobb-Douglas production function, the speed of adjustment of k or y decreases monotonically, while k moves toward the steady-state point. Suppose that we have two countries, North and South, which are identical except that North has a higher initial capital-labor ratio, $k_0^n > k_0^s$. Assuming that they are below the steady-state level, both k_0^n and k_0^s are increasing over time, but that of the South will grow faster because its capital-labor ratio gives a bigger gap between sy/k and $n + \delta$. So the South is catching up until both countries have the same capital-labor ratio and the same growth rate.

Some casual observations could easily suggest that many countries do not show convergence of the growth rates. In fact, there are many countries that show persistent growth rates higher than others, and for many developing countries, there is no sign of catching up with the growth rates of developed countries. Both Romer (1986) and Lucas (1988) cite the lack of convergence of the growth rates of different countries as a sign of the inadequacy of the neoclassical growth theory in explaining the growth experience of countries.

Recently, the views of Romer and Lucas have been challenged, and alternative interpretations of the convergence hypothesis have been suggested. We will have some more discussion of this issue in the last section.

4. The Basic Mechanics of Endogenous Growth[2]

We now present some of the more popular models that endogenize the growth rate of an economy. We will pay particular attention to how these models attempt to address the above criticism on the neoclassical models.

2. The term "mechanics" is borrowed from Lucas (1988).

4.1 Growth with Non-Essential Labor: The Solow-Pitchford AK Model

We have seen that in the neoclassical framework with a Cobb-Douglas production function, because the marginal product of capital approaches zero as the capital-labor ratio becomes infinite; in a steady state, the capital-labor ratio of the economy must be bounded from above. This implies that in the absence of technological progress, the per capita output must also be bounded from above. Solow was aware of this limitation and did give alternative examples of an economy that is "so productive and saves so much that perpetual full employment will increase the capital-labor ratio (and also output per head) beyond all limits." (Solow, 1956, pp. 72 and 77.) On p. 77, he even suggested an example of a CES production function that can give perpetual growth of an economy. Pitchford (1960) was probably the first one to suggest a rigorous, general theory of endogenous growth using CES production functions, and to show how "in some circumstances a rise in the saving ratio can achieve a permanently higher rate of growth of income" (p. 499).

In the Solow-Pitchford model, human capital is also assumed to be constant so that we can write $h = 1$ and $L = M$. Its main feature is that it abandons the Cobb-Douglas production function used in the neoclassical model, and assumes instead a CES production function:

$$Y^\alpha = (aK^\alpha + bL^\alpha), \tag{8}$$

where $a, b > 0$. For perpetual growth, we assume that $0 < \alpha < 1$, which implies that the elasticity of substitution is greater than unity. The reason is given below. The marginal products of capital and labor are, respectively,

$$r = a\big[a + bk^{-\alpha}\big]^{(1-\alpha)/\alpha}, \tag{9.1}$$

$$w = b\big[ak^\alpha + b\big]^{(1-\alpha)/\alpha}. \tag{9.2}$$

By (9.1), if the capital-labor ratio is finite, physical capital accumulation still shows decreasing returns. However, because $0 < \alpha < 1$ if k approaches infinity, the rental rate approaches its lower bound, $a^{1/\alpha}$, while the wage rate approaches infinity. If at this point the saving of the economy is high enough, the economy can experience perpetual growth in terms of its per capita output.

To see this point more rigorously, note that because $0 < \alpha < 1$,

$$\lim_{k\to\infty} \frac{Y}{K} = A, \tag{8'}$$

where $A = a^{1/\alpha}$. In other words, the production function is asymptotically linear. Equation (8′) represents the famous AK model. Assuming no technological progress, the growth rate of output is

$$\widehat{Y} = \widehat{K}, \tag{10}$$

or that of the per capita output is

$$\widehat{y} = \widehat{K} - n. \tag{10′}$$

Making use of the investment equation (3′) and (8′), the growth rate of the economy's per capita income is

$$\widehat{y} = sA - \delta - n. \tag{10″}$$

Equation (10″) has two important implications. First, if the saving rate is high enough, and if A is big enough, the economy can have a positive sustained growth. Second, the growth rate of the economy depends on variables such as saving rate, technology *level*, and population growth rate. Therefore, any government policies that affect these variables will have a growth effect.

The adjustment and steady state of the economy can be illustrated in figure 2. Assuming a constant saving rate, the schedule sy/k is downward sloping, but is bounded from below by the line sA. If saving is sufficiently high, the line sA is above the line $n + \delta$. Equation (3′) implies that the gap between the schedule sy/k and the line $n + \delta$ gives the rate of adjustment of k. Since the schedule sy/k is downward sloping, the rate of adjustment declines over time. Asymptotically, sy/k is equal to sA, and the perpetual growth rate of the economy is equal to $sA - \delta - n$.

This model implies (conditional) convergence, as does the neoclassical model. For example, assume that there are two economies, North and South, which are identical except that the initial capital-labor ratio is higher in the North than in the South, $k_0^n > k_0^s$. In both countries, the capital-labor ratios are increasing over time, but the South has a higher growth rate and is catching up.

When the aggregate production function is of the CES type, as in the Solow-Pitchford model, both labor and capital are non-essential. It is obvious from the above analysis that it is the non-essentiality of labor that is necessary for perpetual growth with physical capital accumulation. Jensen and Wang (1997) provide an elaboration of this point. The required condition for persistent per capita growth is the violation of the Inada condition (the marginal product of capital is bounded below). This is demonstrated in Jensen and Larsen (1987). They also show that

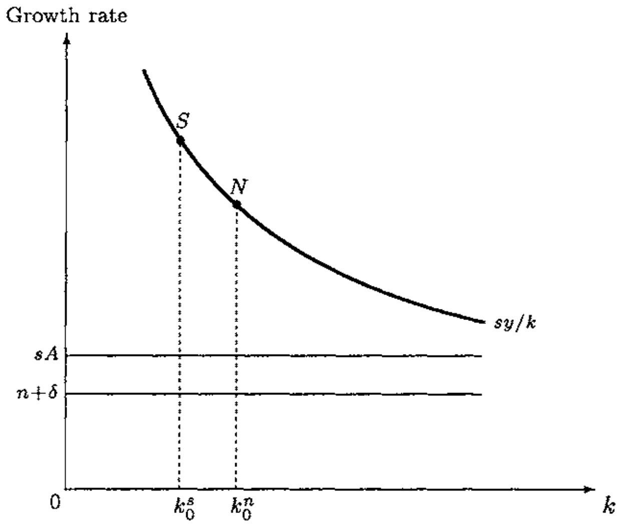


Fig. 2. Growth in the Solow-Pitchford model

depending on the properties of the marginal products of factors, various types of endogeneous growth are possible.

In the above analysis, saving is assumed to be a fixed proportion of the aggregate output. However, if part of the saving in the economy comes from wage earnings, sustaining the growth of the economy may be difficult in the Solow-Pitchford model because the share of labor income approaches zero when the capital-labor ratio approaches infinity. To avoid this problem, Saint-Paul (1992) suggests the use of taxation to redistribute income within the economy.

Other extensions of the Solow-Pitchford endogenous growth model are provided in some recent papers, including Jones and Manuelli (1990) and Rebelo (1991), where a generalization to a two-sector economy is shown to be possible.[3] Even if labor is essential in the production of the consumption good, as long as the investment good can be produced without labor, it is possible to have per capita consumption growing at a constant positive rate forever. To see this, denote the aggregate capital stock by K and the fraction of K working in the consumption sector by ϕ, meaning that $(1-\phi)K$ is the amount of capital working in the

3. For a more general formulation of endogenous growth models with many capital goods, see Dolmas (1996).

investment sector.

Consider the following production functions:

$$Y_c = A(\phi K)^{\alpha} L^{(1-\alpha)}, \tag{11.1}$$

$$Y_i = \beta(1-\phi)K, \tag{11.2}$$

where Y_c and Y_i are the outputs of the consumption sector and investment sector, respectively, and $\beta > 0$ is a parameter. Both sectors are competitive. Using lower-case letters to denote per capita output, $y_c = Y_c/L$ and $y_i = Y_i/L$. Assuming no technological progress, A is a constant. In a steady state, which we are focusing on right now, ϕ is a constant. Using (11.1)–(11.2), we have

$$\widehat{y}_c = \alpha\widehat{k}, \tag{12.1}$$

$$\widehat{y}_i = \widehat{k}. \tag{12.2}$$

Denote the price of the investment good (capital) relative to the consumption good by p. Perfect mobility of capital between the sectors equalizes the rental rates of the two sectors:

$$p\beta = \alpha A k_c^{\alpha-1},$$

where $k_c = \phi K/L$, or, in terms of growth rate,

$$\widehat{p} = -(1-\alpha)\widehat{k}. \tag{13}$$

Condition (13) implies that if the capital-labor ratio is rising, the relative price of capital is falling. Combining conditions (12.2) and (13), we see that py_i is growing at a rate of $\alpha\widehat{k}$, the same as that of y_c.

The per capita national income is equal to $I/L = y_c + py_i$, which grows at a rate of $\alpha\widehat{k}$. The growth rates of capital and the capital-labor ratio depend on saving, which may be chosen by the government in a social planner's problem or by individuals in a decentralized economy. Let s be the ratio of saving to national income. The capital-labor ratio then grows according to

$$\widehat{k} = \frac{sI}{pK} - \delta - n. \tag{14}$$

With enough saving, the economy grows over time, i.e., $\widehat{k} > 0$. Equation (14) thus implies that disparities in growth rates among countries can be explained in terms of the saving rates of the countries. In particular, a country grows faster if it saves more.

4.2 Growth with Knowledge Spillovers and Increasing Returns: The Romer AK Model

Another way of endogenizing growth rates was suggested by Romer (1986). His model has three main features. First, knowledge is used by firms as a capital good. Second, knowledge can be augmented so that we can talk about aggregate knowledge in the economy. Third, firms are competitive, taking prices and the aggregate knowledge as given. Knowledge and other factors are chosen optimally by firms, and knowledge accumulates by sacrificing current consumption. Thus with knowledge as a factor, Romer's model is subject to factor-generated Marshallian externality. Because he did not consider human capital, we can write $h = 1$.

Consider a representative firm which chooses a knowledge input of K_i. Let the aggregate knowledge be $K \equiv \sum_i K_i$. Since the firm takes the technology and the aggregate knowledge as given, in the production function given by (5), we replace A by AK^β, $0 < \beta < 1$. After substitution, the production function reduces to

$$Y_i = AK^\beta K_i^\alpha L_i^{1-\alpha}. \tag{15}$$

This means that the firm treats K, which it can hardly control, and technology as exogenously given, and then chooses K_i and L_i optimally. All firms have the same production function. Adding up the firms' production gives the aggregate production function

$$Y = AK^{\alpha+\beta} L^{(1-\alpha)}. \tag{15$'$}$$

Supposing that $\beta = 1 - \alpha$, and setting L at unity, the production function in (15$'$) reduces to the AK model asymptotically. Alternatively, the production function can be expressed in terms of the per capita output:

$$y = AK^\beta k^\alpha,$$

which gives the growth rate of per capita output:

$$\widehat{y} = \widehat{A} + \beta\widehat{K} + \alpha\widehat{k}. \tag{16}$$

Equation (16) shows the sources of perpetual growth. Assuming a Cobb-Douglas production function with no technological progress, the capital-labor ratio, k, remains constant in a steady state, and the growth of per capita output is proportional to that of knowledge capital.

Romer (1986) actually assumes that knowledge capital displays strictly increasing marginal product, i.e., $\beta > 1 - \alpha$. A problem of this model is that for any constant saving rate $s > 0$, the stock of knowledge capital will become infinite after some finite time if $\dot{K} = sY$. To avoid this problem, Romer assumes a bounded investment technology, $\dot{K} = g(sY/K)$ where $g(.)$ is strictly less than some upper bound. Xie (1991) provides an explicit example of this type, and suggests an alternative formulation:

$$Y_i = K_i^{\alpha}(KL)^{1-\alpha}B(K), \tag{17}$$

where $B(K)$ is positive, increasing, and bounded above by unity. Clearly, this has the same structure as the AK model discussed earlier. Xie shows that for some appropriate initial condition, the growth rate of K will be monotonically increasing and approach an upper bound. For more discussion of possible explosiveness of this type of models, see Solow (1994).

In Romer (1986), K is the stock of public good that enters each firm's production function. Obviously, a flow or a stock of public good produced by the government can also generate perpetual growth. See Barro (1990) and Turnovsky (1997).

4.3 Growth with Education: The Uzawa-Lucas Model

We now turn to endogenous growth models which explicitly examine the accumulation of human capital. There are two main channels through which individuals acquire human capital: education and learning by doing. In this subsection, we focus on education.

Earlier efforts that analyze human capital in a dynamic model generally have a limited success in explaining perpetual growth.[4] To allow for a perpetual growth, Uzawa (1965) proposes to treat the skill level of workers as a variable which can increase over time.[5] Lucas (1988) extends this idea and allows for external effects of human capital. We now present a simple version of their models.

At any time, let h be the average human capital level, which is a general knowledge available to everyone. Individuals possessing this general knowledge can acquire more by receiving education. Each individual is endowed with one unit of nonleisure time. A fraction of this time, denoted by τ, is spent on receiving education, and the rest, $1-\tau$, on work. The increase in human capital depends positively on the amount of time

4. See, for example, Razin (1972a, 1972b), Manning (1975, 1976), Hu (1976), and Findlay and Kierzkowski (1983).

5. While so much attention has been paid to the growth factors on the production side of economies, Uzawa, in a less known paper (Uzawa, 1969), suggested a growth model that endogenizes the rate of time preferences.

spent on education and the prevailing human capital level. For simplicity, it is assumed that there is no depreciation of human capital. Human capital is postulated to increase according to the following function:

$$\dot{h} = hg(\tau), \tag{18}$$

where $g'(\tau) > 0$. Note that h is both the average human capital stock and the human capital stock each individual acquires through education in the next period. Individuals, taking the existing human and physical capital as given, choose τ to maximize their utility subject to the budget constraint and condition (18). After an individual has accumulated human capital, the new level of knowledge is immediately available to all individuals.[6]

In the presence of human capital, the available efficiency units of labor is $L = (1 - \tau)hM$. The production function (5′) reduces to

$$y = (1 - \tau)hAk^{\alpha},$$

and the growth rate of the per capita output is given by

$$\widehat{y} = \widehat{(1 - \tau)} + \widehat{h} + \widehat{A} + \alpha\widehat{k}. \tag{19}$$

Therefore the growth of per capita output depends on that of τ, h, A, and k.

As explained before, because of diminishing returns of capital, if $\widehat{A} = 0$ then in a steady state (balanced growth path) the capital-labor ratio remains constant, i.e., $\widehat{k} = 0$. Furthermore, τ must remain constant in such a path, i.e., $\widehat{\tau} = 0$. This implies that the growth of human capital in a steady state is equal to

$$\widehat{h} = g(\widetilde{\tau}),$$

where $\widetilde{\tau}$ is the steady-state value of τ. Substitute these growth rates into (19) to give $\widehat{y} = \widehat{h} = g(\widetilde{\tau})$. In other words, the per capita output and human capital grow at the same rate, and this rate depends on the steady-state value $\widetilde{\tau}$, which is chosen endogenously by individuals. See Caballé and Santos (1993) for a rigorous discussion of other properties of the model such as existence of a steady state and dynamics.

In this model, the growth of an economy depends crucially on $\widetilde{\tau}$: Any policy or economic factor that affects $\widetilde{\tau}$ can thus change the economy's long-run growth. This model is much richer than the neoclassical one for explaining the international differences in growth rates. Thus, two

6. Note that the free-rider problem may exist in this type of models.

countries which have the same technology may still grow at different rates in steady states if individuals in different countries choose to spend different amounts of time on education (Azariadis and Drazen, 1990), or if they have different education policies. In particular, two countries may have two different steady states, and in general there is no reason to believe that their growth rates should converge.

Stokey (1991) extends the Uzawa-Lucas model and considers a model with a continuum of individuals with different human capital and a continuum of products with different qualities. Firms are competitive and hire individuals with higher levels of human capital to produce higher quality products. She shows explicitly how human capital accumulation depends negatively on the rate of time preference but positively on the elasticity of intertemporal substitution.

Grossman and Helpman (1991a, section 5.2) extend the model of Findlay and Kierzkowski (1983), endogenizing the determination of education in the presence of innovation. In their model, growth is driven not by education but by innovation, and they show that an increase in the fraction of skilled workers does have positive effect on the rate of innovation. Eicher (1996) takes another approach by considering explicitly an education sector, which, in addition to providing education and human capital accumulation, also generates technological spillovers. He shows that higher rates of technological progress and growth may be accompanied by a higher relative wage but lower relative supply of skilled labor.

The Uzawa-Lucas model and many of its extensions assume that the only cost of education is the opportunity cost of the time spent on education. Some attempts have been made to relax this assumption. For example, Manning (1975, 1976), Shea and Woodfield (1996), Eicher (1996), and Wong (1997) consider an education sector in which students are educated by educators. Ohyama (1991) and Galor and Stark (1994) assume explicitly that investment in human capital requires real resources, while Bond, Wang, and Yip (1996) and Bond and Trash (1997) develop models with an education sector that requires physical capital and labor time in the production process.

4.4 Growth with Learning by Doing

Another channel through which human capital and knowledge accumulates is learning by doing. As Arrow described it, "Learning is the product of experience. Learning can only take place through the attempt to solve a problem and therefore only takes place during activity." (Arrow, 1962, p. 155.) The experience that a worker acquires through learning augments the productivity of the worker, implying that, for any given factor endowments, the production possibility set of the economy expands. This is similar to human capital accumulation through education, except that learning by doing requires very little, if any, resources: At least a worker does not have to (in fact, should not) stop working in order to learn.

However, an increase in workers' productivity may or may not lead to a perpetual growth of the economy. In formalizing the concept of learning by doing, Arrow (1962) postulates that the productivity of a given firm is an increasing function of cumulative investment in the industry. In his model, however, the growth rate of consumption converges to zero, because it is assumed that for the economy as a whole, the marginal product of capital eventually falls to zero. See also Levhari (1966a, 1966b), and Sheshinski (1967) for the same result.

In an alternative formulation, Lucas (1988) drops the diminishing returns assumption made by Arrow, and shows how the growth of an economy may depend positively on the rate of accumulation of human capital through learning by doing. Other formulations of endogenous growth with learning by doing have also been suggested by Stokey (1988) and Young (1991, 1993).

To show how learning by doing may sustain growth, let us refer to the Cobb-Douglas production function used earlier. Since workers do not have to spend time on learning, the function in (5′) reduces to

$$y = hAk^{\alpha}, \tag{20}$$

or, in terms of growth rates,

$$\widehat{y} = \widehat{h} + \widehat{A} + \alpha\widehat{k}. \tag{20′}$$

Again, assuming technological progress in a steady state $\widehat{k} = 0$ and the growth rate of per capita output depends on how human capital grows.

As Arrow (1962) suggests, the experience a worker acquires through learning depends on the amount of activity he/she goes through. However, learning is assumed to occur accidentally, and individuals do not take it into account in making consumption and time-allocation decisions.

Let us postulate that the human capital stock is a positive function of a variable Z, which is an index of accumulated experience, i.e.,

$$h = g(Z), \tag{21}$$

where $g'(Z) > 0$. In terms of growth rates, (21) gives

$$\widehat{h} = \varepsilon_g \widehat{Z}, \tag{21′}$$

where ε_g is the elasticity of the function $g(.)$. It is required that $\varepsilon_g \widehat{Z}$ be endogenously determined and remain constant along a balanced growth path.

We now explain two ways of specifying Z that ensure the above balanced growth condition. The first one is that of Lucas (1988, 1993). Suppose that the initial human capital at time 0 is given, h_0. Let Z be the cumulative human capital, and

$$g(Z) = a \int_0^t uh \, \mathrm{d}v, \tag{22}$$

where $0 < a, u < 1$. The variable u represents the fraction of time a representative worker spends on working,[7] and is determined endogenously. The parameter "a" represents the efficiency unit of labor. By equation (22), the growth rate of human capital is equal to au. Therefore equation (22) implies that the growth rate of human capital is proportional to the amount of time individuals spend on producing the good: the more time they spend on producing good, the faster human capital and per capita output will grow.

In a steady state, u is a constant, and therefore the growth rates of human capital and per capita output in the absence of technological progress are both equal to au.

Another way of modelling human capital accumulation is to assume that Z is the accumulated output, and

$$g(Z) = a \int_0^t Y \mathrm{d}\tau, \tag{23}$$

where $a > 0$ is a parameter, and Y is the output level. Equation (23) implies that the current level of human capital depends on cumulative experience, the latter being represented by the cumulative output level. Equations (20) and (23) give

$$\dot{h} = aAhMk^{\alpha}. \tag{24}$$

If we assume no technological progress or population growth, the growth rate of human capital and that of per capita output in a steady state is $aAM\widetilde{k}^{\alpha}$, where $\widetilde{k}$ is the steady-state capital-labor ratio.[8] Clemhout

7. The rest of the time may be spent on leisure. In the Lucas (1988) model, there are two sectors, and u_i is the fraction of the nonleisure time a worker spends on working in sector i.

8. If population growth rate is positive, the growth rate of human capital as given by (24) will not be constant. To have a constant growth rate of human capital, one can assume instead that Z is a positive function of the cumulative per capita output.

and Wan (1970), Stokey (1988), Young (1991, 1993), and Ishikawa (1992) make similar assumptions of human capital accumulation through learning by doing.[9]

Despite their success in endogenizing the growth rate, the above formulations of learning by doing have their weaknesses. It is widely believed that the learning curve of a person in general rises rapidly initially, but then it slows down, and eventually may become flat. Arrow (1962) is aware of this fact, and also suggested that new goods continually appear while some old goods disappear. The same argument is used by Lucas (1988), but he does not explicitly consider continuing emergence of new goods. Stokey (1988) and Young (1991, 1993) assumed that there are diminishing returns in learning by doing with respect to any given product, but because of emerging new products, growth can be sustained. Both Stokey and Young adopt the model first proposed by Wan (1975), where there is an infinite continuum of produceable goods, of which a finite number are produced at any given time.

Note that some learning-by-doing growth models imply scale effects: A country or an industry that becomes bigger in size will experience a higher growth rate of the economy (e.g., Backus et al., 1992). For example, consider again the growth of an economy due to learning by doing as implied by equation (24). The steady-state growth rate implied is $aAM\widetilde{k}^{\alpha}$. Suppose that a country becomes twice as big as before so that the number of workers increases from M to $2M$. Then the growth rate is doubled. The existence of scale effects is an uncomfortable feature of these models because it is not supported by evidence, including both time-series data for a particular country and cross-sectional data for different countries. For example, the growth rate of the United States does not seem to increase over time with its population, and a country like India does not have a higher growth rate than that of a much smaller country like Singapore. The scale effect has the further implication that countries do not converge.

However, it should be noted that not all endogenous growth models have scale effects. For example, the education model represented by equation (18) and the learning-by-doing model suggested by (22) do not have scale effects. In models in which growth of an economy depends on R&D, however, scale effects are more common.

9. Kemp (1974) suggested an alternative learning function: the human capital level in period t is a positive function of the output level in the previous period. This function may or may not lead to perpetual growth, but he did not provide an analysis of this point.

4.5 Growth with Endogenous Technological Progress

So far we have been assuming that the general technology level of the sectoral production function is either fixed or given exogenously, i.e., variable A has been kept as an exogenous variable. We now relax this assumption.

In economic theory, technology still remains a black box which we know not too much about. Technological progress is usually treated exogenously or is based on some ad hoc process. In the past decade, some new approaches and analyses have been suggested and applied to the growth theory.

Technological progress can take place in one or more of the following forms: (a) It improves the productivity of factors (or lowers the cost of production); (b) It leads to the emergence of new products; and (c) It improves the quality of existing products or productivity of existing intermediate products. Sometimes, a product with a higher quality can be regarded as a new product, and there is hence very little difference between form (b) and form (c) of technological progress. However, distinguishing between forms (b) and (c) is generally useful, because very often the emergence of new products is analyzed when products are assumed to be characterized by horizontal product differentiation, while by definition quality improvement involves vertical product differentiation.

In this subsection, we focus on technological progress that improves the productivity of factors. In terms of the production function (5), which we have been using, this is represented by an increase in the value of variable A.[10]

In a series of papers written in the later fifties and early sixties, Kaldor criticized the neoclassical assumption of exogenous technological progress. Kaldor and Mirrlees (1961–1962) suggest a formal model with perpetual growth that is dependent on new investment and saving. They postulate that the rate of growth of productivity per worker operating on new equipment is a positive function of the rate of growth of investment per worker. Therefore, policies that affect new investment directly could have growth effects.

Chipman (1970) suggests another model of endogenous technological progress. He recognizes the fact that technological improvement requires the use of resources: engineers, researchers, computer programmers, com-

10. For convenience, three types of technological progress can be distinguished: Hicks-neutral, labor-augmenting (Harrod-neutral), and capital-augmenting (Solow-neutral), but in general only the labor-augmenting technological progress is consistent with a balanced growth. See Barro and Sala-i-Martin (1995, pp. 54–55). With a Cobb-Douglas production function, these three types of technological progress when given exogenously are not distinguishable.

puters, laboratories, and so on. He postulates that the rate of technological progress is directly related to the amount of resources devoted to research. As a result, government policies that encourage research have a positive effect on the growth of the economy.

Chipman's model, while explicitly considering the use of resources on improving technology, sidesteps some fundamental questions: Who conducts research? How is the fruit of research appropriated? How is the new technology transferred to the firms?

These questions are related to each other. Specifically, in the absence of technology transfer from abroad, technological improvement can be done through R&D (i) by firms that are currently producing a good, (ii) by firms that are potential producers, (iii) by firms or agents that are specializing on research and development, or (iv) by the government.

In what follows, we provide a simple model that is based on Chipman (1970), assuming explicitly that research is being conducted by the government. Let the aggregate labor force, M, be constant over time. Because human capital accumulation is not considered, we let $h = 1$.

Workers are hired either by the firms in the production sector to produce the homogeneous good or by the government to conduct research. Let the fraction ϕ of the labor force be employed in the production sector, while the rest are hired by the government to conduct R&D. Using the notation defined earlier, the labor input in the production sector is $L = \phi M$, and the production function can be written as

$$Y = AK^{\alpha}(\phi M)^{1-\alpha}, \tag{25}$$

which implies that the per capita output, Y/M, is equal to

$$y = \phi A k^{\alpha}. \tag{25$'$}$$

The growth rate of per capita output is

$$\widehat{y} = \widehat{\phi} + \widehat{A} + \alpha \widehat{k}.$$

By employing $(1 - \phi)$ of the total labor force to carry out R&D, the government is able to improve technology according to

$$\dot{A} = \sigma A(1 - \phi)M, \tag{26}$$

where $\sigma > 0$ is an index representing the effectiveness of labor in R&D. The government then distributes the fruit of R&D to all the firms in the economy. Due to its non-rival property, technology is a public good in the sense that the government can provide the technology to an extra firm without hurting the technological level of other firms. The cost of R&D is $w(1-\phi)M$, where w is the wage rate. It is financed by lump-sum taxes on all individuals.

In a steady state, ϕ and k are constant. Therefore the growth rate of A, and thus that of per capita output, is equal to $\sigma(1-\phi)M$, where $\tilde{\phi}$ is the steady-state value of ϕ.

If the government expands its R&D activity, it can increase the rate of technological progress, but it also lowers the amount of labor available to firms. Thus there is a trade-off. It is assumed that the government chooses the size of the R&D activity, or ϕ, to maximize an objective function such as the steady-state per capita consumption or the sum of the discounted stream of the utility levels of a representative consumer.

This model, though simple, does bring out some important features of R&D. First, the new technology developed by the government is one type of public good: It is provided free to the firms, and is non-rival and excludable. Similar public goods can also have growth effects. See, for example, Barro (1990) and Turnovsky (1997).

The present model can be used to explain the divergence in growth rates of different countries. Because the technological progress is controlled by the government, unless different governments choose the same policies, one would not expect that their countries will grow at the same rate. One can further argue that why a country like the Philippines is not growing as fast as Taiwan is because the technology growth rates of the countries are chosen to be different. This argument is in fact supported by some casual observations: the positive correlation between technological progress and growth (countries that grow rapidly are usually those that experience substantial technological progress), and the important role of government in R&D in many countries. However, further thinking will raise the question: If technological progress is so crucial to growth, why does a government not choose to provide more technological progress?

The most straightforward answer is that different governments have different objective functions. (A government chooses to have a low growth rate because this is what it wants). For example, they have different preferences. In general, a country with a smaller rate of time preference (with a bigger discount of future consumption), other things being equal, will prefer to have a lower growth rate. This thus brings out an often-neglected fact: a higher growth rate does not necessarily mean a higher welfare level.

Another reason for different government R&D policies is that different governments may be subject to different budget constraints. If a government finds it too costly to raise revenue to finance R&D activities, the growth rate of the country has to be compromised. This argument is compatible with the observation that countries that grow fast usually have substantial government budget surpluses. (Of course, growth and budget surpluses may be inter-related.)

5. Innovation and Growth in a Closed Economy

In the previous subsection, we focused on the type of technological progress that improves the productivity of factors. We now turn to two other types of technological progress: the one that leads to the emergence of new products, and one that improves the quality of some existing products. The former type of technological progress is called horizontal innovation and the latter called vertical innovation.

The type of models described in the previous section with one homogeneous good is no longer suitable for analyzing horizontal or vertical innovation. Some simple ways of extending the neoclassical model are now explained.

5.1 Horizontal Innovation

The most common way of extending the neoclassical model to allow for the emergence of new products is to consider a sector of differentiated products as originally suggested by Spence (1976) and Dixit and Stiglitz (1978). The advantage of the Spence-Dixit-Stiglitz approach is that the goods enter the utility function of a representative consumer in an additive way, so that the utility is an increasing function of the number of varieties. Moreover, by treating the goods symmetrically and assuming a large number of varieties, the model can be solved in a simple way. Ethier (1982) extends the Spence-Dixit-Stiglitz model by treating the differentiated products as intermediate inputs used by other firms to produce a final product. This approach is followed by Romer (1987, 1990), Grossman and Helpman (1990b, 1991a), and Rivera-Batiz and Romer (1991a).

Following Romer (1990), Grossman and Helpman (1990b), and Rivera-Batiz (1991a), let us divide the economy into three sectors: the final-good, the intermediate-good, and the research sectors. The final good, which is a consumption good, is homogeneous and produced under perfect competition, while the intermediate goods are differentiated. There is only one type of primary factor, labor. The labor endowment is constant over time. The final good is produced using labor, L_f, and

N intermediate inputs:

$$Y = AL_f^{1-\alpha} \sum_{i=1}^{\infty} X_i^{\alpha}, \tag{27}$$

where X_i is the input of the ith intermediate good. This production function is a modified Cobb-Douglas function. Note that all immediate goods enter the production function symmetrically.

Because at time t there are N intermediate inputs available, in the production function in (27), $X_i = 0$ for $i = N+1, \ldots, \infty$. Note how the additivity of the inputs in the production function allows the inclusion of some intermediate inputs that currently do not exist.

At least three different formulations of the intermediate-good sector have been suggested. They have different economic interpretations, but similar mathematical implications. We briefly describe and compare them. The first one is due to Romer (1990), who assumes that the intermediate inputs are different types of capital. Each type of capital can be produced by sacrificing one unit of the final product. Since the intermediate products are treated symmetrically on both the demand and the supply sides, in equilibrium, equal amounts of each of the intermediate products are produced and used in producing the final good. We let this amount be X. This implies that the production function of the final good reduces to

$$Y = ANL_f^{1-\alpha} X^{\alpha}. \tag{27'}$$

The total amount of capital is $K = NX$. Physical capital accumulation comes from saving:

$$\dot{K} = Y - C, \tag{28}$$

where no depreciation of capital is assumed and where C is the consumption of the final good. Equation (28) is the usual investment equation. Firms in the intermediate-good sector compete in an oligopolistic way. Because the number of firms is restricted by the level of technology, firms may earn positive profits in equilibrium.

The R&D sector consists of a large number of firms. Each of them hires workers to conduct R&D activities which lead to new intermediate goods. The rate of increase in the number of intermediate goods, which for convenience is treated as a real number rather than an integer, depends on the knowledge each research firm possesses and the number of workers they hire. All research firms have access to the same pool of knowledge, which is assumed to be proportional to the existing number of intermediate goods. When a firm discovers a new product, it is granted a patent which lasts forever, meaning that any firms that receive a license from the innovating firm can produce the new product. This is an important feature of this type of models: There is perfect knowledge spillover in the research sector but zero spillover in the intermediate product sector. Denoting the total labor force engaged in research by L_r, the number of intermediate goods is postulated to change according to

$$\dot{N} = \sigma N L_r, \tag{29}$$

where σ is an index representing the productivity of labor in R&D. There is free entry to the research sector. In equilibrium, the cost of developing a new product is equal to the sum of the discounted stream of profits from producing the new product.

In the balanced growth path of the economy, the distribution of labor between the final-good sector and the research sector is constant, and so is the amount of each type of capital. The growth of the number of intermediate goods, as given by (29), provides the sustained growth. Both N, K, and Y are growing at a rate of σL_r.

In the model of Grossman and Helpman (1990b), the production of intermediate products requires labor using a Ricardian-type technology. Firms also compete in an oligopolistic way and earn positive profits. Along a balanced growth path, the distribution of labor among the three sectors is constant. This implies that the total quantity of intermediate products, NX, is constant. This is contrary to the assumption in the model in Romer (1990) in which along a balanced growth path the quantity of each intermediate product, X, is constant.

The production function of the final good can be written as

$$Y = A L_f^{1-\alpha} (NX)^{\alpha} N^{1-\alpha}. \tag{27''}$$

As NX is constant along a balanced growth path, the growth rate of Y is equal to $(1 - \alpha)$ times that of N.

The above two models share one common feature of the research sector: the growth of the number of intermediate products depends on two factors, the amount of labor employed and the existing level of knowledge. Rivera-Batiz and Romer (1991a) call this knowledge-driven (KD) specification of research. They propose an alternative formulation which they call the lab-equipment (LE) model: the technology for research uses the same inputs as the final-good technology, in the same proportions. In other words, research requires both labor and intermediate products, just like the production of the final good, and is independent of the existing knowledge. Thus the change in the number of intermediate goods is given by

$$\dot{N} = BL_f^{1-\alpha} \sum_{i=1}^{\infty} X_i^{\alpha}, \tag{30}$$

where B is an R&D technology index, which is given exogenously. If we keep the assumption that intermediate inputs, which are different types of capital, are produced from the final good, the equilibrium condition of the final-good market is

$$C + \dot{K} + \dot{N}/B = Y = AL_f^{1-\alpha} K^{\alpha} N^{1-\alpha}. \tag{31}$$

Recall that $K = NX$, and that X is constant in a balanced growth path. Thus, $\widehat{K} = \widehat{N}$. In other words, the source of growth in this model, as that in the KD model, is the continuous emergence of new products. Furthermore, equation (31) shows clearly how saving, which is equal to $Y - C$, affects the growth of the economy. In another version of the LE model, the intermediate goods are *non-durable*. (See Barro and Sala-i-Martin, 1995.) Then, $\dot{K}$ in (31) is replaced by NX.

The above models of horizontal innovation can be used to explain why countries may have different growth rates, and why these rates may not converge over time, even if they have the same technology. Saving, which may be determined by market forces or chosen by the government, is the key factor behind an economy's growth. In these models, saving also contributes directly to the growth of the number of new products.

5.2 Vertical Innovation: Schumpeterian Creative Destruction

A rigorous theory of repeated quality upgrades of existing products was first developed by Segerstrom et al. (1990) and Aghion and Howitt (1992). The former paper assumes that the time of arrival of a new invention that replaces an existing product is a deterministic function of the aggregate R&D expenditure in the industry, but the identity of the successful inventor is a random variable. Another assumption is that the patent races take place sequentially in one industry after another, in a predetermined order. Aghion and Howitt, on the other hand, assume that the time of arrival is stochastic, but there is only one firm producing the intermediate good, which is rendered obsolete by the arrival of a new invention. Thus there is a sequence of monopolists, the new one stealing the business of the old one. The following is a simple version of the Aghion-Howitt model.

There are two types of goods: a consumption good and an intermediate good; and two types of workers of fixed stocks over time: unskilled workers and skilled workers. Unskilled workers are used to produce the consumption good, while skilled workers can be used to produce the intermediate good or to perform R&D. Let y, x and A denote the output of the consumption good, the amount of the intermediate input, and the quality of the latter. With a fixed endowment of the unskilled workers, the production function for the final good is written as

$$y = AF(x), \tag{32}$$

where F is strictly concave and increasing. Let L_r and $L_x = L - L_r$ denote the amounts of (skilled) labor employed in R&D and in intermediate good production, respectively, with the amount of skilled labor L fixed over time. The consumption good sector is perfectly competitive. The production of the intermediate good requires labor only. Assuming a linear technology, $x = L_x$.

The quality of the intermediate good is measured in terms of its productivity in producing the consumption good. Its quality can be upgraded, and each upgrade represents a constant multiple of the original productivity. Thus we write

$$A = \gamma^m, \tag{33}$$

where $\gamma > 1$ and m (an integer) is the number of times the intermediate good have been upgraded (which is the same as the number of innovations that have occurred).

Because the intermediate good, no matter what its quality level is, is produced with one unit of skilled worker per unit of output, the firm that has the technology of producing the good with the highest quality will capture the whole market, and is therefore a monopolist. Taking the wage rate of skilled workers as given, it chooses the price of the intermediate good it produces to maximize its profit.

Quality improvement of the intermediate good is done by research firms. Note that these outside firms get a bigger return from a successful research than the existing monopolist has, because they do not have to pay the price of losing the profit from the prevailing quality. Thus the monopolist chooses to do no research. While the amount of each upgrade is fixed, the time at which an innovation occurs is random. Let $\pi(m,t)$ denote the probability that there will be m innovations up to time t. The expected output of the consumption good at time t is

$$Z(t) = \sum_{m=0}^{\infty} \pi(m,t)\gamma^m F(L - L_r). \tag{34}$$

Assume that the innovation process is Poisson with parameter $\lambda\phi(L_r)$ representing the arrival rate. Then

$$\pi(m,t) = \frac{[\lambda\phi(L_r)t]^m e^{-\lambda\phi(L_r)t}}{m!}. \tag{35}$$

Substitute (35) into (34) to obtain

$$Z(t) = F(L - L_r)e^{\lambda\phi(L_r)(\gamma-1)t}. \tag{36}$$

A research firm that is successful in its innovation gets patent protection from the government and sells its technology to a (new) intermediate producer. Being the only firm with this new technology, it extracts all the monopolist rent from the new intermediate good producer. Therefore when a research firm chooses to do research, the value of the next innovation is the expected present value of the flow of monopolist profit generated by this new innovation over an interval whose length is exponentially distributed with parameter $\lambda\phi(L_r)$. Note that even though the patent prevents any horizontal spillover between firms, there are intertemporal spillovers, as each successful innovation raises the general knowledge base, helping the next innovation.

The research firms choose the amount of skilled labor to do research, taking the wage rate as given, to maximize its expected profit. The equilibrium of the economy is characterized by the labor market equilibrium.

The steady state of the economy requires that the distribution of skilled labor is constant over time. Denote the steady-state value of L_r in this decentralized economy by $\widetilde{L}_r$. From (36), the instantaneous rate of growth of expected consumption is $\lambda\phi(\widetilde{L}_r)(\gamma - 1)$. Note that $\gamma - 1$ is approximately the same as $\ln\gamma$.

It is immediately clear that any policies that directly increase the employment of labor in the research sector will increase the growth rate of the economy in the sense that future innovations tend to arrive sooner.

The above model has important welfare implications. Suppose there exists a social planner who chooses labor distribution to maximize

$$W = \int_0^\infty e^{-rt} Z(t)\, \mathrm{d}t, \tag{37}$$

where r is great enough to ensure the convergence of the integral. Denote the optimal amount of labor in the research sector by L_r^*. Whether the growth rate (of the expected consumption) of the decentralized economy is higher or lower than the optimal growth rate depends on whether $\widetilde{L}_r$ is greater or smaller than L_r^*. However, in general, the sign of $\widetilde{L}_r - L_r^*$ is ambiguous. For example, consider the special case in which $\phi(L_r) = L_r$ and $F(x) = x^\alpha = L_x^\alpha$. It can be shown that

$$\widetilde{L}_r = \frac{\lambda\gamma(1-\alpha)L - \alpha r}{\lambda[\gamma(1-\alpha)+\alpha]}, \tag{38.1}$$

$$L_r^* = \frac{\lambda(\gamma-1)L - \alpha r}{(1-\alpha)(\gamma-1)\lambda}. \tag{38.2}$$

Which of these two labor employment is bigger is ambiguous. If $\gamma = 2$, $\alpha = 0.5$, $\lambda = 1$, then $\widetilde{L}_r < L_r^*$, but if γ is close to unity and α is very small, then $\widetilde{L}_r > L_r^*$.

Aghion and Howitt (1992) offer four reasons for the difference between the two growth rates: the intertemporal-spillover effect (private research firms attaching no weight to the benefits that accrue beyond the succeeding innovation), the appropriability effect, the business-stealing effect (the private research firm not internalizing the loss to the previous monopolist caused by an innovation), and the monopoly-distortion effect. The intertemporal-spillover and appropriability effects tend to make the laissez-faire average growth rate less than optimal, whereas the other two effects affect the laissez-faire average growth rate in an opposite direction.

Building on the work of Segerstrom et al. (1990) and Aghion and Howitt (1992), Grossman and Helpman (1991d) suggest an alternative model of vertical innovation. They postulate a continuum of final goods, each with its own quality ladder. Patent races take place simultaneously and are risky.

The intertemporal utility of a representative consumer is

$$U = \int_0^\infty e^{-\rho t} \ln u(t)\, \mathrm{d}t, \tag{39}$$

where ρ is the rate of utility discount, and $\ln u(t)$ represents the flow of utility at time t and is defined as

$$\ln u(t) = \int_0^1 \ln\left[\sum_{j=0}^{\infty} q_j(s) X_j(s)\right] \mathrm{d}s, \tag{40}$$

where $X_j(s)$ is the consumption of quality j of product s. If a quality is not yet available, its price is infinity. The consumer chooses the consumption bundles to maximize her utility as given by (40), subject to the intertemporal budget constraint

$$\int_0^\infty D(t)E(t)\, \mathrm{d}t \leq M(0), \tag{41}$$

where $E(t)$ is the expenditure at $t, D(t)$ is the discount factor, and $M(0)$ is the present value of the consumer's income stream. In equilibrium, the safe interest rate is $r(t) = -\dot{D}/D$. The solution to the utility maximization problem is

$$\frac{\dot{E}}{E} = r(t) - \rho. \tag{42}$$

For convenience, the problem of the consumer can be broken up into two steps. First, she chooses the streams of expenditure, $E(t)$, to maximize her intertemporal utility; then taking the expenditure $E(t)$ as given, she chooses the consumption of each product to maximize her instantaneous utility.

On the production side, several assumptions are made to give a tractable model. First, quality of a product is measured in fixed increments: quality j of product s is given by $q_j(s) = \gamma^j$, where $\gamma > 1$ is the same for every s. Second, one unit of labor is needed to manufacture one unit of any product, regardless of quality. Third, firms compete in a Bertrand fashion. Fourth, the leader always stands exactly one step ahead of its nearest rival.

These assumptions have several implications. If there are several firms producing the same product of the same quality, Bertrand competition implies that all of them earn zero profit. If there is one leader in each industry with some potential firms being able to produce the product with inferior quality, the leader can set the price low enough to drive the followers out of the market, leading to only one producer in each market. Also, being only one step ahead of the nearest rival, the leader will set the "limit" price as

$$p = \gamma w, \tag{43}$$

where w is the wage rate. Note that the same price is set for all products of the leading quality. Condition (43) further implies that the demand for the product is equal to $E(\gamma w)^{-1}$, and that the flow of profit of each monopolist producing each product of the highest quality is equal to $(1 - \gamma^{-1})E$. This profit disappears when the product of a higher quality is invented and produced.

Research for quality improvement is done by potential competitors of the existing monopolist. By the same argument presented above, the return of a successful innovation is bigger to an outside firm than to an existing firm. A research firm, knowing enough about the state of knowledge, hires a_r units of labor per unit of R&D activity per unit of time, producing a probability of success of $\tilde{\iota}dt$, where $\tilde{\iota}$ is the R&D

intensity for a time interval of $\mathrm{d}t$. The research success is thus Poisson with the arrival rate dependent on the level of R&D activity.

With symmetry between the industries, we let ι be the aggregate research intensity, and L_r be the aggregate labor employment. Therefore $\iota = L_r/a_r$. While in each industry there is randomness in R&D success or failure, for the economy as a whole the law of large numbers ensures that in the aggregate there is virtual certainty.

There is free entry into the patent race. In equilibrium, no arbitrage implies that the expected rate of return of investing in a research firm is equal to the safe interest rate. Making use of this "no arbitrage" condition and condition (42), we get the adjustment of a consumer's expenditure:

$$\frac{\dot{E}}{E} = \left(1 - \frac{1}{\gamma}\right)\left(\frac{E}{a_r}\right) - \rho - \left(\frac{L_r}{a_r}\right). \tag{44}$$

It is noted that the manufacturing employment is $L_x = L - L_r = E/\gamma$. Using this condition, (44) reduces to

$$\frac{\dot{E}}{E} = -\rho - \frac{L - E}{a_r}. \tag{44$'$}$$

Since the adjustment of E as given by (44′) is unstable, it is argued that the system jumps to that steady state instantaneously. It follows that E is a constant, and hence, from (42), $r = \rho$ always. From this, we can solve for the steady state L_r which is positive, provided L is sufficiently large.

The growth rate of the instantaneous utility can be obtained from (40), after simplification, and given the fact that the research success is Poisson-distributed, it is equal to

$$G_u = \frac{L_r \ln \gamma}{a_r}. \tag{45}$$

Condition (45) implies that the growth rate of the instantaneous utility is proportional to the employment in the research sector.

The welfare implications of the Grossman-Helpman model are similar to those of the Aghion-Howitt model. In particular, the R&D expenditure in the market economy may be smaller or greater than the socially optimal expenditure. The latter case occurs when γ is close to unity, or when it is quite large. This result is consistent with the findings of Aghion and Howitt (1992).

5.3 Comparing Different Types of Technological Progress

We have distinguished between several types of technological progress: factor productivity improvement, horizontal innovation, and vertical innovation. These three types of technological progress enlarge the production and consumption possibilities of an economy in different ways. They therefore have different implications for both the production and the consumption of the economy.

As surveyed above, models describing different types of technological progress vary a lot in terms of the underlying preferences, market structures, production technology, features of the research sector, extent of technology spillover, the role of the government, and so on. The results obtained also vary a lot. Moreover, the growth of an economy is usually measured in different ways. For factor productivity improvement, growth of an economy is represented by the growth rate of the per capita income or output. For horizontal innovation, it is the growth of the number of varieties, and for vertical innovation, the growth rate of the (instantaneous) utility of a representative consumer is a good measure of the growth of the economy.

Despite the differences between their economic interpretations, these models have very similar mathematical expressions, especially the expression for the growth rate of an economy. In particular, the growth rate of an economy in a steady state, as one may note from these models, can always be expressed as an increasing function of the employment engaged in the research activity. (See more discussion below.)

These models also have very different implications on empirical studies. Suppose one wants to determine the growth of factor productivity of an economy. The straightforward way is to compare the growth rate of per capita output and that of capital-labor ratio (assuming a two-factor, one-sector economy). See, for example, Young (1994). However, to measure the other two types of technological progress is much more difficult.

5.4 Scale Effects of R&D

The R&D models introduced above carry the implication that an increase in the size of the economy or the size of the R&D sector will increase the growth rate of the economy. This effect, which is called the scale effect of R&D, is embedded in equations (26) (for improving factor productivity), (29) and (30) (for increasing the number of varieties), (36) (for the growth rate of expected consumption), and (45) (for growth rate of the instantaneous utility). Since these growth rates are directly related to the growth rate of the economy, these equations imply that an increase in the level of employment in the research sector will increase the growth rate of the economy.

The existence of scale effects of R&D comes from the appealing idea that the bigger the knowledge base and the more resources devoted to research, the easier it is to accumulate more knowledge. This idea reflects three important features of knowledge. First, knowledge has an intertemporal spillover effect, which allows the economy to accumulate knowledge and sustain growth. Second, knowledge is non-rival, meaning that it can be used by more than one agent simultaneously without affecting the benefit each of them gets from using the knowledge. Third, in many cases, knowledge is non-excludable; for example, the general knowledge reported in scientific journals. These three features imply that when an innovator introduces new knowledge, it not only improves its own competitiveness, but also raises the knowledge base of the economy and thus helps other and future firms in their R&D efforts. These effects thus have the implication that a large country, or a large research sector, will lead to a higher growth rate. They also have policy implications. For example, policies that encourage the employment in the research sector have positive effects on growth.

These R&D models, however, have been under criticism recently, because the implications of these scale effects are not supported by observed data. For example, Backus, Kehoe, and Kehoe (1992) find little empirical evidence of a relation between the growth rate of GDP per capita and several measures of scale implied by the theory. They do find a significant relation between the growth rate of output per worker and the relevant scale variables. Jones (1995a) points out that the U.S. growth rates exhibit no large persistent changes, even though there have been permanent changes in certain government policies that, according to the endogenous growth theory, should have effects on growth. Similarly, there are little or no persistent changes in growth in other OECD countries. Jones (1995b) further points out that while the number of scientists and engineers employed in R&D in the United States grows by more than five times from 1950 to 1988, the total factor productivity growth for the same period is constant or even negative.

It is noted that the scale effects come from the formulation of technological progress: growth caused by R&D is directly proportional to the amount of resources (such as the number of engineers) engaged in R&D. The scale effects go away if growth is written as a function of some scale-free variables, such as the share of labor working in the research sector. This alternative formulation, however, is not satisfactory, because it is contrary to the belief that innovation is tied to the number of people engaged in the research activity, and moreover, it is also rejected by the

U.S. evidence because, as Jones (1995b) shows, the share of scientists and engineers in the total labor force has also gone up.

Several efforts have been made to eliminate the scale effects in R&D. Jones (1995b) modifies the R&D equation by allowing declining rate of innovation with the level of knowledge and externalities due to duplication in the R&D process. Segerstrom (1995), following Lucas (1988), introduces human capital which grows through education and knowledge spillover. [See equation (18).] An alternative formulation is introduced by Young (1995), where there are both vertical innovation (quality improvement) and horizontal innovation (increase in the number of varieties). To avoid scale effects, he assumes intertemporal knowledge spillover in the vertical dimension, but not in the horizontal dimension. A larger market will lead to an increase in the number of horizontal product varieties, thus affecting the *level* of utility, but not the growth *rate*. Eicher and Turnovsky (1996) extend Jones' approach and develop a more general model that may or may not have scale effect.

Even though these papers suggest models with no scale effects, it seems that this is achieved at a cost of eliminating the endogeneity of growth due to R&D. Because Jones (1995b) assumes a declining rate of innovation, the growth of the economy decreases over time until it reaches a level that is directly proportional to the growth rate of population, the proportionality constant being dependent on some exogenous parameters. In other words, Jones' model, though assuming endogenous R&D, implies exogenous growth: government policies such as R&D subsidy have no growth effect. Jones describes his model as "semi-endogenous." This feature is also shared by the model of Eicher and Turnovsky (1996), when scale effect is absent. Segerstrom's (1995) model has endogenous growth, but endogeneity comes from education and human capital accumulation, not from R&D.[11] Thus, education subsidies have growth effects, but R&D subsidies do not. In Young's model, the absence of scale effects implies exogenous growth, even though vertical innovation and horizontal innovation are determined endogenously. Thus government R&D subsidies or trade policies have no growth effect, even though the number of varieties and welfare may change.

6. Trade and Endogenous Growth

So far, we have examined growth of closed economies. We now try to see how the above models can be extended to open economies. In this

11. Segerstrom's result is not surprising, because from equation (18) we know that human capital accumulation through education could have endogenous growth without scale effects.

section, we focus on international trade in goods. In the next section, we will look at international factor mobility.

To analyze trade and growth, note that the neoclassical one-sector, homogeneous-good model is not suitable for considering trade. Either a multi-sectoral model or product differentiation has to be considered. This can be done easily by extending the models introduced above.

6.1 Trade and Growth with Physical Capital Accumulation

We first consider models where growth is driven by capital accumulation alone. Fisher (1995) extends the two-sector AK model of Jones and Manuelli (1990) and Rebelo (1991) to an overlapping-generations model. Individuals live for two periods, inheriting nothing when born except being endowed with one unit of labor, and leaving no bequest when dead. Each individual works, saves, and consumes only when young, and consumes when old. Thus in this model saving of the economy comes entirely from workers when they are young. Population and labor force are constant over time.

The consumption good is produced by labor and capital using a Cobb-Douglas production function, and the investment good is produced with capital only and with constant marginal product of capital. Markets are perfectly competitive. Fisher shows that with sufficient saving, the growth rate of the capital-labor ratio of a closed economy is

$$\widehat{k} = \frac{s(1-\delta+\beta)(1-\alpha)}{\alpha + s(1-\alpha)}, \tag{46}$$

where s is every individual's savings as a fraction of the wage rate, and α, β are technology parameters defined by (11.1)–(11.2). Assuming a Cobb-Douglas utility function for every individual, s is constant.

Now consider two countries with identical technology and preferences, except with different time preferences. In particular, they have different values of s. An important feature of the present two-sector AK model is that the investment good is infinitely capital intensive, because it employs no labor. This has two very important implications when free trade is allowed. First, because the more thrifty country (with a bigger value of s) has a higher growth rate of capital-labor ratio, it tends to have a comparative advantage in the investment good; for example, if they begin with the same capital-labor ratio, then in the next period, the thrifty country will become capital abundant. Second, if a country is completely specialized under free trade, irrespective to the trade patterns, it must produce the labor-intensive consumption good only. These two points combined together imply that if the two countries have substantially different factor endowment ratios so that complete specialization occurs under free trade, then the less thrifty country will be completely specialized in producing the consumption good, while the other country is diversified. In this case, the less thrifty country has a lower wage-rental ratio. All investment will occur in this country and none in the thrifty country. This means that the former has a growing capital-labor ratio while the ratio in the latter country is constant. Sooner or later, the capital-labor ratios of the countries are close enough so that both countries are diversified.

When both countries are diversified, the usual argument shows that factor price equalization (FPE) exists, meaning that the countries reach an integrated equilibrium of the world under free trade. Both countries have the same capital-labor ratio, which grows over time according to equation (46), except that the saving rate is the weighted average of those of the countries. Two implications can be drawn. First, at this integrated equilibrium, each country's share of the world wealth remains constant. Second, the growth rate of the world is in between the autarkic growth rates of the countries. As a result, the more thrifty country experiences a drop in its growth rate, while the less thrifty country gets a faster growth rate. The possibility is that trade can reverse the autarkic growth path of a country (Fisher, 1995). Third, because both countries have the same capital-labor ratio and grow at the same rate, they will remain diversified, with FPE forever.

Fisher and Vousden (1995) extend Fisher's model to analyze the effects of changes in tariffs and of the formation of customs unions and free trade areas. They show that policies that encourage the import of the consumption good by countries with high saving rates will provide a source of increased outward foreign investment and stimulate growth.

Jones and Manuelli (1990) show that in an AK model with infinitely lived agents, trade liberalization can have growth effects. In their model, there are no externalities, and laissez-faire is therefore optimal for the world as a whole.

6.2 Trade and Growth with Human Capital and Learning by Doing

Lucas (1988) extends his one-sector model of accidental learning by doing to a two-good model, and examined the roles of human capital accumulation in international trade. His model illustrates some of the features of dynamic models that we find in other papers. (See, for example, Ishikawa, 1992.) So we present a brief description of his model and results.

There are two consumption goods. Consumers have homothetic preferences so that the ratio of the demands for the goods is a function of the relative price. The two good sectors are characterized by perfect competition and one input, labor. Workers can accumulate experience, or human capital, by working in a firm. As assumed earlier, learning is accidental in the sense that no one will take the learning process into consideration in choosing employment or production. Following equation (22), the growth rate of human capital accumulation in sector i, $i = 1, 2$, is postulated to be $a_i u_i$, where u_i is the fraction of the labor force working in the sector, and $a_i > 0$ is a measure of the efficiency of learning. Without loss of generality, assume that sector 1 is the "high-technology" sector with $a_1 > a_2$. Ricardian technologies are assumed, i.e., the output of a good is equal to the efficiency units of labor input (by a choice of labor unit). This means that the marginal product of labor in a sector is equal to the level of human capital specific to that sector. If both goods are produced, profit maximization implies that the price ratio is equal to the reciprocal of the ratio of skill levels in the two sectors.

The first question we can ask is whether the economy, if closed, will be diversified. This question can also be asked for a static model, but for a dynamic model, this is a more interesting question because the price ratio may change over time. Suppose that an economy is diversified in a steady state, with the price ratio staying stationary. This requires that the two types of human capital grow at the same rate, or that $a_1 u_1 = a_2 u_2$. This will indeed be an equilibrium if at the corresponding price ratio the good markets clear. Note that because the technological coefficients in both sectors are determined endogenously, the autarkic price ratio depends on both the technology and preferences of the economy.

Analyzing the stability of the steady state is less straightforward. It turns out that it depends on the elasticity of substitution between the goods. If the goods are poor substitutes, the steady state tends to be stable with diversification in production, because consumers prefer to consume positive quantities of both goods. If the goods are good substitutes, then the steady state with diversification is unstable. For the case of CES preferences, the critical value of the elasticity of substitution is unity.

Lucas (1993) extends the above model to trade, assuming a continuum of small countries facing exogenously given world prices under free trade. The comparative advantage of a country depends on the country's autarkic price ratio and the world's price ratio. As in a static model, countries tend to be completely specialized, but what is different in the present dynamic model is that a country will accumulate only the type of human capital that is specific to the good produced. Therefore, when different countries are producing different goods under free trade, they will have different growth rates: Countries do not converge, even if they have the same technologies, as long as they have different preferences and different autarkic price ratios.

For the case of CES preferences and if the elasticity of substitution is greater than unity, then countries that produce the "high-technology" good will grow faster. Over time, the growth of this sector tends to drive down the relative price of this good, and if this terms-of-trade effect dominates the direct effect of productivity improvement, then those countries with faster technological improvement will have slower real income growth, a phenomenon analogous to immiserizing growth. Furthermore, if the relative price of this "high-technology" good is decreasing over time, there may come a time at which countries that are producing this good may switch to producing the other good.

Lucas' model has some interesting policy implications. Consider a country which has a long-run comparative advantage in the "high-technology" good. Suppose that currently it is under autarky, but has not reached its steady state, and that it shows a short-run comparative advantage in the "low-technology" good. If the country adopts a free-trade policy, it will export the latter good, become completely specialized in it and never produce the "high-technology" good. In terms of the economy's growth, the "right" policy for this country is to restrict (or even prohibit) trade at first and let the economy adjust closer to its steady state. When the economy has gained a comparative advantage in the fast growing good, trade can then be liberalized. A similar argument is also presented by Krugman (1984). However, neither Lucas (1988) nor Krugman (1984) provides a welfare analysis.

Other models of trade with learning by doing have been suggested. For example, Young (1991) also consider accidental learning and allows for spillovers across goods. He shows that less developed countries (LDC) would experience higher growth rates under autarky than under free trade. This loss from trade may be compensated for by the usual static gains from trade. The fall in growth when an LDC is opened to trade is due to the fact that static comparative advantage causes the LDC to specialize mostly in traditional goods, where learning has been exhausted. However, in the special case where the initial gap between an LDC and a DC (developed country) is small, under free trade, the LDC can overtake the DC if it has a greater work force. This result reflects the assumption that learning is an increasing function of the scale of production. An implication of this model is that in a world with two identical economies, temporary subsidies to high-tech industries in one country will give the country a permanent advantage. In a recent hybrid model, Young (1993) combines invention with learning by doing as complementary activities, but the implications for trade have not been explored.

Stokey (1991) distinguishes individual human capital from the social stock of knowledge. The former disappears when the individual dies, but private investment in human capital raises the social stock of knowledge. There is a continuum of goods, already invented, with quality ranging from zero to infinity. High quality goods can only be produced by workers with a higher stock of human capital. (Two workers with human capital level one may not produce the good that one worker with human capital level two can, in sharp contrast with the Lucas (1988) formulation.) Along a balanced growth path, human capital and the index of the highest quality good in existence grow at the same rate. What will happen to the growth rate of a backward country that decides to renounce autarky and embrace free trade? Assuming that there are no international knowledge spillovers, it can be shown that the investment in human capital in that country will fall. The reason is simple: free trade reduces the reward to the highly skilled labor in the backward country. This in turn reduces the incentive to accumulate human capital in that country. This does not necessarily mean that trade is harmful, because the usual static gains from trade may outweigh the loss caused by a fall in the growth rate of human capital.

A more recent work that analyzes the relationship between technological transfer through learning by doing and trade is Van and Wan (1997). Drawing upon the contagion theory suggested by Findlay (1978), they argues that technological progress, foreign trade, and factor accumulation are complements in the growth of an economy. Thus, foreign trade provides a channel to an economy through which it learns from other economies, and physical capital accumulation, instead of being a source of growth, is the consequence as the economy grows.

Bond and Trash (1997), making use of the Uzawa-Lucas model of education and extending the work of Bond, Wang, and Yip (1996), analyze the growth and trade of an economy that is characterized by human capital and physical capital accumulation. They show that under free trade between the economy and another one, both economies may experience balances or unbalanced growth. In the case with balanced growth, they derive a result related to the patterns of trade similar to the static Heckscher-Ohlin theorem.

Wong and Yip (1997) analyze the effects of industrialization and international trade on economic growth in a two-sector model with learning by doing. The interesting feature of their model is that the two sectors grow at different rates (in fact, zero growth for the agricultural sector) in a balanced growth path, thus making the relative price of manufacturing decline over time. This is in sharp contrast to most multisector models in the literature, where sectors grow at a uniform rate in a balanced path, with constant relative prices. Whether the economy is diversified under trade in the Wong-Yip model has important implications on the growth of the economy.

6.3 Technological Progress, Trade, and Growth

The models on technological progress and endogenous growth described above can be extended to analyze trade and growth. As we explained above, technological progress, either in the form of an improvement in the productivity of factors, emergence of new products, or quality improvement is due to the R&D efforts made by either profit-seeking entrepreneurs or the government. Because R&D activities require the explicit use of resources, they must be supported and financed by savings (or taxes). Thus, when we bring two economies together, and allow the flow of goods (or ideas), and analyze the effects of trade and other policies on growth, we focus on two major issues: how these policies may affect the R&D efforts through a change in the amount of resources allocated to the research sector and the productivity of these resources in conducting research activities, and whether international knowledge spillover occurs. As will be shown later, there are no unanimous answers to the above questions.

Let us first consider the case of trade with horizontal innovation. Suppose that there are two identical economies that are initially separated and are at their balanced growth paths. Two separate ways of trade between the economies are considered: free trade in goods (at least intraindustry trade in the differentiated intermediate goods) but not ideas (i.e., no international knowledge spillover and complete patent protection in the world), and free trade in ideas (perfect international knowledge spillovers) but not goods. In these cases, how would trade affect the growth rates of the countries?

Consider first the knowledge driven (KD) models. Recall that in this type of model, the growth of the economy comes from the growth of new products, while the increase in the number of new products depends on the existing knowledge base and the amount of labor employed in the research sector. Whether the growth of each economy changes is dependent on how the knowledge base and/or the research employment may change.

If there is no trade in ideas, i.e., no international knowledge spillover, then the current knowledge base of each country will not change. How may the research employment change? The answer to this question depends on whether the production of intermediate goods requires sacrifice of the final good or requires primary inputs, and whether the final goods in the two countries are homogeneous. If there is no trade in the final good because of homogeneity, and if production of the intermediate capital goods requires the final good, the employment in the research sector is not affected by trade. As a result, trade has no effect on the growth of each country (Rivera-Batiz and Romer, 1991a).

Suppose we consider an alternative case, in which there is free trade in ideas, i.e., perfect international knowledge spillover, but no trade in goods. Suppose further that the ideas in the two countries are nonintersecting. Then the international knowledge spillover will double the knowledge stock in each country. Even if the research employment does not change, the growth rate of each country will be doubled. In fact, because of the increase in profitability in the research sector, firms will employ more labor, meaning that the growth rate of each country will be more than doubled (Rivera-Batiz and Romer, 1991a).

In the lab-equipment (LE) model of Rivera-Batiz and Romer (1991a), the rate of change of the number of new products is independent of the existing knowledge stock, implying that free trade in ideas between the countries will have no economic effect. Free trade in goods (only intra-industry in the differentiated capital goods), however, will have a positive growth effect. The reason is that the intra-industry trade increases the profitability of research, thus drawing more labor into the research sector and creating a higher saving rate.[12]

In the case where there are perfect international knowledge flows, countries converge to a common growth rate and global stability is assured; see Wälde (1996) for a proof. What happens if there are only *partial* international knowledge flows? Feenstra (1996) shows that if the domestic knowledge is the sum of its past innovations and a *positive* fraction of past innovations abroad, then countries will have a common growth rate in the long run; however, if spillovers depend on the volume of foreign inputs used at home, then countries will in general differ in their long run growth rates.

We now turn to some other issues related to horizontal innovation and the above models. The first one is about the stability of a steady state. Many papers have not paid much attention to this issue, but an exception is Devereux and Lapham (1994). They note one important

12. In the knowledge-driven model, free trade in goods also causes an increase in profitability in the research sector, but does not lead to an increase in the employment in the research sector because the positive effect is exactly offset by the increase in the marginal product of labor in the final-good sector.

feature of the KD model of Rivera-Batiz and Romer (1991a) without international knowledge spillover. They show that if the home country's initial stock of knowledge is smaller than that of the foreign country, then the opening of trade will cause the home country to devote more human capital to manufacturing, and its share of knowledge in the world stock of knowledge will eventually go to zero. However, with trade, world growth rate will exceed the autarkic growth rate, because the foreign country, which has an initial comparative advantage in R&D, will devote more resources to this sector. The Devereux-Lapham instability result holds only when there are no international knowledge flows. For a somewhat different model with a similar instability result, see Grossman and Helpman (1991a, Chapter 8).

Another issue analyzed in the above models is about policy implications. Rivera-Batiz and Romer (1991b) study the effects of trade restrictions on growth in a world with two identical countries that produce non-overlapping intermediate goods. Both countries impose a tariff on all imported intermediate goods. They show that the growth rate is a non-monotone function of the tariff rate: it declines when the tariff rate rises from zero, but after some positive critical value of the tariff rate, the growth rate rises, though it never reaches the growth rate in the free trade regime. This non-monotonicity is a rather surprising result. Essentially, the tariff has two effects, a trade distortion effect and a R&D resource reallocation effect. When two effects work in opposite directions, the size of the tariff rate may determine their relative strength.

Grossman and Helpman (1990b) also study the effects of tariffs. As explained earlier, the growth effect of a policy depends on how it affects the amount of the resources (labor) devoted to the R&D sector. Suppose that country 1 has a comparative advantage in R&D. If country 2 imposes a tariff on country 1's export, more labor will be driven to the R&D sector, thus improving the latter country's growth. In the presence of international knowledge spillover, both countries grow at the same rate in the long run, and the tariff can improve this growth rate. For the same reason, an R&D subsidy imposed by country 2 could hurt the growth of both countries if international knowledge spillover is present. It is because the R&D subsidy draws resources from the country's production sector to the R&D sector. This policy thus encourages country 2's export but discourages that of country 1, hurting the R&D activity in the latter country which has a comparative advantage in R&D. As a result, the world's growth rate tends to be hurt by the subsidy.

However, a faster growth does not necessarily imply a higher welfare, a point made clearly in Grossman and Helpman (1991e). They show that a trade policy that speeds up growth may reduce welfare if, for example, it causes a fall in the outputs of the intermediate goods that are undersupplied due to monopolistic pricing.

So far we have been focusing on trade under technological progress with horizontal innovation. How would the above results be different if instead vertical innovation exists? The several models of vertical integration introduced in the previous section can be extended to open economies. Consider the vertical innovation model of Grossman and Helpman (1991d). Let us modify it so that there are two primary factors, skilled labor (H), and unskilled labor (L). There is also an outside good that does not benefit from innovation, and is assumed to use unskilled labor intensively. Suppose the foreign country is relatively well endowed with unskilled labor. Under certain assumptions (such as identical technology and diversification), the two countries without international factor mobility achieve a world integrated equilibrium. The production pattern is then identical to that which would obtain under international factor mobility.

Does this equilibrium achieve a higher growth rate than the autarkic growth rate? Grossman and Helpman (1991d) show that the answer is in the affirmative if the elasticity of substitution in the production of the outside good is greater than one. This is because (a) an increase in H will increase the supply of skilled labor for R&D, and (b) an increase in L will increase the w_H/w_L ratio, and the outside good sector will release skilled labor (despite the Rybczynski effect, which implies that, at constant factor prices, more labor of both types will be demanded by the outside good sector). On the other hand, if the elasticity of substitution is less than one, then the Rybczynski effect may dominate, causing a worldwide contraction of the R&D sector, thus slowing growth.

Tariff policies for a small open economy are the subject of study in Grossman and Helpman (1991e). The protection of a final good that uses human capital intensively will raise the reward to human capital and make R&D costly, thus slowing growth. However, faster growth does not necessarily mean higher overall welfare for this economy. A trade policy that speeds up growth may reduce welfare if it causes a fall in the output of the intermediate goods, that are undersupplied due to monopolistic pricing.

Issues related to trade patterns and specialization with vertical innovation are examined by Taylor (1993). He generalizes the Grossman-Helpman quality-ladder model by allowing asymmetry among the continuum of goods. Under the Ricardian technology, the interaction between the comparative advantage rankings in production and in innovation determine the long-run pattern of trade.

6.4 Technology Transfer

In the previous subsection, two polar cases in terms of the flow of technology between countries are examined: the one with costless and instantaneous knowledge spillover, and the one with absolutely no knowledge spillover. Both cases are not realistic in the world. Production technology, in the form of knowledge that can be described in blueprints or embedded in finished products, have many properties of a public good: It is non-rival and non-excludable. It may be transferable from one firm to another, whether the firms are in the same or different countries, and the use of it by an additional firm does not affect the use of it by the existing firm. However, because the technology, if it is advanced, allows the user to produce a new or better product or to improve the productivity of the employed factors, the firm that has the sole possession of it wants to guard its secrecy or to prevent other firms from using it (through legal protection, for example), while other firms have incentives to try to learn the technology, a process called imitation, and use it in their production. Obviously, guarding a possessed technology from its rivals and trying to copy an advanced technology are costly, but in the literature more attention is paid to the cost of imitation.

In the present context, we are interested in possible technology transfers between countries so that we simply assume that domestic protection of a new technology is perfect through perfect patent protection, for example. Once technology transfer between countries becomes the focus of analysis, several issues arise. The first one is the process and costs of imitating the technology in the advanced countries by the firms in the backward countries. The second issue is about the interactions between innovation and imitation. The third is the analysis of the product cycle theory, and the fourth one is the analysis of some government policies that directly affect the rate of innovation and/or the rate of imitation. These policies include research (either innovation or imitation) subsidies and intellectual property rights protection. These four issues are interrelated. We present a brief discussion about them.

The product cycle theory as suggested by Vernon (1966) provides a rigorous theory that postulates the invention and initial production of new products in countries such as the United States, and later the shift of production of these products to countries with lower wage rates. This paper provide many new ideas and observations, and despite the lack of a mathematical model, it refers to a dynamic environment in which new products continually emerge, and production continually shifts from the United States to less developed countries. In his model, he emphasizes the investment of the U.S. firms in less developed countries as the major vehicle of transferring the technology of producing new products from the United States to other countries.

Vernon's product cycle theory has been extended and formalized by many papers. Krugman's North-South model (Krugman, 1979) provides a rigorous model of innovation and imitation. He shows that in the steady state of the world, there is a constant gap between the number of products produced in the North and that in the South. In his model, the channel of technology transfer is not foreign direct investment but imitation. His model is later extended by Feenstra and Judd (1982), who examines several welfare and policy issues. These two models, however, consider only exogenous innovation.

Endogenizing innovation and imitation is a natural step in the endogenous growth literature. Segerstrom (1991), by extending the model of Grossman and Helpman (1991c), examines the interactions between endogenous innovation and imitation in a closed, growing economy. However, to examine the product cycle theory with endogenous growth, two countries are the lowest dimension of a suitable model. In a series of papers, Grossman and Helpman (1991b, 1991c, 1991d) investigates innovation, imitation, and product cycle, using several different models.[13]

A simple version of Grossman and Helpman (1991c) is now presented to illustrate how imitation in the presence of vertical innovation can be introduced. Suppose that there are two countries labelled North and South. North has a comparative advantage in innovation, while South has a lower wage rate. If both countries have the same access to technologies, products will be produced in the South only (at least in the short-run before wages adjust). Assuming Bertrand competition, three types of firms may exist in equilibrium: (i) Northern leaders (firms that can produce the state-of-the-art products) that are competing with another Northern firm that can produce the second-to-top quality; (ii) Northern leaders that are competing with a Southern firm that can produce the second-to-top quality; (iii) Southern firms that are able, via imitation, to produce the state-of-the-art products. In the presence of imitation threats, the Northern leaders have incentives to conduct research: to master the next generation technology as a safeguard against future imitation; to deter rival firms from targeting its product for imitation; and to try to gain a two-step advantage over its nearest rival.

Imitation is treated as a process similar to innovation in the sense that it is risky, and it requires resources. Southern firms choose products

13. Grossman and Helpman (1991b) assume horizontal innovation, while the other two papers consider vertical innovation. A survey of some of the results in these papers and some further extensions are given in Grossman and Helpman (1991a, Chapter 11).

to imitate. The probability of success of an imitation is represented by a Poisson distribution, with the arrival rate dependent on the amount of a resource (such as labor) that a firm chooses to conduct the research. In a steady state, the difference between the measures of products manufactured in the North and that in the South is zero, and the composition of Northern products remains constant.[14]

Two types of equilibria may arise. In the first type, leaders enjoy a large technological advantage over followers in research, and only the leaders engage in R&D. The equilibrium involves alternating phases of Northern and Southern production of each good. In the second type, followers are relatively efficient in innovation, and both the leaders and the followers engage in R&D. The path followed by any particular good can be complex, because it may pass from the leader to another Northern firm or to a Southern firm.

Note that because no learning by doing or human capital accumulation is assumed in the Grossman and Helpman model, the South conducts only imitation and is always behind the North in the technology race.

Another paper that models product cycles of products is Dinopoulos et al. (1993). They use the Heckscher-Ohlin framework and showed how differences in relative factor endowment may explain product cycles in the presence of factor price equalization. This is in contrast to Grossman and Helpman (1991c) where product cycles are due to lower wages in the South.

Grossman and Helpman (1991b) suggest a product-cycle model with horizontal innovation. The results are closer to what Vernon observed: New products are being invented in the North, which are later imitated by the South. With a wage advantage, the South eventually becomes the sole producer in the world.

Another issue related to technology transfer between two countries is trade-related intellectual property rights (TRIPs) protection. As mentioned earlier, every technology leader has an incentive to protect the secrecy of its technology knowledge while other firms (especially those in another country with other advantages such as lower wages) have incentives to imitate and produce a similar product. For a closed economy, imitation may be prevented by patent laws, but in a two-country model with a leader in one country and many potential imitators in another country, patent protection is less effective.

Helpman (1993), by extending the Krugman (1979) model of ex-

14. In an alternative setting, Segerstrom (1991) show that imitation by Northern firms is possible if firms collude by trigger strategies, rather than compete a la Bertrand.

ogenous innovation and the Grossman and Helpman (1991b) model of endogenous innovation, analyze the effects of intellectual property rights (IPRs) protection on the welfare of both countries, and the effects on innovation and imitation. IPRs protection is modeled as an increase in the cost of imitation by firms in the South (the backward country). He show that an IPRs protection hurts the South, but its effects on the welfare of the North and that of the world is ambiguous. Under certain conditions, the North benefits, but in some cases, both the North and the South are hurt by the protection. Helpman also examines the effects of the IPRs protection on the growth rate of innovation, and showed some cases in which the protection hurts, not helps, the Northern firms' innovation.

Taylor (1994) extends his previous paper of quality ladder to examine the implications of TRIPs. He shows that the failure to provide patent protection reduces R&D activities worldwide and slows growth. These results are different from those in Helpman. For an alternative formulation of the same issue, see Rivera-Batiz and Romer (1991b).[15]

6.5 Poverty Traps, Trade, and Growth

Development economists have argued that a poor country may remain poor forever, unless there is a big push to industrialize it. A poverty trap is a stable steady state with low per capita consumption. See Lewis (1954), Barro and Sala-i-Martin (1995, Chapter 1), Murphy, Schleifer, and Vishny (1989), and Azariadis and Drazen (1990).

Does trade create an opportunity to escape from the poverty trap? The answer is "yes" and "no," depending on the assumptions. Majumdar and Mitra (1995) assume that capital is the only factor of production that is mobile between two sectors, the consumption good sector and the investment good sector. In the former sector, marginal product of capital is constant. The production function of the latter sector exhibits increasing returns at low levels of capital, and diminishing returns beyond a certain threshold, with zero marginal product of capital in the limit. It follows that it is not possible for the closed economy to have positive growth forever. If the country is open to trade and the rest of the world has a better technology for the investment good sector, then growth becomes possible: the country can import the investment good and specialize increasingly in the production of the consumption

15. While these papers analyze the effectiveness of IPRs, an important question has not been raised or answered: Since a country (e.g., a less developed country) usually benefits from learning from advanced firms in another country, why would it be willing to protect the intellectual property rights (IPRs) of the technology leader in another country?

good. In fact, trade has effectively endowed the country with the AK technology with which it can indirectly produce the investment good.

Long, Nishimura, and Shimomura (1997) adopt the Heckscher-Ohlin framework, but allow for variable returns to scale of the S-shaped type. They show that there is a threshold level of non-consumable and non-depreciating capital stock, above which the country will choose to grow perpetually, thanks to a high marginal productivity of capital, like in other AK models. Below that threshold level, the country will run down its capital stock to zero, by selling its capital in exchange for the consumption good. This contrasts sharply with the autarkic case in which the capital stock is a positive constant in the long run. In the free trade case, the country will eventually specialize in one good, but during the transition phase, it may produce both goods. The country switches in and out of diversification by discrete jumps, because it is never efficient to produce a good on a small scale.

7. Growth and International Factor Mobility

In this section, we examine the roles of international factor movement in the neoclassical and endogenous growth models. We will first consider international capital movement, and then international labor migration.

In the trade literature, international factor mobility occupies an important part. However, previous work on the factor movement among countries usually assumes static frameworks with given factor endowments in countries, even though it is recognized that factor endowments may change over time due to investment and population growth. The assumption of given factor endowments is sometimes justified by the argument that only steady states are considered. In the endogenous growth literature, this argument may no longer be valid because the factor endowment ratios of countries may change along balanced growth paths.

In this section, we examine how the theory of international factor movement may change when growth is endogenous. We will pay more attention to several issues: how factor mobility may affect growth, how it may affect convergence of countries' growth rates, and how growth may affect international factor mobility.

7.1 International Capital Movement

We first begin with the neoclassical framework. As we showed earlier, the steady-state growth rate of an economy analyzed in a neoclassical model is given exogenously. With given technologies, the movement of capital therefore does not change the steady-state growth rates. In fact, if two countries are identical, they will have the same steady state with the same factor prices. This means that in a steady state, capital will not move.

However, if two countries have not reached their steady states, capital may move even if they have identical technology, as long as they have different initial capital-labor ratios. Capital may also affect the adjustment of the economies.

To see this point, consider figure 3. There are two countries, North and South. They have identical and fixed technology, but the North has a higher initial capital-labor ratio, $k_0^n > k_0^s$. These capital-labor ratios are lower than the countries' steady-state ratios.

If the two countries are isolated, then they will grow over time until the steady state is reached, as explained in section 3. Suppose now that international capital movement is allowed by both countries. For simplicity, we assume no risk and negligible moving costs. However, capital movement takes time so that any rental differential between the countries cannot be eliminated by capital movement instantaneously. The higher capital-labor ratio in the North implies a higher wage-rental ratio. Thus capital flows gradually from the North to the South. Let us denote the amount of capital that comes from the North to the South by Z, and its rate of change by $\dot{Z}$.

The presence of capital movement requires modification of the neoclassical model examined in previous sections. First, national income includes not just the domestic output but also the repatriation of national capital working abroad (or less the payment to foreign capital working locally). Second, the change in domestic capital stock comes not only from domestic investment but also from more foreign capital inflow (or less domestic capital outflow).

To analyze the adjustment of an economy, let us focus on the North for the time being. Its capital stock at any time grows over time according to

$$\dot{K}^n = sY^n - \delta K^n - Z,$$

where s is the saving rate, which is assumed to be a constant fraction of the domestic output Y^n (superscript n for the variables of the North and superscript s for those of South).[16] A similar equation holds for the South. The growth rates of the capital-labor ratio in the countries are

$$\widehat{k}^n = sy^n/k^n - n - \delta - \widehat{z}^n, \tag{47.1}$$

16. The assumption that saving is a constant fraction of the domestic output is made for convenience. A probably more realistic assumption is that it is a constant fraction of the national income or is chosen by either the government or individuals to maximize some objective functions.

$$\widehat{k}^s = sy^s/k^s - n - \delta + \widehat{z}^s, \tag{47.2}$$

where $\widehat{z}^n = \dot{Z}/K^n$ and $\widehat{z}^s = \dot{Z}/K^s$. For simplicity, the two countries are assumed to have the same saving rate.

The effect of international capital movement on the growth rates is illustrated in figure 3. Points N and S represent the initial points of the North and the South. Without capital movement, they adjust along schedule sy/k until the balanced-path point B is reached. In the absence of capital movement, the gap between schedule sy/k and line $n + \delta$ represents the speed of adjustment.

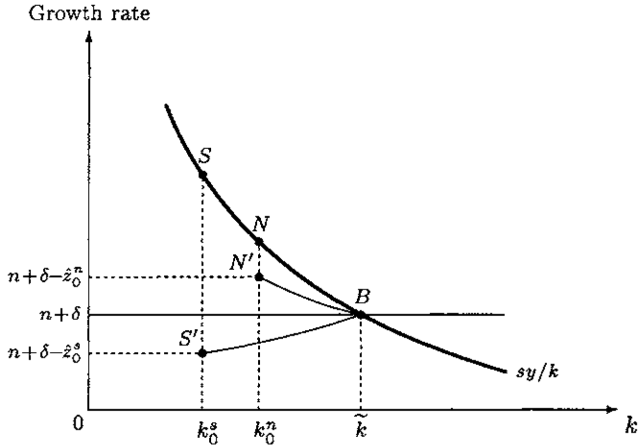


Fig. 3. Growth in the presence of international capital movement

When capital moves, construct schedule $n + \delta + \widehat{z}^n$ (shown as schedule $N'B$ in figure 3) for the North and schedule $n + \delta - \widehat{z}^s$ (shown as schedule $S'B$) for the South. If we assume that the growth rate of capital movement is an increasing function of the rental rate differential between the countries, both $\widehat{z}^n$ and $\widehat{z}^s$ decrease over time as more capital flows from the North to the South. In other words, schedules $n + \delta + \widehat{z}^n$ and $n + \delta - \widehat{z}^s$ converge and meet at point B, where capital movement ceases.

The speed of adjustment of the North depends on the gap between schedules sy/k and $n + \delta + \widehat{z}^n$ while that of the South depends on the gap between schedules sy/k and $n + \delta - \widehat{z}^s$. Thus international capital movement slows down the growth of the North and speeds up that of the South. This allows the South to catch up, and the growth rates of the countries converge faster.

How would the above conclusion be different if we have an endogenous growth model? Let us consider the Solow-Pitchford AK model. In this model, even though the growth rate of an economy is endogenously determined, international capital movement between two countries with identical and fixed technology has *no* effect on the steady-state growth rate of each country. The reason is that in a steady state, the rental rate is equal to A. In other words, there is no international capital movement in a steady state, and the growth rate of each country is given by (10″).

International capital movement in this model has the same positive effect on the rate of convergence of the countries' growth rate as it does in the neoclassical framework. It is because the growth rate of the countries' capital-labor ratios are still given by equations (47.1)–(47.2). Thus, the above analysis also applies to the AK model.

International capital mobility could lead to perpetual growth of an economy which, when closed, has no growth in the long run. This result was first established by Deardorff (1994). To see this point, consider two neoclassical economies with identical technology, North and South, with the North's exogenous and constant population growth rate being lower than that of the South. Suppose that the North is a small economy so that free capital mobility anchors the rental rate in North to that in the South, thus avoiding diminishing marginal product of capital in the North. If the saving in the North is high enough, then the North can grow perpetually. As long as the North's savings rate is not too high, the North can remain a small open economy for ever. If the North's saving is high enough, then it will sooner or later own a significant share of the world capital stock. Because it has a lower population growth rate, asymptotically the share of its labor force in the world drops to zero. In the long run, the capital-labor ratio in the world is constant, with the North owning a constant share of the world's capital, meaning that asymptotically both countries' capital stocks grow at the same rate as that of the South's population. Thus the North's capital-labor ratio is rising while that of the South is constant.

7.2 International Labor Migration

Although many papers on international migration consider only static models, there have been efforts to analyze migration in a dynamic context, especially in models in which education and training are explicitly examined. Some of the more important papers include Bhagwati and Hamada (1974), Rodriguez (1975), Miyagiwa (1991), Galor and Stark (1994), and Shea and Woodfield (1996). These papers determined the transformation of unskilled workers to skilled workers through education in the presence of international labor migration. However, because these papers assumed that the skill level of the skilled workers is fixed, knowledge does not accumulate. Thus growth of the economy is not sustained.

To see how international labor migration can be introduced into growth theory, let us begin with the neoclassical model we described in section 3. Consider again a one-sector closed economy with a Cobb-Douglas production function. It has been shown that without technological progress, the per capita output remains stationary in a steady state. The steady-state equilibrium is represented by equation (7).

Suppose now that the economy allows an inflow of foreign workers at a rate of m. Right after their arrival, foreign workers become permanent residents in the economy. For simplicity, assume that foreign workers do not bring physical capital with them, and that they have the same saving rate as the domestic residents.[17] With the inflow of foreign workers, the local population and thus the labor force grow at a rate of $n + m$. The new steady state equilibrium condition is

$$sy/k = n + m + \delta. \tag{48}$$

Differentiation of equation (48) shows that an increase in m decreases k and thus the local wage rate.[18]

As explained before, in the absence of technological progress, the per capita output of the economy remains stationary in a steady state. Therefore, its growth rate is not affected by labor inflow. International migration, however, does have effects on the convergence of the growth rates of two economies, when they are currently off their steady states.

Consider two economies labelled North and South. Suppose that they have the same technology that is stationary, the same depreciation rate, the same population growth rate, and the same saving rate. Thus, they have the same steady state.

Suppose that currently the capital-labor ratios of both countries are below their steady-state level, $\tilde{k}$, with the North, having a higher capital-labor ratio, i.e., $\tilde{k} > k_0^n > k_0^s$. If the economies are closed, both capital-labor ratios will move up over time until the steady state is reached.

17. These two assumptions can be relaxed easily. See Barro and Sala-i-Martin (1995, Chapter 9) for an analysis of cases in which foreign workers bring physical capital with them. Galor and Stark (1990) argue that foreign temporary workers, who are facing the possibility that they may leave soon, may save more.

18. While these effects of labor immigration are similar to those in a static model, a major difference should be noted. If there is a once-and-for-all inflow of foreign workers, as is assumed in a static model, there will be no effect on the steady-state capital-labor ratio and factor prices. The reason is that the steady-state equilibrium is still described by (7). The intuition is that as foreign workers come in, saving of the economy goes up until the steady-state capital-labor ratio climbs up back to its original level.

With $k_0^n > k_0^s$, the North has a higher wage rate, meaning that if migration is allowed, workers will move from the South to the North. Suppose that the rate of migration is m.[19] The vertical gap in figure 3 between the sy/k schedule and line $n+\delta+m$, shown as NN', represents the speed of increase in the North's capital-labor ratio, while SS' represents that of the South. The diagram shows that migration has slowed down the growth rate of the North but speeded up that of the South, allowing the latter to catch up faster.[20]

This model, though simple, does not imply a perpetual growth of the economies. A more interesting approach is to include human capital and permit endogenous accumulation of human capital. Galor and Stark (1994), and Shea and Woodfield (1996) are two recent attempts. The former paper, by considering an economy with multiple steady-state equilibria, presents cases in which admitting foreign workers who are slightly less skilled than the average native could move the economy to a steady state with a substantially lower human capital level. The latter paper derives the optimal immigration policy when skilled and unskilled workers come at the same time. The growth of the economies in these two papers, however, is not sustained, because in a steady state human capital does not accumulate. Barro and Sala-i-Martin (1995, Chapter 9) assume that a country can maintain a constant growth rate of migration, m (at least for a certain period of time). Then migration can have a growth effect.

Wong (1995, Chapter 14) considers three types of international labor migration – permanent migration, temporary migration, and brain drain – and discusses two channels through which human capital accumulates: learning by doing and education. His main concern is the choice between the three types of migration, but he does not examine explicitly the growth rates of the host and source countries.

An attempt to analyze the inter-relationship between international labor migration and growth rate of an emigration economy was given in Wong (1997). By extending the Uzawa-Lucas model of education and human capital accumulation, he analyzes how growth rate affects and is affected by each of the three types of migration. By allowing workers to choose the type of migration, i.e., when and where to work and to get education, he shows some cases in which permanent migration switches

19. The migration rate can be regulated exogenously by either government, or it may depend endogenously on the wage differential between the countries. This point is not crucial in the present analysis.

20. However, if the migration rate drops as the growth rates of the countries are getting closer to each other, the gap between their adjustment rates will decrease, too.

to temporary migration as the emigration economy grows. A deeper analysis of the case of brain drain was given by Wong and Yip (1996).

8. Concluding Remarks

In this chapter, we have surveyed the major models and issues of endogenous growth and international trade. We first described major endogenous growth models, and then turned to the literature of growth and trade.

Endogenizing and explaining growth of economies has become a major focus in the literature recently. The main feature of this literature, as explained in sections 4–5 above, is to link the growth of an economy with some of the features of economies such as preferences, technologies, and government policies. Several factors of growth have been outlined: accumulation of factors, external effects, learning by doing, education, and R&D.[21] This survey uses a unified model to present the main features of some of the endogenous growth models and their mathematical similarities.

It has been realized that even though most papers on endogenous growth were written in the past decade, there had already been papers in the sixties and early seventies that have dynamic models with growth rates endogenously determined by individuals or government policies. Even Solow mentioned the conditions for perpetual growth of economies.

It is thus interesting to ask why these "old" papers on perpetual growth did not generate the kind of interest in endogenous growth like what was experienced in the past decade.

Several reasons can be suggested. First, one major objective of the papers of Solow, Swan, and others in the fifties and sixties was to introduce production substitution possibilities in order to solve the instability problem in the Harrod-Domar growth models. The growth rate per se was not the main focus of the analysis, and these papers were by and large content with models that suggested a steady state with no perpetual growth for an economy.

In the past decade, however, the growth rates of countries were a much bigger issue. On the one hand, countries showed wide disparities in their growth rates. It is interesting to explain why many countries have different growth rates and whether these rates tend to converge over time. The neoclassical model of Solow and Swan is not the right tool because it implies that countries with identical and fixed technolo-

21. To the extent that government regulations may divert talents away from the R&D sector, the extent of regulations may also affect growth rate. See Goff (1996) and Berger (1996), for example.

gies and preferences will converge in terms of their growth rates until they reach the steady state with the same (exogenous) growth. On the other hand, people are interested in knowing the implications of different government policies on growth. Again the neoclassical model is not the appropriate tool, as long as long-run growth is concerned.

If we judge the recent endogenous growth literature by the three points of criticism on the neoclassical model mentioned in section 3, we can see that its biggest success is its endogenous determination of economies' growth rates. By providing different rigorous mathematical models, these papers highlighted several important factors that may affect the growth of economies. The more practical implication of these models was that the government has a role in economic growth.

Empirically, the endogenous growth models can easily be adopted to explain why countries do not have the same growth rates and why their growth rates do not converge. However, how much success these models really have in passing empirical tests is debatable. First, as we explained earlier, there is the uncomfortable implication in many of these models that the size of a country or an industry holds a paramount influence over the country's growth. This implication is not supported by both time-series and cross-country data. Second, it has been suggested that the Solow model with exogenous growth rates, when suitably augmented, can explain the growth rates of countries at least as well as some endogenous growth models do (Mankiw et al., 1992; Jones, 1995a, 1995b). Third, most of the endogenous models are based on some ad hoc assumptions about how human capital or technology accumulate, how growth is determined, and whether scale effects are present. In many cases, the results depend crucially on the range of a particular parameter: whether it is zero or positive, or whether it is greater than unity. Sometimes the functional form of a function is important. Fourth, most models on endogenous growth consider only an economy with one homogeneous final good. Mathematically, this assumption allows tractability of the algebra and simplifies the non-essential elements of the model in order to highlight different factors of growth. Empirically, this assumption could be misleading because it neglects structural changes, interactions between sectors, and different distributions of sectors in different countries. In particular, very little work has been done to examine the empirical relevance of some of the microfoundation equations of the models such as the R&D equation, the education equation, and so on. Fifth, despite the work on how R&D, education, learning by doing, factor accumulation and so on may affect growth, we still have little knowledge about why economics like Taiwan, Hong Kong, Singapore, and Korea grew so rapidly in the past several decades, while countries like the Philippines and India did not experience such growth.[22] Sixth, nearly all empirical work in the endogenous growth literature (for example, Young, 1994; Barro and Sala-i-Martin, 1995; Jones, 1995a, 1995b) used the growth rate of per capita income (or output) of countries as a measure of growth. However, we saw above that growth can be due to horizontal innovation (increase in the number of varieties) and vertical innovation (quality improvement of existing products). How important these factors of growth are in the growth experience of economies such as Hong Kong and Taiwan is unknown, but neglecting them in empirical studies could give misleading results.

Another issue that has become controversial is the convergence hypothesis. As explained earlier, several papers cited the persistent divergence in countries' growth rates and the lack of convergence of their growth rates as evidence that the neoclassical growth theory is inadequate. This view, however, has been challenged recently. For example, Barro and Sala-i-Martin (1992) observed convergence among the 48 states of the United States in terms of the growth rates of their per capita income and per capita gross state product. A similar convergence among the 47 Japanese prefectures has also been observed (Sala-i-Martin, 1996). However, convergence among different countries was less obvious (Barro and Sala-i-Martin, 1992, 1995). Several concepts of convergence have been introduced. First, it has been argued that the neoclassical growth theory implies only convergence (called conditional convergence) among those countries with the same economic structure (technologies, preferences, saving policies, and so on), not convergence (called absolute convergence) among all countries, possibly with different economic structures. Barro and Sala-i-Martin (1992) and Sala-i-Martin (1996) do observe conditional convergence.[23] Second, while many countries have persistent gaps between their growth rates, countries with similar economic structures seem to have their growth rates converging over time.[24] This is confirmed by Sala-i-Martin (1996). Furthermore, Quah (1996) and Galor (1996) argue that under certain conditions, countries

22. The literature on indeterminacy [for example, Xie (1994) and several other papers in the same *JET* issue] tells us that, starting from the same initial conditions, different countries can move along different paths with different growth rates, depending on agents' expectations about the future. This literature does not explain why expectations differ, and/or how they can be manipulated.

23. The conditional convergence hypothesis is supported in Barro and Sala-i-Martin (1992) and Sala-i-Martin (1996) only if the technology and preference parameters of the countries are assumed to depart substantially from the usual benchmark cases. For example, the capital share is required to be in the neighborhood of 0.8.

24. Sala-i-Martin (1996) calls this β-convergence in the sense that the ratio of the per capita income of the North to that of the South declines over time.

can show multiple steady states, and different countries with similar economic structures can converge to different steady states and thus different growth rates, a phenomenon called club convergence.[25]

The literature on trade and growth, with its diversity of results, suggests that no simple policy recommendations should be made without a thorough understanding of the structure and the key features of the economies under consideration. The results and the relationship between growth rates and international trade in general are sensitive to the structures of the economic models. The opening of trade can increase growth (Rivera-Batiz and Romer) or retard growth (Young). Moreover, faster growth may imply higher or lower welfare. The classical gains from trade theorem relies on the absence of externalities. Growth, on the other hand, is largely associated with dynamic spillovers and externality.

It is no doubt that the recent endogenous growth literature has improved our understanding of some of the factors that may affect countries' growth. Despite the voluminous literature in the past decade, however, there remain many unanswered questions.

Acknowledgement. Long's research is supported by SSHRC (Canada) and FCAR (Québec). This chapter has also been presented at the University of Konstanz, Lund University, and Academia Sinica. Thanks are due to Alan Deardorff, Elias Dinopoulos, Theo Eicher, Oded Galor, Bjarne S. Jensen, Murray C. Kemp, Hiroshi Ohta, Ping Wang, and participants at the Denmark conference and these seminars for helpful comments.

References

Aghion, Philippe, and Peter Howitt (1992). "A Model of Growth Through Creative Destruction," *Econometrica* 60: 323–351.

Arrow, Kenneth (1962). "The Economic Implications of Learning by Doing," *Review of Economic Studies* 29: 155–176.

Azariadis, Costas, and Allan Drazen (1990). "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105: 501–526.

Backus, David K., Patrick J. Kehoe, and Timothy J. Kehoe (1992). "In Search of Scale Effects in Trade and Growth," *Journal of Economic Theory* 58: 377–409.

25. In the neoclassical framework introduced in section 3, the steady state of a closed economy must be unique. However, Quah (1996) and Galor (1996) argued that multiple steady states may exist in the presence of overlapping generations, threshold externality, capital market imperfections, heterogeneity, country size, or club formation. See the papers cited in these two papers for more details.

Barro, Robert (1990). "Government Spending in a Simple Model of Endogenous Growth," *Journal of Political Economy* 98: S103–S125.

— and Xavier Sala-i-Martin (1992). "Convergence," *Journal of Political Economy* 100: 223–251.

— and Xavier Sala-i-Martin (1995). *Economic Growth.* New York: McGraw-Hill.

Berger, H. (1996). "Regulation in Germany: Some Stylized Facts about Its Time Path, Causes and Consequences." Mimeo, University of Munich.

Bhagwati, Jagdish N., and Koichi Hamada (1974). "The Brain Drain, International Integration of Markets for Professionals and Unemployment: A Theoretical Analysis," *Journal of Development Economics* 1: 19–24.

Bond, Eric W., and Kathleen Trask (1997). "Trade and Growth with Endogenous Human and Physical Capital Accumulation." (this volume).

Bond, Eric W., Ping Wang, and Chong K. Yip (1996). "A General Two-Sector Model of Endogenous Growth with Human and Physical Capital: Balanced Growth and Transitional Dynamics," *Journal of Economic Theory* 68: 149–173.

Caballé, Jordi, and Manuel S. Santos (1993). "On Endogenous Growth with Physical and Human Capital," *Journal of Political Economy* 101: 1042–1067.

Chipman, John S. (1970). "Induced Technical Change and Patterns of International Trade." In Raymond Vernon (ed.), *The Technology Factor in International Trade.* New York: National Bureau of Economic Research, 95–127.

Clemhout, Simone, and Henry Y. Wan, Jr. (1970). "Learning-by-Doing and Infant Industry Protection," *Review of Economic Studies* 37: 33–56.

D'Autune, Antoine, and Phillipe Michel (1994). "Politique Economique, Recherche et Developement et Croissance." Presented at *Journees,* AFSE.

Deardorff, Alan (1994). "Growth and International Investment with Diverging Population," *Oxford Economic Papers* 46: 477–491.

Devereux, Michael, and Beverley Lapham (1994). "The Stability of Economic Integration and Endogenous Growth," *Quarterly Journal of Economics* 109: 299–305.

Dinopoulos, Elias, James F. Ochmke, and Paul S. Segerstrom (1993). "High-Technology-Industry Trade and Investment," *Journal of International Economics* 34: 49–71.

Dixit, Avinash, and Joseph E. Stiglitz (1977). "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67: 297–308.

Dolmas, Jim (1996). "Endogenous Growth in Multi-sector Ramsey Models," *International Economic Review* 37: 403–421.

Domar, E. (1946). "Capital Expansions, Rate of Growth and Employment," *Econometrica* 14: 137–147.

Eicher, Theo (1996). "Interaction Between Endogenous Human Capital and Technological Change," *Review of Economic Studies* 63: 127–144.

— and Stephen J. Turnovsky (1996). "Toward a General Model of Economic Growth." Mimeo, University of Washington.

Ethier, Wilfred (1982). "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review* 72: 389–405.

Feenstra, Robert C. (1996). "Trade and Uneven Growth," *Journal of Development Economics* 49: 229–256.

— and Kenneth L. Judd (1982). "Tariffs, Technology Transfer, and Welfare," *Journal of Political Economy* 90: 1142–1165.

Findlay, Ronald (1970). "Factor Proportion and Comparative Advantage in the Long Run," *Journal of Political Economy* 78: 27–34.

— (1974). "Relative Price, Growth and Trade in a Simple Ricardian Model," *Economica* 41: 1–13.

— (1978). "Relative Backwardness, Direct Foreign Investment, and the Transfer of Technology: A Simple Dynamic Model," *Quarterly Journal of Economics* 92: 1–16.

— (1984). "Growth and Development in Trade Models." In R.W. Jones and P.B. Kenen (eds.), *Handbook of International Economics*, Volume I. Amsterdam: North-Holland, 185–236.

— and Henryk Kierzkowski (1983). "International Trade and Human Capital: A Simple General Equilibrium Model," *Journal of Political Economy* 91: 957–978.

Fisher, Eric O'N. (1995). "Growth, Trade and International Transfers," *Journal of International Economics* 39: 143–158.

— and Neil Vousden (1995). "Dynamic Trade Creation." Unpublished typescript.

Galor, Oded (1986). "Time Preference and International Labor Migration," *Journal of Economic Theory* 38: 1–20.

— (1994). "Tariff, Income Distribution, and Welfare in a Small Overlapping-Generations Economy," *International Economic Review* 35: 173–192.

— (1996). "Convergence? Inferences from Theoretical Models," *Economic Journal* 106: 1056–1069.

— and Oded Stark (1990). "Migrants' Savings, the Probability of Return Migration and Migrants' Performance," *International Economic Review* 31: 463–467.

— and Oded Stark (1994). "Migration, Human Capital Formation, and Long-Run Output." In Horst Siebert (ed.), *Migration: A Challenge for Europe.* Tübingen: J.C.B. Mohr, 59–68.

Goff, Brian (1996). *Regulation and Macroeconomic Performance.* Boston, Mass.: Kluwer Academic Publisher.

Grossman, Gene M., and Elhanan Helpman (1990a). "Trade, Innovation, and Growth," *American Economic Review* 80: 86–91.

Grossman, Gene M., and Elhanan Helpman (1990b). "Comparative Advantage and Long Run Growth," *American Economic Review* 80: 796–815.

Grossman, Gene M., and Elhanan Helpman (1991a). *Innovation and Growth in the Global Economy.* Cambridge, Mass.: MIT Press.

Grossman, Gene M., and Elhanan Helpman (1991b). "Endogenous Product Cycles," *Economic Journal* 101: 1241–1229.

Grossman, Gene M., and Elhanan Helpman (1991c). "Quality Ladder and Product Cycles," *Quarterly Journal of Economics* 106: 557–586.

Grossman, Gene M., and Elhanan Helpman (1991d). "Quality Ladder in the Theory of Growth," *Review of Economic Studies* 58: 43–61.

Grossman, Gene M., and Elhanan Helpman (1991e). "Growth and Welfare in a Small Open Economy." In E. Helpman and A. Razin (eds.), *International Trade and Trade Policy.* Cambridge, Mass.: MIT Press.

Grossman, Gene M., and Elhanan Helpman (1993). "Hysteresis in the Trade Pattern." In W. Ethier, E. Helpman, and P. Neary (eds.), *Theory, Policy and Dynamics in International Trade.* Cambridge University Press, 268–290.

Grossman, Gene M., and Elhanan Helpman (1995). "Technology and Trade." In G. Grossman and K. Rogoff (eds.), *Handbook of International Economics,* Volume III. Amsterdam: Elsevier Science.

Harrison, Ann E. (1994). "Productivity, Imperfect Competition and Trade Reform," *Journal of International Economics* 36: 53–73.

Harrod, R.F. (1939). "An Essay in Dynamics Theory," *Economic Journal* 49: 14–33.

Helpman, Elhanan (1993). "Innovation, Imitation, and Intellectual Property Rights," *Econometrica* 61: 1247–1280.

Hu, Sheng Cheng (1976). "Education and Economic Growth," *Review of Economic Studies* 43: 509–518.

Ishikawa, Jota (1992). "Learning by Doing, Changes in Industrial Structure and Trade Patterns, and Economic Growth in a Small Open

Economy," *Journal of International Economics* 33: 221–244.

Jensen, Bjarne S. (1994). *The Dynamic Systems of Basic Economic Growth Models.* Dordrecht: Kluwer Academic Publishers.

— and Mogens E. Larsen (1987). "Growth and Long-Run Stability," *Acta Applicandae Mathematicae* 9: 219–237. (reprinted in Jensen (1994)).

— and Chunyan Wang (1997). "General Equilibrium Dynamics of Basic Trade Models for Growing Economies." (this volume).

Jensen, Richard, and Marie Thursby (1986). "A Strategic Approach to Product Life Cycle," *Journal of International Economics* 21: 269–284.

Jones, Charles I. (1995a). "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics* 110: 495–525.

— (1995b). "R&D-Based Models of Economic Growth," *Journal of Political Economy* 103: 759–784.

Jones, Larry E., and Rodolfo Manuelli (1990). "A Convex Model of Equilibrium Growth," *Journal of Political Economy* 98: 1008–1037.

Kaldor, Nicholas, and James A. Mirrlees (1961–1962). "Growth Model with Induced Technical Progress," *Review of Economic Studies* 29: 174–190.

Kemp, Murray C. (1974). "Learning by Doing: Formal Tests for Intervention in An Open Economy," *Keio Economic Studies* 11: 1–7.

Krugman, Paul R. (1979). "A Model of Innovation, Technology Transfer, and the World Distribution of Income," *Journal of Political Economy* 87: 253–265.

— (1984). "Import Protection as Export Promotion: International Competition in the Presence of Oligopoly and Economies of Scale." In Henryk Kierzkowski (ed.), *Monopolistic Competition and International Trade.* Oxford: Clarendon Press, 180–193.

Levhari, David (1966a). "Extensions of Arrow's Learning by Doing," *Review of Economic Studies* 33: 117–131.

— (1966b). "Further Implications of Learning by Doing," *Review of Economic Studies* 33: 31–38.

Lewis, William Arthur (1954). "Economic Development with Unlimited Supplies of Labour," *Manchester School of Economics and Social Studies* 22: 139–191.

Long, Ngo Van, Kazuo Nishimura, and Koji Shimomura (1997). "Endogenous Growth, Trade, and Specialization Under Variable Returns to Scale." (this volume).

Lucas, Robert E. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22: 3–42.

— (1993). "Making a Miracle," *Econometrica* 61: 251–272.

Majumdar, Mukul, and Tapan Mitra (1995). "Patterns of Trade and Growth Under Increasing Returns: Escape From the Poverty Trap," *Japanese Economic Review* 46: 207–225.

Mankiw, N. Gregory, David Romer, and David N. Weil (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107: 407–437.

Manning, R. (1975). "Optimal Aggregative Development of a Skilled Workforce," *Quarterly Journal of Economics* 89: 504–511.

— (1976). "Issues in Optimal Educational Policy in the Context of Balanced Growth," *Journal of Economic Theory* 13: 380–395.

Miyagiwa, Kaz (1991). "Scale Economies in Education and the Brain Drain Problem," *International Economic Review* 32: 743–759.

Murphy, Kevin M., Andrei Schleifer, and Robert W. Vishny (1989). "Industrialization and the Big Push," *Journal of Political Economy* 97: 1003–1026.

Ohyama, Michihiro (1991). "Human Capital and Endogenous Economic Growth," *Keio Economic Studies* 28: 1–14.

Oniki, H., and Hirofumi Uzawa (1960). "Patterns of Trade and Investment in a Dynamic Model of International Trade," *Review of Economic Studies* 32: 15–38.

Pitchford, John D. (1960). "Growth and the Elasticity of Factor Substitution," *Economic Record* 36: 491–504.

Quah, Danny (1993). "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics*, 427–443.

— (1996). "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal* 106: 104–1055.

Ramsey, Frank P. (1928). "A Mathematical Theory of Savings," *Economic Journal* 38: 543–559.

Razin, Assaf (1972a). "Optimum Investment in Human Capital," *Review of Economic Studies* 39: 455–460.

— (1972b). "Investment in Human Capital and Economic Growth," *Metroeconomica* 24: 101–116.

Rebelo, Sergio (1991). "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy* 99: 500–521.

Rivera-Batiz, Luis A., and Paul M. Romer (1991a). "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics* 106: 530–555.

Rivera-Batiz, Luis A., and Paul M. Romer (1991b). "International Trade with Endogenous Technological Change," *European Economic Review* 35: 971–1004.

Rodriguez, Carlos A. (1975b). "Brain Drain and Economic Growth: A Dynamic Model," *Journal of Development Economics* 2: 223–247.

Romer, Paul M. (1986). "Increasing Returns and Long-Run Growth," *Journal of Political Economy* 94: 1002–1037.

— (1990). "Endogenous Technological Change," *Journal of Political Economy* 98: S71–S101.

— (1994). "The Origins of Endogenous Growth," *Journal of Economic Perspectives* 8: 3–22.

Saint-Paul, Gilles (1992). "Fiscal Policy in An Endogenous Growth Model," *Quarterly Journal of Economics* 108: 1243–1259.

Sala-i-Martin, Xavier (1996). "The Classical Approach to Convergence Analysis," *Economic Journal* 106: 1019–1036.

Segerstrom, Paul S. (1991). "Innovation, Imitation, and Economic Growth," *Journal of Political Economy* 99: 807–827.

— (1995). "Endogenous Growth without Scale Effects." Mimeo, Michigan State University.

— , T.C.A. Anant, and Elias Dinopoulos (1990). "A Schumpeterian Model of Product Life Cycle," *American Economic Review* 80: 1077–1092.

Shea, K.-L., and A.E. Woodfield (1996). "Optimal Immigration, Education and Growth in the Long Run," *Journal of International Economics* 40: 495–506.

Sheshinski, Eytan (1967). "Optimal Accumulation with Learning by Doing." In K. Shell (ed.), *Essays on the Theory of Optimal Growth.* Cambridge, Mass.: MIT Press.

Smith, Alasdair M. (1984). "Capital Theory and Trade Theory." In R.W. Jones and P.B. Kenen (eds.), *Handbook of International Economics,* Volume I. North-Holland, Amsterdam: 289–324.

Solow, Robert M. (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70: 65–94.

— (1994). "Perspectives on Growth Theory," *Journal of Economic Perspectives* 8: 45–54.

Spence, Michael (1976). "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies* 43: 217–236.

Stokey, Nancy L. (1988). "Learning by Doing and the Introduction of New Goods," *Journal of Political Economy* 96: 701–717.

— (1991). "Human Capital, Product Quality, and Growth," *Quarterly Journal of Economics* 106: 587–616.

Swan, T.W. (1956). "Economic Growth and Capital Accumulation," *Economic Record* 32: 334–361.

Taylor, M. Scott (1993). "Quality Ladders and Ricardian Trade," *Journal of International Economics* 34: 225–243.

— (1994). "TRIPs, Trade and Growth," *International Economic Review* 35: 361–381.

Turnovsky, Stephen J. (1995). *Methods of Macroeconomic Dynamics.* Cambridge, Mass.: MIT Press.
— (1997). "Public and Private Capital in an Endogenously Growing Open Economy." (this volume).
Uzawa, Hirofumi (1965). "Optimal Technical Change in an Aggregative Model of Economic Growth," *International Economic Review* 6: 18–31.
— (1969). "Time Preference and the Penrose Effect in a Two-Class Model of Economic Growth," *Journal of Political Economy* 77: 628–652.
Van, Pham Hoang, and Henry Y. Wan, Jr. (1997). "Interpreting East Asian Growth." (this volume).
Vernon, Raymond (1966). "International Investment and International Trade in the Product Cycle," *Quarterly Journal of Economics* 83: 190–207.
Wälde, Klaus (1996). "Proof of Global Stability, Transitional Dynamics and International Capital Flows in a Two-Country Model of Innovation and Growth," *Journal of Economics* (forthcoming).
Wan, Henry Y., Jr. (1975). "Trade, Development and Inventions – The Evolutionary Path," Cornell University. Revised version appears in Horst Herberg and Ngo Van Long (eds.), *Trade, Welfare, and Economic Policies.* Ann Arbor, Michigan: University of Michigan Press, 1993, 239–254.
Wang, Jian-Ye (1990). "Growth, Technology Transfer, and the Long-Run Theory of International Capital Movement," *Journal of International Economics* 29: 255–271.
Wong, Kar-yiu (1995). *International Trade in Goods and Factor Mobility.* Cambridge, Mass.: MIT Press.
— (1997). "Endogenous Growth and International Labor Migration: The Case of A Small, Emigration Country." (this volume).
— and Chong K. Yip (1996), "Education, Economic Growth and Brain Drain." Mimeo.
— and Chong K. Yip (1997). "Industrialization, Economic Growth and International Trade." Mimeo.
Xie, Danyang (1991). "Increasing Returns and Increasing Rates of Growth," *Journal of Political Economy* 99: 429–435.
— (1994). "Divergence in Economic Performance: Transitional Dynamics with Multiple Equilibria," *Journal of Economic Theory* 63: 97–112.
Young, Alwyn (1991). "Learning by Doing and the Dynamic Effects of International Trade," *Quarterly Journal of Economics* 106: 369–405.
— (1993). "Invention and Bounded Learning by Doing," *Journal of*

Political Economy 101: 443–472.
— (1994). "Lessons from the East Asian NICs: A Contrarian View," *European Economic Review* 38: 964–973.
— (1995). "Growth without Scale Effects," NBER working paper No. 5211.

Part II

Dynamics of Basic Trade Models

CHAPTER 3

General Equilibrium Dynamics of Basic Trade Models for Growing Economies

Bjarne S. Jensen and Chunyan Wang

1. Introduction

In the literature on the pure theory of trade, the two-factor, two-sector, two-country framework has provided a fundamental general equilibrium structure for static and comparative-static analyses of many issues related to factor allocation, output composition, relative prices, and trade patterns. Although the work on two-sector growth models has long ago been extended to trading economies, *stability* issues and *non-steady-state dynamics* have to be further analyzed. As is well known, many results and theorems of both static and dynamic trade theory rest on the *assumption* of *incomplete* specialization. A major purpose of this chapter is to study the *conditions* that will in the *long-run* preserve the *diversification* of a *small* trading economy and two *large* trading economies. However, when the parametric conditions of diversified steady state growth are not satisfied, we give, for a small country, special attention to various forms of *endogenous* (persistent) growth per capita. Basic (prototype) dynamic trade models will be analyzed in detail.

A small trading economy, owing to the given terms of trade, can be conceived of as being restricted (by the outside world and competitive pricing) to operating with fixed-coefficient technologies. Nevertheless, the long-run stability of the *diversification* of a *small* trading country does *not* depend entirely, as in a closed economy, on the *ranking* of sectorial *factor* intensities. Alternative *trade patterns* give the growing two-sector economy some opportunities to remain diversified. However, besides *technology*, the *domestic demand* composition is of critical importance for preserving *domestic production* of both tradable goods. For two large trading countries, the long-run stability of incomplete specialization is likewise critically dependent on demand side parameters.

Dynamic trade theory was initiated by Oniki and Uzawa (1965), who studied the effects of capital accumulation and labor growth on international equilibrium over time for two large countries. In this area, other studies were Bardhan (1966, 1970), Kemp (1969), Findlay (1970), Takayama (1972), Woodland (1982), Gandolfo (1994). Our work may especially be seen as extensions of the contributions of Stiglitz (1970), Deardorff (1971, 1973, 1974, 1978, 1994), and Smith (1976, 1977, 1984).

In section 1, we present the general equilibrium structure of the trade models. Section 2 is devoted to dynamic two-sector models for *small* trading economies. As to capital accumulation, we consider *proportional*, *classical*, and *optimal* saving models. Section 3 presents a dynamic analysis of a two-sector growth model for two *large* countries with the *terms of trade* of international equilibrium *endogenously* determined. It may be called a two-factor Ricardian trade model, as we also allow for different sector technologies in the two countries. The *parametric conditions* of preserving *diversification* in both countries are obtained. Final comments are offered in section 4.

2. Structure of Two-Sector Trade Models

The structure of the two-sector trade models are formed by the *basic assumptions* of international immobile production factors, full employment, competitive prices, and trade balance equilibrium. The elements of a *competitive* two-sector economy with *homogenous* production functions will subsequently determine and impose important *restrictions* upon the character and *parameters* of the actual *homogenous dynamic* systems for a small or large economy trading in both goods.

2.1 Domestic Production and Factor Endowments

Consider an economy consisting of a *capital* good industry (sector) and a *consumer* good industry (sector), labelled 1 and 2, respectively. The two-sector general equilibrium model is characterized by the following assumptions.

The *sector technologies* are described by production functions exhibiting *constant returns* to scale,

$$Y_i = F_i(L_i, K_i) = L_i f_i(k_i) = L_i y_i, \quad i = 1, 2, \tag{1}$$

where $f_i(k_i)$, $i = 1, 2$, have the properties

$$\forall k_i > 0: \ f_i'(k_i) = df_i(k_i)/dk_i > 0, \ f_i''(k_i) = d^2 f_i(k_i)/dk_i^2 < 0, \tag{2}$$

$$\lim_{k_i \to 0} f_i'(k_i) = \overline{\beta} \le \infty, \ \lim_{k_i \to \infty} f_i'(k_i) = \underline{\beta} \ge 0, \ f_i'(k_i) \in J = [\underline{\beta}, \overline{\beta}]. \tag{3}$$

Thus, the *intensive function* f is a strictly concave monotonic increasing function on the nonnegative real line, with its slope decreasing from $\overline{\beta}$ at $k = 0$ to $\underline{\beta}$ at $k = +\infty$.

The *allocation ratios* (sector fractions) of labor are

$$L_1/L + L_2/L \equiv \mathrm{l}_1 + \mathrm{l}_2 \equiv 1. \tag{4}$$

Then, the full employment condition may be rewritten as, cf. (4),

$$k \equiv \mathrm{l}_1 k_1 + \mathrm{l}_2 k_2 \equiv k_2 + (k_1 - k_2)\mathrm{l}_1. \tag{5}$$

Thus, the *overall* capital-labor ratio k is the weighted average of the *sectorial* capital-labor ratios, k_1 and k_2, with the labor allocation ratios as weights. We note that, cf. (4)–(5)

$$\mathrm{l}_1 = \frac{k - k_2}{k_1 - k_2}, \qquad \mathrm{l}_2 = \frac{k_1 - k}{k_1 - k_2}. \tag{6}$$

The factor endowments belonging to the *diversification cone* $C_k \subset \mathbf{R}^2_+$ are:

$$C_k = \{(L, K) \in \mathbf{R}^2_+ \mid k_1 < K/L < k_2 \ \vee \ k_2 < K/L < k_1\}. \tag{7}$$

The total *domestic* (and per capita, Y_i/L) *production* of the two goods is, cf. (4),

$$Y_i = L_i y_i = L y_i \mathrm{l}_i, \quad i = 1, 2. \tag{8}$$

2.2 Prices, Incomes, Savings, and Trade Balance

The open two-sector economy is assumed to operate under *perfect competition (zero profit condition)*; absolute (money) *factor* prices (w, r) are the same in both sectors, *output* prices (P_1, P_2) represent unit cost, and *revenue* (total *cost*) is *shared* $(\epsilon_{L_i}, \epsilon_{K_i})$ between the factors.
Hence, we have the competitive *general equilibrium* relations, $i = 1, 2$,

$$P_i Y_i = rK_i + wL_i, \quad \epsilon_{K_i} = \frac{rK_i}{P_i Y_i}, \quad \epsilon_{K_i} + \epsilon_{L_i} = 1, \tag{9}$$

$$\epsilon_{K_i} = \frac{k_i}{k_i + \omega} = \frac{k_i f_i'(k_i)}{y_i}, \ \epsilon_{L_i} = \frac{\omega}{k_i + \omega}, \ \omega = \frac{w}{r} = \frac{f_i(k_i)}{f_i'(k_i)} - k_i, \tag{10}$$

$$P_i = \frac{w}{y_i \epsilon_{L_i}} = \frac{rk_i}{y_i \epsilon_{K_i}}, \quad \frac{r}{P_i} = \frac{y_i}{k_i + \omega} = f_i'(k_i), \quad \frac{w}{P_i} = \frac{y_i \omega}{k_i + \omega}, \tag{11}$$

$$p = \frac{P_1}{P_2} = \frac{y_2}{y_1}\frac{\epsilon_{L_2}}{\epsilon_{L_1}} = \frac{y_2(k_1 + w/r)}{y_1(k_2 + w/r)} = \frac{f_2'(k_2)}{f_1'(k_1)}, \quad P_i \neq 0. \tag{12}$$

The *common* wage-rental ratio, ω, becomes by (12),

$$\omega = \frac{w}{r} = \frac{k_1(y_2/y_1) - k_2(P_1/P_2)}{P_1/P_2 - y_2/y_1} = \frac{k_2(\epsilon_{K_1}/\epsilon_{K_2} - 1)}{1 - (y_2/y_1)(1/p)}. \tag{13}$$

An *open* competitive two-sector economy, trading at international prices determined in the world market, can only remain incompletely specialized (*diversified*, produce both goods, have a common *positive* ω), if the *range* of the *terms* of *trade*, $p = P_1/P_2$, is confined to the zero profit *price interval* with the limits, as seen from (13); cf. Rybczynski lines, Wong (1995)

$$k_1 > k_2: \quad \frac{y_2}{y_1} < p < \frac{y_2}{y_1}\frac{k_1}{k_2} = \frac{Y_2/K_2}{Y_1/K_1}, \tag{14}$$

$$k_2 > k_1: \quad \frac{y_2}{y_1}\frac{k_1}{k_2} < p < \frac{y_2}{y_1}. \tag{15}$$

When p is given as a *fixed* number, we use – for factor endowments within the diversification cone C_k, (7) – the symbols $\tilde{y}_i = f_i(\tilde{k}_i[\tilde{\omega}(p)])$, $i = 1, 2$, in (14)–(15), cf. figure 1.

National income (product), Y, is the monetary *value* of outputs from both sectors and represents aggregated factor incomes, cf. (8)–(9)

$$\begin{aligned} Y &= P_1Y_1 + P_2Y_2 = L(P_1y_1l_1 + P_2y_2l_2) \\ &= rK + wL = L(rk + w) = Ly. \end{aligned} \tag{16}$$

Proposition 1. *(Deardorff). With given prices (P_1, P_2) and the monotonicity and concavity conditions (2)–(3), the per capita revenue (GNP) function, $y(k)$, (16), is a concave C^1-class function on $[0, \infty[$, and $y(k)$ has a linear segment (flat) in the diversification cone C_k (7), (14)–(15).*

Proof. The proposition is proved and geometrically illustrated as a convex envelope theorem in Deardorff (1971, pp. 10; 1974, p. 297). □

The factor *income distribution* is defined by

$$\delta_K = \frac{rK}{Y} = \frac{k}{\omega + k}, \quad \delta_L = \frac{wL}{Y} = \frac{\omega}{\omega + k}, \quad \delta_K + \delta_L = 1. \tag{17}$$

Let Q_i, $i = 1, 2$, denote the quantitative size of the *domestic demand* (absorption level) for good 1 (investment) and good 2 (consumption), and they are respectively equal to domestic production, Y_i, minus net exports, $X_i \lesseqgtr 0$, i.e.,

$$Q_1 = Y_1 - X_1, \quad Q_2 = Y_2 - X_2. \tag{18}$$

The *trade balance* is assumed to satisfy the constraint

$$P_1X_1 + P_2X_2 = 0, \quad \text{i.e.} \quad Y = P_1Q_1 + P_2Q_2, \tag{19}$$

i.e., trade equilibrium prevails with no foreign borrowing/lending allowed.

The export of goods by one country, A, is import for another country, B (rest of world), i.e.,

$$X_{iA} = -X_{iB}, \qquad i = 1, 2. \tag{20}$$

Policy analyses based on open economy models, relaxing (19) with current account transmissions, have often just a single good, and in case of two goods, only one is traded, Obstfeld (1982), or nonspecialization in production is assumed, see Obstfeld (1989). A purely aggregative set-up in which every country produces the same, single good (and a rich country just producing more of it) can be useful in analyzing economies, interacting in a world of international trade, see Lucas (1993). Processes of factor accumulation, including growth of human capital, and international convergence issues are often in focus.

One of our main objectives, however, is to *merge* "old" and "new" *growth* theory with international *trade* theory. Accordingly, for our dynamic analysis below, with ***diversification***, ***specialization***, and ***trade patterns*** as ***endogenous*** issues, we need at least *two tradeable* goods. International trade in goods produced with labor or human capital as the sole factor of production is not considered, cf. Young (1991).

As to the division of *income* between *consumption* and *saving*, we shall first employ very simple aggregate saving functions, viz., *proportional* saving and *classical* saving that have been the standard polar opposites in much of both the growth and trade literature. Hence we alternatively use the *monetary saving* functions, cf. (16)–(17),

$$S = sY = P_1 Q_1, \qquad 0 < s < 1, \tag{21}$$

$$S = s_K \delta_K Y = s_K r K = P_1 Q_1, \qquad 0 < s_K \leq 1. \tag{22}$$

The Ramsey approach to consumer optimization over time (*optimal intertemporal saving*) has a prominent role in the literature on *two-sector optimal* growth models of *closed* economies, cf. Uzawa (1964), Srinivasan (1964), Cass (1965), Wan (1971), Drabicki and Takayama (1975), and in *open one-sector optimal* growth models with *financial* asset trading, see Becker and Foias (1987), Obstfeld and Rogoff (1995). We shall employ *Ramsey saving* to a small two-sector economy *trading* in *both* goods.

International trade is the difference between domestic production and domestic demand. Combining (21), (19), and (16) gives the export value of capital goods as

$$P_1 X_1 = (1 - s) P_1 Y_1 - s P_2 Y_2. \tag{23}$$

Hence, per capita *export* ("excess supply") of capital goods, X_1/L, positive or negative, may be written as, cf. (23), (8), (6),

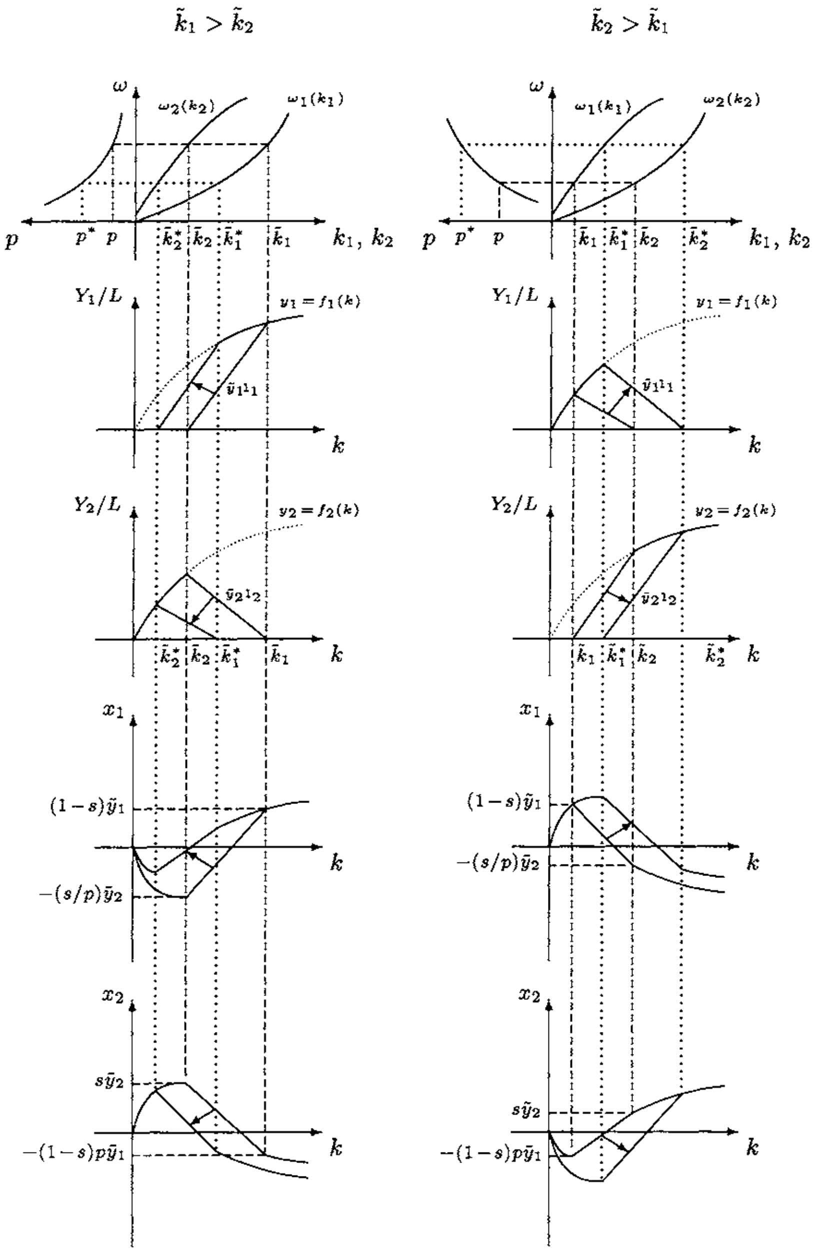


Fig. 1. Wage-rental ratio, factor allocation, sector outputs, and trade patterns with proportional saving and given terms of trade

$$\begin{aligned}
x_1 &= X_1/L = (1-s)Y_1/L - s(P_2/P_1)Y_2/L \\
&= (1-s)y_1 l_1 - (s/p)y_2 l_2 \\
&= \tfrac{1}{k_1-k_2}\big[-\{(s/p)y_2 k_1 + (1-s)y_1 k_2\} \\
&\qquad + \{(s/p)y_2 + (1-s)y_1\}k\big],
\end{aligned} \tag{24}$$

and per capita export of the consumer good, $x_2 = -px_1$, positive or negative, cf. (19), (24).

Within the *diversification cone*, figure 1 summarizes diagrammatically the core of the *Heckscher-Ohlin model*, which "identifies a *mapping* from exogenously given factor supplies and exogenously given external product prices (determined in the international market place) into internal factor prices, output levels and consumption levels, the difference between the last two items being international trade", Leamer and Levinsohn (1995, p. 1345). By using homogenous production functions of degree one and hence representing the factor supplies by the capital-labor ratio, k, this mapping and its various derivatives (theorems) are all represented by the shape of the respective curves within the diversification cone of figure 1. To obtain a coherent structural description of a small trading economy, the curves in figure 1 are extended to any factor endowments and are shown for two values of $p = P_1/P_2$.

Moreover, the curves Y_i/L, x_i, $k \in [0, \infty[$, in figure 1 will become trajectories related to the *state variable* $k(t)$ of our *dynamic* systems of factor accumulation. It is evident that these trajectories are not closed curves. Accordingly, besides the determination of the *speed* (rapidity) of motion (growth), a dynamic analysis must address the issue of whether $k(t)$ moves toward some *interior critical point* ("long-run equilibrium") within the *diversification* cone or towards the *endpoints* and then into regions with complete *specialization.* It is the possibilities of passing into or out of complete specialization that has always complicated the dynamic analysis of growing trading economies.

3. General Equilibrium Dynamics of a Small Trading Country

The momentary (timeless) general equilibrium relationships between *exogenous* factor endowment variations and domestic production and trade patterns were described above. The actual size or evolution of the factor endowments, however, was not explained. Moreover, international trade itself will through time have feed-back effects upon the available factor endowments. For *endogenous* factor endowment variations, proper dynamic laws governing the process of factor accumulation are needed to supplement the momentary general equilibrium equations.

3.1 Dynamics by Proportional Saving and Exogenous Labor Growth

3.1.1 Dynamic Solutions, Stability, and Parameter Regions

As to a dynamic model of a small trading economy, we must now perform a rigorous analysis of some specifications of factor accumulation. The domestic labor force of the small trading economy will by convention first be assumed to grow exponentially at an exogenous rate, $n > 0$, that is,

$$\dot{L} \equiv dL/dt = dL_1/dt + dL_2/dt = \dot{L}_1 + \dot{L}_2 = Ln \equiv L\,\mathfrak{f}(k). \tag{25}$$

The domestic stock of capital goods increases by savings (investment demand). Depreciation of capital is ignored. Hence, with proportional saving, domestic capital accumulation (absorption) is described by the equation, cf. (21), (16),

$$\begin{aligned}\dot{K} \equiv dK/dt &= Q_1 = sY/P_1 = s[Y_1 + Y_2/p] = L(s/P_1)y \\ &= Ls[y_1 \mathrm{l}_1 + (y_2/p)\mathrm{l}_2] \equiv L\,\mathfrak{g}(k).\end{aligned} \tag{26}$$

Thus, with the *factor endowments*, L and K, as *state variables*, the complete description of the growth process in the small trading economy is given by the *dynamic system*, (25)–(26).

This system (25)–(26) applies to growth processes with "fixed coefficient" sector technologies operating within the diversification cone (7) as well as to flexible sector technologies (1) operating with a domain of admissible endowment ratios, $k \in [0, \infty[$, cf. figure 1.

Within the *diversification* cone, we have, without loss of generality, that the dynamic system (25)–(26) of a small trading economy with proportional saving (and classical saving, cf. below) is always a *linear* system with constant coefficients, $\tilde{k}_1$, $\tilde{k}_2$, $\tilde{y}_1 = f_1(\tilde{k}_1)$, $\tilde{y}_2 = f_2(\tilde{k}_2)$, cf. figure 1.
With complete *specialization*, the *nonlinear* system (25)–(26) with flexible neoclassical technology represents a simple extension of the standard nonlinear Solow model to a small trading economy.

The differential equations (25)–(26), linear or nonlinear, represent a *homogenous dynamic* system of degree one. Such systems on $\mathbf{R}^2$ for any degree $m \in \mathbf{R}$ were studied in Jensen (1994), upon which we shall draw below. A seminal study of homogenous dynamics in $\mathbf{R}^n$ is found in Solow and Samuelson (1953).

Of course, the linearity of (25)–(26) within the diversification cone allows general closed form (quantitative) solutions. However, the main *economic issues* raised by (25)–(26) are *qualitative*, such as a partition of the *parameter space* (system parameters) into *regions* with roughly *similar* stability properties (structurally stable) of the family of solutions, cf. Gandolfo (1996, pp. 338), Jensen (1994, pp. 237).

A natural and useful tool in analyzing ***homogenous*** dynamics is a function, $h(k)$, called director function in Jensen (1994, pp. 195), pertaining to the time derivative of the *ratio* of the state variables.

In case of (25)–(26), it is easily seen that,

$$\begin{aligned} \dot{k} &= h(k) \equiv \mathfrak{g}(k) - k\,\mathfrak{f}(k) = (s/P_1)y - nk \\ &= s[y_1 \mathbb{1}_1 + (y_2/p)\mathbb{1}_2] - nk, \qquad k \in [0, \infty[. \end{aligned} \tag{27}$$

Lemma 1. *The director function, $h(k)$, (27), is of C^1-class on $[0, \infty[$, irrespective of $k_1 \gtrless k_2$ and diversification ($0 < \mathbb{1}_i < 1$) or specialization ($\mathbb{1}_i = 0$ or $\mathbb{1}_i = 1$); $h(k)$ is concave under the condition (2), and has a linear segment (flat) within C_k (7), (14)–(15), as depicted in figure 2.*

Proof. The Lemma follows from Proposition 1. The C^1-class property and the concavity of $y(k)$, $\mathfrak{g}(k)$, (27), and the linearity of nk establish the Lemma. □

As to the *parameter regions* of (27) and the long-term possibilities of diversification, we state:

Theorem 1. *For a small, competitive, trading economy, facing given terms of trade p, within the limits of (14)–(15), with proportional saving and exogenous labor growth, the necessary and sufficient conditions for the existence of a long-run factor endowment ratio (capital-labor ratio), $k = \kappa$, with incomplete specialization (diversification), are given by the parameter conditions, see figure 2.3 and figure 2.7,*

$$k_1 > k_2 : \quad \frac{\tilde{Y}_1}{\tilde{K}_1} = \frac{\tilde{y}_1}{\tilde{k}_1} < \frac{n}{s} < \frac{\tilde{y}_2/p}{\tilde{k}_2} = \frac{\tilde{\epsilon}_{K_1}}{\tilde{\epsilon}_{K_2}} \frac{\tilde{Y}_1}{K_1}, \tag{28}$$

$$k_2 > k_1 : \quad \frac{\tilde{y}_2/p}{\tilde{k}_2} < \frac{n}{s} < \frac{\tilde{y}_1}{\tilde{k}_1}. \tag{29}$$

When the existence conditions (28), (29) are satisfied, then the family of solutions for $k(t)$ to (27) have asymptotic stability in the diversification cone, C_k (7), (14)–(15), i.e.,

$$\exists k_0 = \kappa \in C_k \quad : \quad k(t) = \kappa, \quad \forall t, \tag{30}$$

$$\forall k_0 \in C_k \setminus \{\kappa\} : \quad k(t) \to \kappa \text{ as } t \to \infty. \tag{31}$$

Hence, for a small trading economy with an initial diversified state, the unique diversified steady state, κ, is an attractor in C_k, irrespective of the sector capital intensities: $k_1 \lesseqgtr k_2$.

Furthermore, the unique steady state, κ, obtained under diversification conditions (28)–(29) is a global attractor for any initial capital-labor ratio, i.e.,

$$\forall k_0 \in \mathbf{R} \setminus \{\kappa\} : \quad k(t) \to \kappa \text{ as } t \to \infty. \tag{32}$$

Thus, for any initial specialization, the small trading economy will, with the parameters (28)–(29) in the long run combine trade with competitive domestic production of both goods.
The long run (steady state) capital-labor ratio, κ, of the diversified economy in figure 2.3 and figure 2.7 is given by

$$\begin{aligned}\kappa &= \frac{\tilde{k}_2(p - \frac{\tilde{y}_2}{\tilde{y}_1}\frac{\tilde{k}_1}{\tilde{k}_2})}{p - \tilde{y}_2/\tilde{y}_1 - (np/s)(\tilde{k}_1 - \tilde{k}_2)/\tilde{y}_1} \\ &= \frac{\tilde{k}_2(1 - \tilde{\epsilon}_{K_2}/\tilde{\epsilon}_{K_1})}{1 - (\tilde{\epsilon}_{K_2}\tilde{k}_2/\tilde{\epsilon}_{K_1}\tilde{k}_1) - (n/s)(\tilde{k}_1 - \tilde{k}_2)/\tilde{y}_1}.\end{aligned} \tag{33}$$

A specialized steady state κ is given by either $f_1(\kappa)/\kappa = n/s$ or $f_2(\kappa)/\kappa = np/s$. The steady state – diversified or specialized – proportional growth rate is $\mathfrak{f}(\kappa) = \mathfrak{g}(\kappa)/\kappa = n$.

Proof. The procedure of proving (28)–(29) is the same, whence we only give case: $k_1 > k_2$.
With diversification, the director function $h(k)$, (27), becomes, cf. (6), (14)

$$\dot{k} = h(k) = \frac{s}{\tilde{k}_1 - \tilde{k}_2}\left[(\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2\right] + s\left[\frac{(\tilde{y}_1 - \tilde{y}_2/p)}{\tilde{k}_1 - \tilde{k}_2} - \frac{n}{s}\right]k. \tag{34}$$

From the diversification restriction (14), we get two inequalities

$$(i) \quad \tilde{y}_1 - \tilde{y}_2/p > 0, \qquad (ii) \quad (\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2 > 0. \tag{35}$$

By $k_1 > k_2$ and (ii), (35), it is immediately seen that the *intercept* of the line (34) is always *positive*. The *slope* of (34) is clearly *positive*, when the parameters satisfy the inequality stated in figure 2.1. When the latter inequality is not satisfied, the *slope* of (34) is *negative*.
With *negative* slope and *positive* intercept, the line (34) either *crosses* the k-axis *in* the diversification *interval* or cuts the k-axis *beyond* the diversification interval.

Incidently, note that the successive parameter intervals for n/s (left to right) in figure 2 are associated with *increasing* values of n/s, i.e., *larger* n and/or *smaller* s.

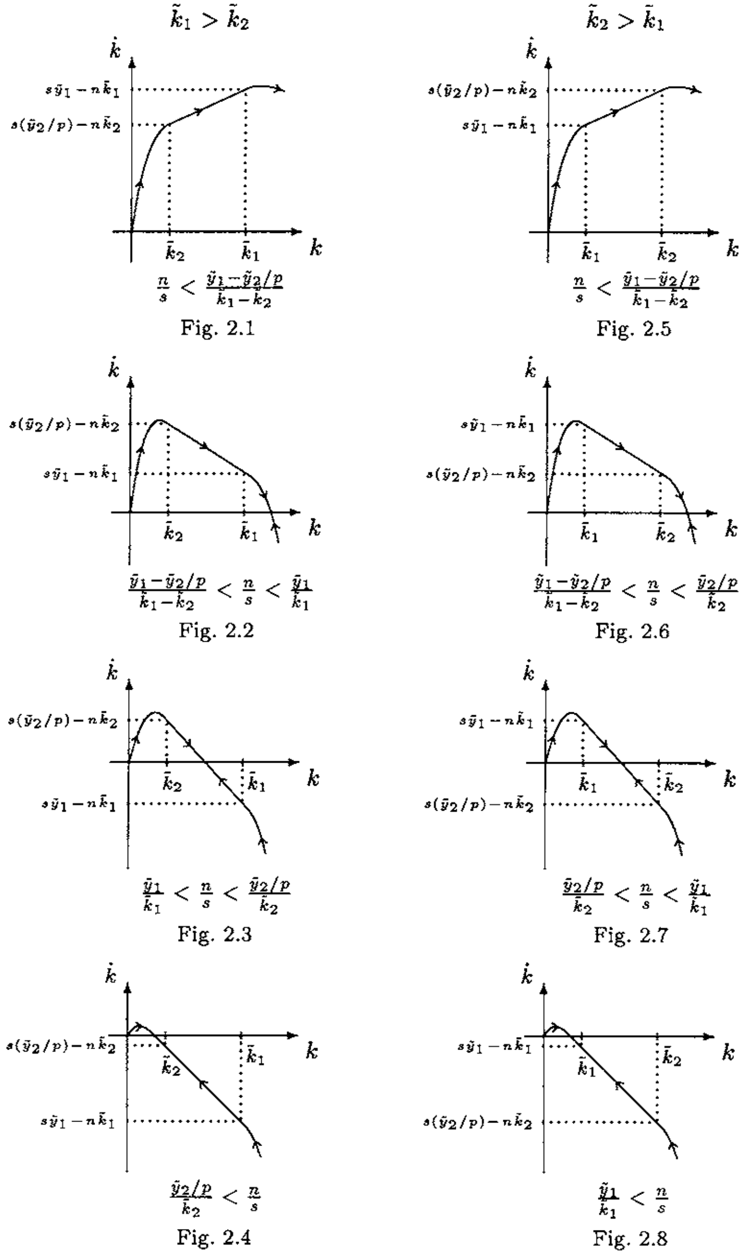


Fig. 2. The director function, $h(k)$, (27), with proportional saving and alternative parameter intervals for n/s

To determine the location of the line, (34), we may calculate $h(k)$ at the specialization points

$$k = \tilde{k}_1 : \quad h(k) = s\tilde{y}_1 - n\tilde{k}_1, \tag{36}$$

$$k = \tilde{k}_2 : \quad h(k) = s(\tilde{y}_2/p) - n\tilde{k}_2. \tag{37}$$

The *value* of $h(k)$, (36), is *positive*, when $(n/s) < \tilde{y}_1/\tilde{k}_1$, which is depicted in figure 2.2. The *value* of $h(k)$, (37), is *negative* when $(n/s) > (\tilde{y}_2/p)/\tilde{k}_2$, which is depicted in figure 2.4.

By (35), it can be seen that $\frac{\tilde{y}_1 - \tilde{y}_2/p}{\tilde{k}_1 - \tilde{k}_2} < \frac{\tilde{y}_1}{\tilde{k}_1} < \frac{\tilde{y}_2/p}{\tilde{k}_2}$. Hence – with the value of $h(k)$, (36), *negative* and the value of $h(k)$, (37), *positive* – the line $h(k)$, (34), evidently passes through the diversification interval with the parameter restriction depicted in figure 2.3. This establishes (28). Using (9), (11), we also have: $(\tilde{Y_2/p})/\tilde{k}_2 = (\tilde{Y_2/K_2})(\tilde{P_2/P_1}) = (\tilde{Y_2P_2/rK_2})(\tilde{r/P_1}) = (\tilde{Y_1/K_1})(\tilde{\epsilon}_{K_1}/\tilde{\epsilon}_{K_2})$.

The director function $h(k)$ on $k \in [0, \infty[$ and the respective parameter restrictions (belonging to the diversification interval) are shown in figure 2 – with figure 2.7, illustrating the restriction (29).

The proof of (30)–(31) follows immediately from the phase diagram, figure 2.3 and figure 2.7.

To prove (32), we first note that $h(k)$ with *specialization* and flexible technologies have the nonlinear forms, cf. (36)–(37)

$$h(k_1) = sf_1(k_1) - nk_1, \tag{38}$$

$$h(k_2) = (s/p)f_2(k_2) - nk_2. \tag{39}$$

With $n/s > f_1(k_1)/k_1$, it is seen that $h(k_1)$, (38), is always negative for $k_1 > \tilde{k}_1$ as illustrated in figure 2.3. This case is alternatively illustrated with solid lines in the traditional diagram, $k_1 > \kappa$, figure 3.1 below; see Wan (1971). Hence, an initial specialization in good 1 will eventually terminate, when the relative factor endowments attain the *upper* endpoint $\tilde{k}_1$ of the diversification interval in figure 2.3.

With $n/s < (f_2(k_2)/p)/k_2$, it is seen that $h(k_2)$, (39), is positive for small values of k_2, as illustrated in figure 2.3. This case is alternatively illustrated with solid lines in figure 3.2. As (39) coincides with (37) at the *lower* endpoint $\tilde{k}_2$ of the diversification interval, an initial specialization in good 2 eventually terminates as well.

The long-run diversified capital-labor ratio, κ, (33) is immediately obtained as the root of $h(k)$, (34), and a specialized κ follows from (27) with either $\mathfrak{l}_i = 0$.

The expression $\mathfrak{f}(\kappa) = \mathfrak{g}(\kappa)/\kappa$, $\kappa \neq 0$ for the proportional growth rate in steady state follows from (25) and $h(\kappa) = 0$, (27). $\square$

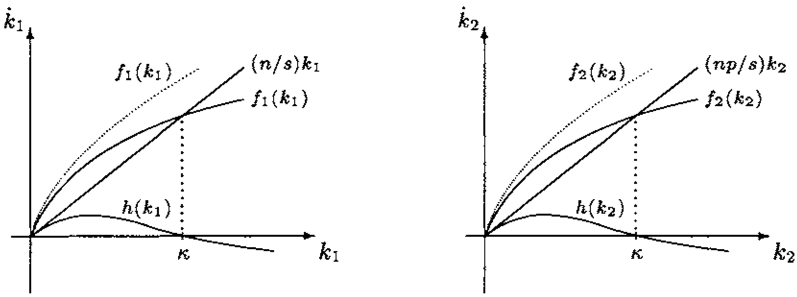


Fig. 3. Phase diagrams and alternative shapes of $f_i(k_i)$

Remark 1. In a *closed* economy, the Rybczynski theorem and $k_2 > k_1$ imply that a high k – and hence a large production of the consumer good – will subsequently reduce the high momentary capital-labor ratio. This is not necessarily the case in a *trading* economy, as the large production of consumer goods implies a high export of consumer goods, which in turn, by the balance of trade equilibrium then implies a large import of capital goods – and hence the high momentary domestic capital-labor ratio may in fact further increase. Accordingly, in figure 2, $k_2 > k_1$ does *not* anywhere appear as a sufficient *stability* condition for a steady state. ∇

Remark 2. Note, as stated in Theorem 1, that all the inequalities for n/s in figure 2 assume the given prices (terms of trade) to belong to the zero profit price intervals (14)–(15). If p did not belong to these intervals, we would never even temporarily and hence neither in the long-run (steady state) experience diversification. Figure 2 would neither contain any line segments, and complete specialization will prevail for any factor endowment ratio, k. ∇

Corollary 1. *When the parametric restrictions (28)–(29) of the dynamic system (25)–(26) are not satisfied, the small competitive trading economy will eventually be* specialized *in the* most capital intensive *good, except when both the saving rate, s, is very small, and the labor growth rate, n, is very high, cf. figure 2.4, figure 2.8.* A specialized steady state *will not always exist, in figures 2.1–2.2 and figures 2.5–2.6. As to the cases of specialized non-steady state growth, see Theorem 2.*

Proof. The proof of the corollary is analogous to the proof of (32) combined with the parameter regions given in figures 2.1–2.2 and figures 2.5–2.6. The shape of $f_i(k_i)$ will decide the existence of a steady state, cf. figure 3. □

For given terms of trade, p and labor growth n, it is evidently, besides *technology* conditions, the *domestic demand* composition, s, that is of decisive importance for the character of any steady state of the trading economy. Clearly, large domestic demand for capital goods relative to consumer goods will in the long run prevent competitive domestic production of both goods, cf. Corollary 1. Only if the domestic demand for capital goods, s, is properly restrained to the respective intervals (28)–(29), will *trade* be compatible with a *diversified steady state* of a small, growing, fully employed, *competitive* economy.

The possibilities for satisfying the diversification *criterion* (28)–(29) are directly *observable* in a small *currently diversified* competitive economy. Irrespective of the specification of the sectorial production functions, the average *productivity* of *capital* in the *capital good sector* (Y_1/K_1), (or capital output ratio, K_1/Y_1, measured in the same units) and the *cost* shares, $\tilde{\epsilon}_{K_i}$, are *observable* in the *diversification interval* of figure 2, where, cf. figure 1, the *allocation ratios* (1_i), output *mix* (Y_1/Y_2), income *distribution* (δ_K), and *trade pattern* are continuously changing with $k(t)$, but sectorial factor combinations $(\tilde{k}_i)$ and cost shares $(\tilde{\epsilon}_{K_i})$ remain time-invariant.

Regarding the actual values of the parameters in (28), the country tables in Obstfeld and Rogoff (1995), Summers and Heston (1991), Simon (1990) show the range of s (fraction of GNP) roughly as

$$\begin{array}{ccc} 1885\text{–}1913 & 1921\text{–}1939 & 1973\text{–}1991 \\ 0.1 \le s \le 0.2 & 0.15 \le s \le 0.2 & 0.15 \le s \le 0.3 \end{array} \tag{40}$$

The parameter n ("natural rate of proliferation"), the capital-output ratio (K_1/Y_1) and cost shares ϵ_{K_i}, roughly, have the range

$$0.005 \le n \le 0.3, \quad 2 \le K_1/Y_1 \le 6, \quad 0.25 \le \epsilon_{K_i} \le 0.8. \tag{41}$$

If we choose $K_1/Y_1 = 5$, then very *high* values of n and *low* values of s $(n/s = 0.02/0.1 = 0.2)$ are needed to have $(n/s) > Y_1/K_1 = 0.2$. With a lower $K_1/Y_1 = 4$, it is evidently more difficult to satisfy $(n/s) > Y_1/K_1$, cf. (28). Although the latter may be met by, e.g., $(n = 0.028, s = 0.13, K_1/Y_1 = 5)$, the actual parametric possibilities of n/s complying with the lower end point of the diversification *condition* (28) are very limited, especially in the later periods, cf. (40).

But when the *consumer* good sector is more *capital intensive*, it is evident that the *criterion* (29) is easier to satisfy within the range of the parameters in (40)–(41), especially in case of large differences between $\tilde{k}_1$ and $\tilde{k}_2$, hence between $\tilde{\epsilon}_{K_2}$ and $\tilde{\epsilon}_{K_1}$. If we still keep $K_1/Y_1 = 5$, then $n = 0.02$, $s = 0.12$ together with $\epsilon_{K_1} = 0.28$, $\epsilon_{K_2} = 0.35$ will satisfy (29). Furthermore, if modern agriculture with $\epsilon_K \geq 0.85$ is seen as the consumer good sector, then, e.g., the parameters: $K_1/Y_1 = 4$, $\epsilon_{K_1} = 0.3$, $\epsilon_{K_2} = 0.8$, and $n/s = 0.1\,(= 0.02/0.2 = 0.015/0.15$, etc.) will comply with (29).

Thus, although not necessary, the *ranking* $\tilde{k}_2 > \tilde{k}_1$ is more *conducive* to *diversification.* Parametrically, the *scope* of figure 2.7 is *empirically* somewhat *larger* than figure 2.3.

3.1.2 Specialization and Rapidity of Endogenous Growth

The dynamic implications of the *nonexistence* of a *steady state* solution to the neoclassical one-sector growth model of a closed economy, which formally looks similar to the *specialized trading economy* with $\dot{k}_i = h(k_i)$, (38)–(39), were studied in Jensen and Larsen (1987). We will here briefly discuss the dynamics of a trading economy, specialized in the consumer good and hence importing all its capital goods, i.e., cf. (39)

$$\dot{k}_2 = h(k_2) = (s/p)f_2(k_2) - nk_2, \quad k_2 = k > \tilde{k}_2. \tag{42}$$

The character of the solutions to (42) – when no root (steady state) exists, cf. figure 2.5 – depends on the shape of $h(k_2)$ and essentially $f_2(k_2)$. We may summarize the *endogeneity* conditions and the *rapidity* of per *capita* income *growth* as follows.

Theorem 2. *With constant returns to scale, (1), and exogenous labor growth, (25), a necessary condition for the endogenous growth of per capita income in a small specialized trading economy is that labor is inessential in production, i.e., cf. (1)*

$$\forall K_2: \quad F_2(0, K_2) > 0. \tag{43}$$

Under the condition (43) and the following alternative assumptions

$$\begin{aligned} f_2(k_2) - (np/s)k_2 &\rightarrow \varepsilon \geq 0 \text{ as } k_2 \rightarrow \infty, &(44)\\ f_2'(k_2) &\rightarrow \underline{\beta} = (np/s) \text{ as } k_2 \rightarrow \infty, &(45)\\ f_2'(k_2) &\rightarrow \underline{\beta} > (np/s) \text{ as } k_2 \rightarrow \infty, &(46) \end{aligned}$$

the capital-labor ratio $k_2(t)$ and per capita income, $y_2(t)$ will in the long run have time paths with the property of unbounded, linear, polynomial, *exponential growth, respectively. With (44)–(46) as giving alternative shapes for $f_2(k_2)$, cf. figure 3, corresponding properties of $h(k_2)$, (42), are shown in figure 4.*

Proof. The first part of the proof is given in Appendix A. Using figure 3 above, we note that $f_2(k_2)$ by, respectively, (44)–(46), has the line $(np/s)k_2$ as asymptote, or eventually becomes "parallel" to, or always remains steeper than $(np/s)k_2$. The consequences of (44)–(46) for the rapidity of growth were proved in Jensen and Larsen (1987). It follows from $k_2(t)$ going to infinity and l'Hospital that *per capita income*, $y_2(t) = f_2[k_2(t)]$, grows, in the long-run, with the same rapidity as the *capital-labor ratio*, $k_2(t)$. □

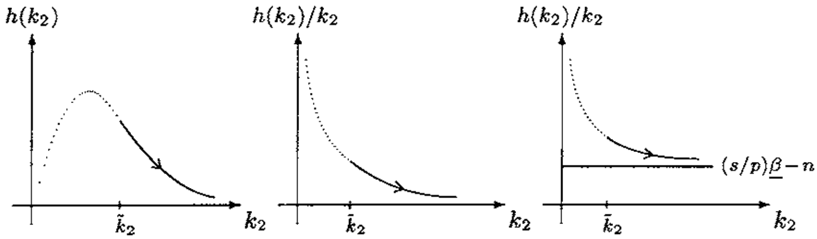


Fig. 4.1 Fig. 4.2 Fig. 4.3

Fig. 4. Rapidity of endogenous growth of per capita income. Bounded, polynomial, and exponential growth

The case of figure 4.3 with exponential growth is customarily referred to as *endogenous growth*, Barro and Sala-i-Martin (1995). See Romer (1986, 1994), Lucas (1988), Jones and Manuelli (1990), Rebelo (1991), Yang and Borland (1991). However, unbounded, linear or a polynomial increases of *per capita* income are relevant as growth prospects, too.

We want to emphasize that the production function, $F(L_2, K_2)$ must have the property: $\forall K_2:\ F(0, K_2) > 0$, i.e., *labor* is *inessential*. Whether capital is inessential or not, $\forall L_2:\ F(L_2, 0) \geq 0$, is immaterial for the endogenous growth properties in figure 4. When both capital and labor are inessential, the CES production function with substitution elasticity larger than 1 may give endogenous growth of per capita income, cf. Pitchford (1960), Long and Wong (1997).

The general intensive form of the CES functions is, cf. (1)

$$y_i = f_i(k_i) = \gamma_i \left[(1 - a_i) + a_i k^{\frac{\sigma_i - 1}{\sigma_i}}\right]^{\frac{\sigma_i}{\sigma_i - 1}},$$
$$\gamma_i > 0, \qquad 0 < a_i < 1, \qquad 0 < \sigma_i < \infty. \tag{47}$$

Corollary 2. *For a CES production function with a substitution elasticity, $\sigma_2 \leq 1$ – or $\sigma_2 > 1$ and $\gamma_2 a_i^{\frac{\sigma_2}{\sigma_2 - 1}} < np/s$ – the solutions, $k_2(t)$ to (42) cannot exhibit any endogenous (persistent) growth. The specialized steady state capital-labor is*

$$\kappa = \left[(1-a_2)^{-1}[(np/\gamma_2 s)^{\frac{\sigma_2 - 1}{\sigma_2}} - a_2]\right]^{\frac{\sigma_2}{1-\sigma_2}}. \tag{48}$$

With $\sigma_2 > 1$, the solutions, $k_2(t)$, to (42) exhibit endogenous (persistent) per capita growth as follows:

$$\textit{polynomial: } \gamma_2 a_2^{\frac{\sigma_2}{\sigma_2-1}} = np/s, \quad \textit{exponential: } \gamma_2 a_2^{\frac{\sigma_2}{\sigma_2-1}} > np/s. \tag{49}$$

Proof. From (47), we have

$$f_i'(k_i) = a_i \gamma_i \left[a_i + (1-a_i) k_i^{\frac{1-\sigma_i}{\sigma_i}}\right]^{\frac{1}{\sigma_i - 1}}. \tag{50}$$

Hence, by (47) and (50), we obtain the limits

$$\sigma_i < 1: \quad \lim_{k\to\infty} f_i(k) = \gamma_i [1-a_i]^{\frac{\sigma_i}{\sigma_i - 1}}, \quad \lim_{k\to\infty} f_i'(k) = \underline{\beta} = 0, \tag{51}$$

$$\sigma_i > 1: \quad \lim_{k\to\infty} f_i(k) \approx \gamma_i a_i^{\frac{\sigma_i}{\sigma_i - 1}} k, \qquad \lim_{k\to\infty} f_i'(k) = \underline{\beta} = \gamma_i a_i^{\frac{\sigma_i}{\sigma_i - 1}}. \tag{52}$$

With (51), or $\sigma_2 > 1$ and $\gamma_2 a_2^{\frac{\sigma_2}{\sigma_2-1}} < np/s$, a root (48) exists to (42), preventing endogenous (persistent) growth. The endogenous growth with rapidity (49) follows from (52), (45)–(46), figure 4, and $k = k_2$. □

With the CES function, the cases (44) are excluded, where linear growth is requiring $0 < \varepsilon \leq \lim_{k\to\infty} h(k) < \infty$. Either endogenous growth does not occur, or its rapidity will be at least polynomial.

Polynomial growth occurs with, e.g. $\gamma_2 = 1$, $a_2 = 0.3$, $\sigma_2 = 1.7$, $n = 0.01$, $p = 1$, $s = 0.186$, and hence exponential growth occurs with, e.g., $s = 0.2$. Accordingly, rather high substitution elasticities are generally required together with $0.2 < a_2 < 0.35$ and (40). The "total factor productivity" parameter, γ_2, and the terms of trade p may have ameliorating roles, cf. (49).

Regardless of the actual capital-intensive good of specialization and hence trade pattern, *persistent* growth of *per capita* national income of the trading economy occurs with a *high saving* rate s, *low* n, and a *low* $p = P_1/P_2$ (relatively *cheap capital* goods), which create the long-run opportunities for the necessary *specialization* and *rapid accumulation* of capital goods, cf. figure 2 and (26)–(27), and facilitate the *technology* (CES) to meet the *parametric conditions* (49).

Competitive factor pricing, (45)–(46), and $k(t)$ approaching infinity imply that capital's (labor's) *share*, (9)–(10), converges to one (zero)

$$\lim_{k_i\to\infty} \epsilon_{K_i} = \lim_{k_i\to\infty} \frac{k_i f_i'(k_i)}{f_i(k_i)} = \underline{\beta} \lim_{k_i\to\infty} \frac{k_i}{f_i(k_i)} = \underline{\beta}(1/\underline{\beta}) = 1. \tag{53}$$

Despite (53), the real wage exhibits persistent growth. With CES, (47), the wage-rental ratio, (10), is, $w/r = \omega = [(1-a_i)/a_i]k_i^{1/\sigma_i}$. Hence, the latter, (42), (46), and (52) give, cf. figure 4.3,

$$\lim_{t\to\infty} \dot{w}/w(t) = (1/\sigma_2) \lim_{t\to\infty} \dot{k}_2/k_2(t) = (1/\sigma_2)[(s/p)\underline{\beta} - n]. \tag{54}$$

The capital stock $K(t)$ grows at the exponential rate $(s/p)\underline{\beta}$, which is higher than the combined exponential rates of the labor force $L(t)$ and of the wage rate $w(t)$. This explains the limits (53) of the factor shares.

3.2 Dynamics by Proportional Saving and Endogenous Labor Growth

Any mathematical growth model is inevitably based on some fundamental set of assumptions, and the most critical ones should be fully elucidated. In this section, we therefore replace the assumption of exogenous labor growth by an endogeneity assumption. For economic growth, as observed over two centuries, the inclusion of increases in both population and per capita product is indispensable, Kuznets (1966, p. 20). As seen below, *persistent* per *capita* growth essentially requires that the *classical* (Malthusian) law of *population* (labor) growth is *terminated.*

For a small country with *fixed* terms of trade, and hence a fixed *real wage*, the latter cannot within the diversification cone endogenize the labor growth as in the *classical* canonical model of *closed* economies, cf. Niehans (1963), Samuelson (1978), Jensen (1994). In this section, we modify our former analysis by allowing the *growth* rate of the *labor* force to depend in a "*semi-classical*" way on *living standards* [consumption per capita, $c = Q_2/L$, cf. (18)–(19), (21), (16)], i.e. [1]

$$\dot{L}/L = nc = nQ_2/L, \tag{55}$$

$$\begin{aligned}\dot{L} &= nQ_2 = n(1-s)Y/P_2 = Ln[(1-s)/P_2]y \\ &= L\,n(1-s)(py_1 1_1 + y_2 1_2) \equiv L\,\mathfrak{f}(k).\end{aligned} \tag{56}$$

The governing function of capital accumulation remains the same as (26)

$$\dot{K} = s(Y/P_1) = L(s/P_1)y = Ls(y_1 1_1 + (y_2/p)1_2) \equiv L\,\mathfrak{g}(k). \tag{57}$$

Hence, the director function $h(k)$ becomes

$$\begin{aligned}\dot{k} &= h(k) \equiv \mathfrak{g}(k) - k\,\mathfrak{f}(k) = (s/P_1)y - [n(1-s)/P_2]yk \\ &= [y_1 1_1 + (y_2/p)1_2][s - np(1-s)k], \qquad k \in [0, \infty[.\end{aligned} \tag{58}$$

1. The parameter n, (55), (56) is inherently and numerically different from the natural growth rate, n, (25), (41).

Theorem 3. *For a small, competitive, trading economy with proportional saving and endogenous ("semi-classical") labor growth, the necessary and sufficient conditions for the existence of a long-run capital-labor ratio, $k = \kappa$, with incomplete specialization are given by the parameter restrictions, see figure 5.2:*

$$\tilde{k}_1 > \tilde{k}_2: \quad \frac{np\tilde{k}_1}{1+np\tilde{k}_1} < s < \frac{np\tilde{k}_2}{1+np\tilde{k}_2}, \tag{59}$$

$$\tilde{k}_2 > \tilde{k}_1: \quad \frac{np\tilde{k}_2}{1+np\tilde{k}_2} < s < \frac{np\tilde{k}_1}{1+np\tilde{k}_1}. \tag{60}$$

When the existence conditions, (59)–(60), are satisfied, then the family of solutions for $k(t)$ to (58) have asymptotic stability in the diversification cone, C_k (7), (14)–(15), i.e.,

$$\exists k_0 = \kappa \in C_k \quad : \quad k(t) = \kappa, \ \forall t, \tag{61}$$

$$\forall k_0 \in C_k \setminus \{\kappa\}: \quad k(t) \to \kappa \text{ as } t \to \infty. \tag{62}$$

Hence, for small trading economies with an initial diversified state, the unique diversified steady state, κ, is an attractor in C_k, irrespective of the sector capital intensities: $k_1 \gtrless k_2$.
Furthermore, the unique steady state, κ, obtained under the conditions (59)–(60) is a global attractor for any initial capital-labor ratio, i.e.

$$\forall k_0 \in \mathbf{R} \setminus \{\kappa\}: \quad k(t) \to \kappa \text{ as } t \to \infty. \tag{63}$$

Thus, for any initial specialization of the small trading economy, it will with the parameters (59)–(60), in the long run combine trade with competitive domestic production of both goods.
The long-run capital-labor ratio, κ, of the diversified – and specialized – economy is

$$\kappa = \frac{s}{np(1-s)}. \tag{64}$$

The proportional growth rate in a diversified steady state is

$$\begin{aligned} \mathfrak{f}(\kappa) &= \mathfrak{g}(\kappa)/\kappa = [n(1-s)/P_2][\tilde{r}\kappa + \tilde{w}] = (s/P_1)(\tilde{r} + \tilde{w}/\kappa) && (65)\\ &= n(1-s)p\tilde{y}_1[\tilde{\epsilon}_{L_1} + \tilde{\epsilon}_{K_1}/\tilde{k}_1] \\ &= \tilde{y}_2[n(1-s)\tilde{\epsilon}_{L_2} + (s/p)(\tilde{\epsilon}_{K_2}/\tilde{k}_2)]. && (66)\end{aligned}$$

The specialized steady state growth rate is given by (65), without tilde, evaluated at κ, (64). The trajectories of the phase portrait of (56)–(57) are straights lines everywhere, parallel to the ray with the slope (64).

Proof. The procedure of proving (59)–(60) is the same, whence we only give: Case $k_1 > k_2$, see figure 5.
Again, the diversification restrictions are, cf. (35)

$$(i) \quad \tilde{y}_1 - (\tilde{y}_2/p) > 0, \qquad (ii) \quad (\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2 > 0. \tag{67}$$

With diversification, the director function (58) becomes, cf. (6),

$$\begin{aligned}\dot{k} = h(k) = & \frac{1}{\tilde{k}_1 - \tilde{k}_2}\{ - np(1-s)(\tilde{y}_1 - \tilde{y}_2/p)k^2 \\ & - \left[s(\tilde{y}_2/p - \tilde{y}_1) + np(1-s)((\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2)\right]k \\ & + s((\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2)\}.\end{aligned} \tag{68}$$

The director function, $h(k)$, (68), is a parabola with a positive discriminant

$$\begin{aligned}\triangle^2 = & \frac{1}{4}\Big[\frac{1}{\tilde{k}_1 - \tilde{k}_2}\big(s\{(\tilde{y}_2/p) - \tilde{y}_1\} \\ & -np(1-s)\{(\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2\}\big)\Big]^2 > 0,\end{aligned} \tag{69}$$

and hence has two real roots. The product of the roots, κ_1 and κ_2, is seen to be, cf. (68)

$$\kappa_1\,\kappa_2 = -\frac{(\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2}{(n/s)p(1-s)(\tilde{y}_1 - \tilde{y}_2/p)} < 0, \tag{70}$$

which is always negative with diversification conditions (67). As the coefficient of the quadratic term of (68) is always negative with (67), the parabola has a shape with a maximum value. The parabolas are depicted in figures 5.1–5.3, with solid lines in the diversification interval and with dotted lines outside.

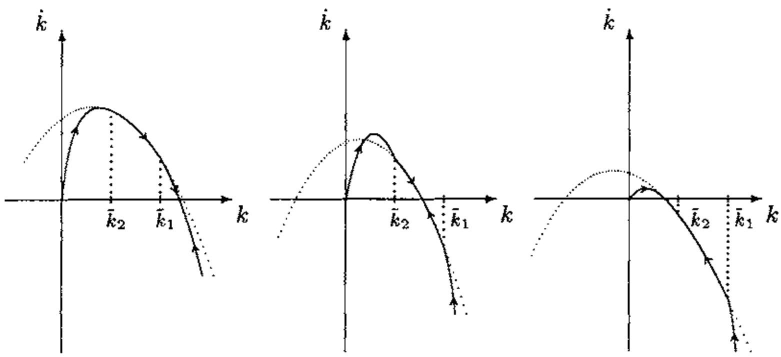
The director function, $h(k)$, (58), at the specialization points becomes

$$k = \tilde{k}_1: \quad h(k) = \tilde{y}_1[s - n(1-s)p\tilde{k}_1], \tag{71}$$

$$k = \tilde{k}_2: \quad h(k) = (\tilde{y}_2/p)[s - n(1-s)p\tilde{k}_2]. \tag{72}$$

The *value* of $h(k)$, (71) is *positive* when $s > n(1-s)p\tilde{k}_1$, which is equivalent to the condition stated in figure 5.1. The value of $h(k)$, (72) is *negative* when $s < n(1-s)p\tilde{k}_2$, which is equivalent to the condition stated in figure 5.3.
With a saving ratio s between these two limits, the director function $h(k)$, (58), passes through the diversification interval as depicted in figure 5.2. This establishes (59). The proof of (60) is analogous. The proof of (61)–(62) follows directly from the phase diagrams, figure 5.2.

The proof of (63) is straightforward, as the replacement of $\tilde{y}_1$ by $f_1(k_1)$ in (71) and $\tilde{y}_2$ by $f_2(k_2)$ in (72) does not affect the value of $h(k)$.
The long-run capital-labor ratio, κ, (64), can be obtained as the positive root of (58). However, it can be immediately verified, by noting that $dK/dL \equiv \dot{K}/\dot{L} = s/[np(1-s)]$, cf. (56)–(57). The latter explains the shape of the phase portrait. The diversification conditions (59)–(60) also follow directly from (64) and the requirement that κ is between $\tilde{k}_1$ and $\tilde{k}_2$.
The expression (66) follows from (56)–(57) and (10)–(11), (16). □



$\frac{np\tilde{k}_2}{1+np\tilde{k}_2} < s$ Fig. 5.1

$\frac{np\tilde{k}_1}{1+np\tilde{k}_1} < s < \frac{np\tilde{k}_2}{1+np\tilde{k}_2}$ Fig. 5.2

$s < \frac{np\tilde{k}_1}{1+np\tilde{k}_1}$ Fig. 5.3

Fig. 5. The director function, $h(k)$, (58), with parameter intervals for s

The diversification possibilities (59)–(60) may be wide, when $\tilde{k}_1$ and $\tilde{k}_2$ differ considerably; the crucial parameter, n, is not, as is s, tied to the values (41).

Corollary 3. *When the parametric restrictions (59)–(60) are not satisfied, the small competitive trading economy will be* specialized *in the* most capital intensive *good, except when the saving rate, s, is too small, cf. figure 5.3. A specialized steady state in the capital intensive good will always exist. Hence, there is no possibility of endogenous (persistent) per capita growth with the endogenous labor growth (56).*

Proof. The proof of the corollary is analogous to the proof of (63). The necessary existence of a specialized steady state growth follows again from (71)–(72), which for $k > \tilde{k}_1$ always have roots irrespective of the shape of $y_1 = f_1(k_1)$ and $y_2 = f_2(k_2)$. □

Remark 3. One might be inclined to see the absence of endogenous per capita growth in Corollary 3 as hinging on the particular specification of endogenous labor growth in (56). As to Corollary 3 being generally upheld, let us briefly examine the classical alternative to (56)–(57), viz.

$$\dot{L} = Ln(w/P_2) \equiv L\,\mathfrak{f}(k). \tag{73}$$

Then by (73) and (57), we get, cf. (16)

$$\begin{aligned}\dot{k} &= h(k) \equiv \mathfrak{g}(k) - k\,\mathfrak{f}(k) = (s/P_1)y - n(w/P_2)k \\ &= n(r/P_2)[(s/np) - w/r]k + s(w/P_1).\end{aligned} \tag{74}$$

Imposing the diversification restrictions (14) upon (74), the *parameter conditions* for *diversification* become, after some calculations,

$$\frac{\tilde{y}_1}{\tilde{k}_1}\,\frac{\tilde{k}_1 - \tilde{k}_2}{\tilde{y}_2\tilde{k}_1 - \tilde{y}_1\tilde{k}_2 p} < \frac{n}{s} < \frac{\tilde{y}_2/p}{\tilde{k}_2}\,\frac{\tilde{k}_1 - \tilde{k}_2}{\tilde{y}_2\tilde{k}_1 - \tilde{y}_1\tilde{k}_2 p}. \tag{75}$$

Apart from a multiplicative factor, the condition (75) corresponds to (28). It is a bit more complicated than before, as (14) involves restriction on the fixed parameter (w/P_2) in (74).

If the condition (75) is not met, and we get *specialization* in e.g., good 2, then (74) can be reduced to, cf. (10)–(11)

$$\dot{k} = h(k) = f_2(k_2)[(s/p) - (n + \epsilon_{K_2})k_2]. \tag{76}$$

Clearly, a steady state always exists, and hence *persistent* growth *per capita* is not possible with classical endogenous labor growth (73). ∇

3.3 Dynamics with Classical Saving

In the literature on open economies and capital accumulation, the classical saving function has often been applied in the dynamic models, cf. Bardhan (1965, 1966), Stiglitz (1970), Deardorff (1971), Smith (1984). For a small trading economy with classical saving, domestic capital accumulation is given by, cf. (22), (11), (13),

$$\begin{aligned}\dot{K} &= Q_1 = Ls_K(r/P_1)k = Ls_K[\tilde{y}_1/(\tilde{k}_1 + \tilde{\omega})]k \\ &= \frac{L}{\tilde{k}_1 - \tilde{k}_2}[s_K(\tilde{y}_1 - \tilde{y}_2/p)k] \equiv L\,\mathfrak{g}(k).\end{aligned} \tag{77}$$

Hence, (77) and *exogenous labor* growth, (25), give, cf. (11), (13),

$$\begin{aligned}\dot{k} &= h(k) \equiv \mathfrak{g}(k) - k\,\mathfrak{f}(k) = [s_K(r/P_1) - n]k \\ &= s_K[\frac{\tilde{y}_1 - \tilde{y}_2/p}{\tilde{k}_1 - \tilde{k}_2} - \frac{n}{s_K}]k.\end{aligned} \tag{78}$$

The last expression in both (77)–(78) applies to $k \in C_k$.

Except for the singular case, when the line segment of $h(k)$, (78) coincides with the k-axis, only specialized steady states exist with classical saving. With specialization in the capital intensive good, and $r/P_1 = f_1'(k_1) = (1/p)f_2'(k_2)$, cf. (77), (11), the *conditions* and *rapidity* of persistent growth of per capita income with *classical saving* are the *same* as stated in (43)–(46) in Theorem 2.

The general conclusion that *classical* saving always leads to *complete* specialization in the long run would be *premature*, as the assumption of *exogenous* labor growth may again be critical in this respect.

We shall again use the *endogeneity* of *labor* growth formulation (56) modified with a classical demand for Q_2. We get, cf. (22), (19), (16),

$$\begin{aligned}\dot{L} &= nQ_2 = n/P_2[(1-s_K)rK + wL] \\ &= Ln/P_2[(1-s_K)rk + w] \equiv L\,\mathfrak{f}(k).\end{aligned} \tag{79}$$

Then the dynamic processes (77), (79) give,

$$\begin{aligned}\dot{k} = h(k) &\equiv \mathfrak{g}(k) - k\,\mathfrak{f}(k) = (s_K r/P_1)k - n/P_2[(1-s_K)rk + w]k \\ &= -n(1-s_K)(r/P_2)k^2 + [s_K(r/P_1) - n(w/P_2)]k.\end{aligned} \tag{80}$$

Theorem 4. *For a small, competitive, trading economy with classical saving and endogenous labor growth, the necessary and sufficient conditions for the existence of a long-run capital-labor ratio, $k = \kappa$, with incomplete specialization are given by the parameter restrictions:*

$$\tilde{k}_1 > \tilde{k}_2 : \ \frac{n\tilde{y}_2(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1 + np\tilde{k}_2)} < s_K < \frac{np\tilde{y}_1(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1 + np\tilde{k}_1)}, \tag{81}$$

$$\tilde{k}_2 > \tilde{k}_1 : \ \frac{npy_1(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1 + np\tilde{k}_1)} < s_K < \frac{n\tilde{y}_2(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1 + np\tilde{k}_2)}. \tag{82}$$

When the existence conditions, (81)–(82), are satisfied, then the family of solutions for $k(t)$ to (80) have asymptotic stability in the diversification cone, C_k (7), and globally, irrespective of the sector capital intensities: $k_1 \gtrless k_2$, i.e.

$$\begin{aligned}\exists k_0 &= \kappa \in C_k \ : \quad k(t) = \kappa, \quad \forall t, \end{aligned} \tag{83}$$

$$\begin{aligned}\forall k_0 &\in \mathbf{R} \setminus \{\kappa\} : \quad k(t) \to \kappa \text{ as } t \to \infty.\end{aligned} \tag{84}$$

The diversified steady state capital-labor ratio, κ, and its proportional growth rate are given by

$$\kappa = \frac{(s_K/n)(\tilde{y}_1 - \tilde{y}_2/p) + p[\tilde{y}_1\tilde{k}_2 - (\tilde{y}_2/p)\tilde{k}_1]}{(1-s_K)p(\tilde{y}_1 - \tilde{y}_2/p)}, \tag{85}$$

$$\begin{aligned}\mathfrak{f}(\kappa) &= \mathfrak{g}(\kappa)/\kappa = (n/P_2)[(1-s_K)\tilde{r}\kappa + \tilde{w}] \\ &= (s_K/P_1)\tilde{r} = s_K f_1'(\tilde{k}_1).\end{aligned} \tag{86}$$

Furthermore, the specialized steady state capital-labor ratio, κ, and its proportional growth rate are given by

$$\kappa = \frac{(s_K/n)\epsilon_K(\kappa)}{(1-s_K)\epsilon_K(\kappa) + (1-\epsilon_K(\kappa))p}, \tag{87}$$

$$\mathfrak{f}(\kappa) = \mathfrak{g}(\kappa)/\kappa = s_K f_1'(\kappa). \tag{88}$$

Proof. The proof of (81) is analogous to that of (82). We give case $k_1 > k_2$.
With diversification, the function (80) becomes, cf. (16), (11), (13),

$$\begin{aligned} h(k) = & \frac{n}{\tilde{k}_1 - \tilde{k}_2}\{-p(1-s_K)(\tilde{y}_1 - \tilde{y}_2/p)k^2 + \\ & [(s_K/n)(\tilde{y}_1 - \tilde{y}_2/p) - p(\tilde{y}_1\tilde{k}_2 - (\tilde{y}_2/p)\tilde{k}_1)]k\}. \end{aligned} \tag{89}$$

It is clear that the director function, $h(k)$, (89), has one zero root. As the coefficient of the quadratic term of (89) is always negative with $(\tilde{y}_1 - \tilde{y}_2/p)/(\tilde{k}_1 - \tilde{k}_2) > 0$, (67), the shape of the parabola (89) is the same as in figures 5.1–5.3.
The nonzero root of $h(k)$ can be calculated as given in (85). This root is also the attractive steady state capital-labor ratio, if the root is positive and located between $\tilde{k}_1$ and $\tilde{k}_2$. As the denominator of (85) is always negative in case $\tilde{k}_1 > \tilde{k}_2$, a negative numerator requires that

$$1 > s_K > \frac{np[(\tilde{y}_2/p)\tilde{k}_1 - \tilde{y}_1\tilde{k}_2]}{\tilde{y}_1 - \tilde{y}_2/p}. \tag{90}$$

The condition that κ, (85) is located as $\tilde{k}_2 < \kappa < \tilde{k}_1$ is seen, after some manipulation, to further require that

$$\frac{n\tilde{y}_2(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1+np\tilde{k}_2)} < s_K < \frac{np\tilde{y}_1(\tilde{k}_1 - \tilde{k}_2)}{(\tilde{y}_1 - \tilde{y}_2/p)(1+np\tilde{k}_1)}. \tag{91}$$

Comparing the left hand side of (91) with (90) and using $s_K \leq 1$, it is seen that (90) is always satisfied by (91), which is (81).
The proof of (83)–(84) follows immediately from the shape of the parabola, $h(k)$, (89).

Finally, the long-run capital-labor ratio, (85), is a nonzero root of (89), and the balanced growth rate, (86), follows as usual from (79) and (77); (87)–(88) can be similarly obtained by replacing $r/P_i = f_i'(k)$, $w/P_i = f_i(k) - kf_i'(k)$ in (80), cf. (10). $\square$

Corollary 4. *There is no possibility of endogenous (persistent) per capita growth with classical saving, (77), and semiclassical labor growth, (79), or classical labor growth, (73), except for σ_i infinite in (47).*

Proof. The function $h(k)$, (80) with specialization in, e.g., $k = k_1$, becomes, cf. (10)–(11),

$$h(k) = s_K f'(k)k - nk[(1 - s_K)f'(k)k + pf(k) - pkf'(k)]. \qquad (92)$$

Taking the limit of $h(k)$ when $k \to \infty$, we have

$$\lim_{k\to\infty} h(k) = \lim_{k\to\infty} k^2[-n(1 - s_K)f'(k)] \to -\infty, \qquad (93)$$

so there must exist a root, $h(\kappa) = 0$; hence, no persistent growth.

With classical labor growth (73) and classical saving (77), the function $h(k)$, (78), becomes, cf. (10)

$$\dot{k} = h(k) = s_K(r/P_1)k - n(w/P_2)k = kf'(k)[s_K - np\omega(k)]. \qquad (94)$$

For a CES function, (94) only gives $k(t) \to \infty$ in the *extreme* case of linear isoquants ($\sigma = \infty$), cf. Jensen (1994, p. 45). □

By comparing the endogenous *balanced growth* rates, (65) and (86), it is seen that, if the saving rates, s and s_K, have similar size, then *diversified* economies grow *faster* with *proportional* saving than with *classical* saving. Evidently, savings from wage income contributes to maintaining a larger balanced growth rate. Another noteworthy aspect of (86) is that a *higher saving* rate, s_K unequivocally [irrespective of changes in κ, (85)] increases the balanced growth rate of a small, diversified, *trading* country. A higher saving rate, s_K, in a classical *closed* two-sector growth model with endogenous labor growth will *not* necessarily increase the *balanced* growth rate, as changes in the general equilibrium prices and factor intensities may offset the effects of a larger s_K, Jensen (1994, p. 153). Thus, with *trade* at *fixed* terms of *trade*, high *saving* rates do *increase* the *balanced growth* in a classical setting of endogenous labor growth.

Remark 4. With specialization as before, (42), let us compare savings from capital income (77) with savings from only *labor income*, cf. (10)–(11): $\dot{K} = Q_1 = Ls_w(w/P_1) = (s_w/p)[f_2(k_2) - k_2 f_2'(k_2)] \equiv L\,\mathrm{g}(k)$, which, together with $\dot{L} = nL$, gives $\dot{k}_2 = h(k_2) \equiv (s_w/p)[f_2(k_2) - k_2 f_2'(k_2)] - nk_2$. Using l'Hospital, we get, $\lim_{k_2\to\infty} h(k_2) \approx -nk_2$. Accordingly $h(k_2)$ has a root, preventing $k_2 \to \infty$. Thus, despite $f_2'(k_2)$ being properly *bounded below*, and also *exogenous labor* growth, *savings* from only *labor income* cannot – in contrast to (26), (77) – generate *persistent* per *capita* growth. See hereto (53), (54), and OLG models with all savings from wage income, Galor and Lin (1997), Bertola (1996). ▽

3.4 Dynamics of the Trading Economy with Optimal Saving

Above, we analyzed the trading economy with the saving rate exogenous and constant, proportional and classical, respectively. In this section, the saving rate mostly varies over time, as it will be endogenously determined by infinitely lived consumers maximizing total lifetime utility.

The representative consumer is assumed to have a time additive intertemporal utility function

$$U = \int_0^\infty u[c(t)]e^{-\rho t}dt, \tag{95}$$

where the decision variable, c, is per capita consumption, cf. (55), and ρ is the constant subjective rate of time preference (discount rate). The *momentary* utility (felicity) function, $u(c)$, is assumed to be concave, increasing, and to satisfy the Inada conditions

$$\lim_{c\to 0} u'(c) = \infty, \qquad \lim_{c\to\infty} u'(c) = 0. \tag{96}$$

With $\rho > 0$, total utility, U is bounded, if $u(c)$ is bounded over time.

Remark 5. With $L(t) = L_0e^{nt}$, a social planner may maximize the objective function

$$U = \int_0^\infty u[c(t)]L(t)e^{-\rho t}dt = L_0\int_0^\infty u[c(t)]e^{-(\rho-n)t}dt, \tag{97}$$

where (97) is similar to (95) with $\breve{\rho} = \rho - n$. But with endogenous labor growth (56), the social planning objective analogous to (97) would be intractable. Hence, we stick only to (95); see Barro and Sala-i-Martin (1995), Deaton (1992). ∇

3.4.1 Ramsey Saving and Exogenous Labor Growth

For a small, two-sector trading economy, the domestic capital accumulation can be described by cf. (19), (16), (12), (55)

$$\begin{aligned}\dot{K} &= Q_1 = 1/P_1\,[Y - P_2Q_2] = L\,[y/P_1 - (1/p)Q_2/L] \\ &= L\,[y/P_1 - c/p] \equiv L\,\mathfrak{g}(k,c).\end{aligned} \tag{98}$$

With exogenous labor growth (25), the function $h(k,c)$, becomes, cf. (27) and (98)

$$\dot{k} = h(k,c) \equiv \mathfrak{g}(k,c) - k\,\mathfrak{f}(k,c) = (1/P_1)y(k) - c/p - nk. \tag{99}$$

Proposition 2. *(Stiglitz). For a small, competitive, trading economy with Ramsey saving and an exogenous labor growth, the necessary and sufficient condition for diversification is that the rate of time preference, ρ, has the particular value*

$$\rho = \tilde{\rho} = (y'(k)/P_1) - n = f_1'(k_1) - n = (f_2'(k_2)/p) - n. \tag{100}$$

There are multiple diversified steady states (k, $c = y(k)/P_2 - npk$), but no transition dynamics. If ρ is larger (less) than $\tilde{\rho}$, then the small trading economy will in the long run specialize in the labor (capital) intensive good. The transition dynamics (saddle paths) are shown in figure 6.1.

Proof. The Ramsey optimization problem is

$$\max U = \max_{c(t)} \int_0^\infty u[c(t)]e^{-\rho t}dt \tag{101}$$

$$\text{s.t.} \quad \dot{k} = h(k,c) = (1/P_1)y(k) - nk - c/p, \quad c \geq 0, \tag{102}$$

which is equivalent to maximizing the current value Hamiltonian function

$$\mathcal{H} = u[c(t)] + \pi(t)\left[(1/P_1)y(k) - nk - c/p\right], \tag{103}$$

with a costate variable, $\pi(t)$, and the transversality conditions:

$$k(0) = k_0, \quad \lim_{t\to\infty} \pi(t)e^{-\rho t}k(t) = 0. \tag{104}$$

The first order condition gives, cf. (103)

$$\frac{\partial \mathcal{H}}{\partial c} = u'(c) - 1/p\pi(t) = 0, \tag{105}$$

and the maximum principle gives that

$$\dot{\pi}(t) = -\frac{\partial \mathcal{H}}{\partial k} + \rho\pi(t) = \pi(t)\left[\rho + n - y'(k)/P_1\right]. \tag{106}$$

By derivation and inserting the first order condition (105) into (106), we obtain the differential equation for per capita consumption, c, as

$$\dot{c} = -\frac{u'(c)}{u''(c)}\left[\frac{y'(k)}{P_1} - n - \rho\right] \equiv c\sigma(c)\left[\frac{y'(k)}{P_1} - n - \rho\right] \equiv \eta(k,c), \tag{107}$$

where $\sigma(c)$ is the intertemporal elasticity of substitution.

By (105) the transversality condition (104) becomes

$$pu'(c)e^{-\rho t}k(t) \to 0 \text{ as } t \to \infty. \tag{108}$$

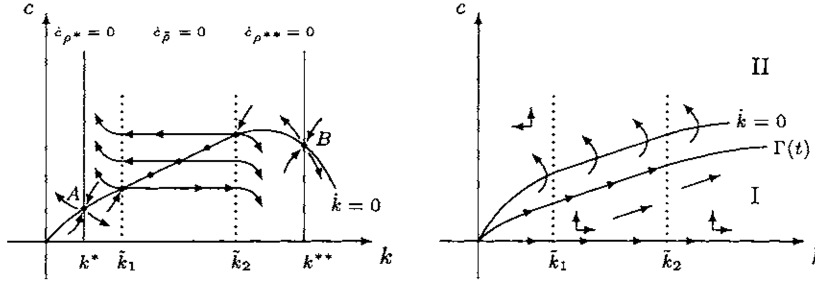


Fig. 6.1 Fig. 6.2

Fig. 6. Dynamics of a small trading economy with optimal saving and exogenous labor growth

The equations (99), (107) define a dynamic system in k and c. As the concave GNP function, $y(k)$, has a linear segment in C_k, cf. Proposition 1, a stationary c occurs in C_k, iff ρ has the *particular value* stated in (100), cf. (107), (16), (12), (11). For every $k \in C_k$, the optimal control variable, c follows from $h(k, c) = 0$, (102). Outside the line $h(k, c) = 0$, $\dot{k}$ is either positive or negative and we will eventually be specialized as indicated in figure 6.1.

When ρ becomes *larger*, $\dot{c} = 0$, (107), will therefore require $y'(k)$ to be larger, which gives a vertical line $k = k^* < \tilde{k}_1$, crossing the $\dot{k} = 0$ curve, and giving a long run steady state A in figure 6.1. Similarly, when ρ becomes *smaller*, (107) will give a vertical line $k = k^{**} > \tilde{k}_2$, which crosses the $\dot{k} = 0$ curve after the diversification cone, implying specialization in the capital-intensive good, cf. point B in figure 6.1. □

Remark 6. The sufficiency of (101)–(106) can be proved simply. It is observed that the objective function $u[c(t)]e^{-\rho t}$ is a concave function in (k, c)-space, cf. (96). Furthermore, it can be shown, by using (105) and (96), that $\pi(t) = pu'(c) > 0$, and that the function, $h(k, c)$, in (102) is also concave in c, k. Therefore the necessary conditions provided by the maximum principle are also sufficient for optimal solutions. ▽

Theorem 5. *With Ramsey saving and an exogenous labor growth, persistent (endogenous) per capita growth can be obtained if the concave per capita GNP function $y(k)$ and the intertemporal substitution elasticity $\sigma(c)$ or the rate of time preference ρ satisfies, respectively*

$$\lim_{k \to \infty} y'(k)/P_1 \equiv \beta/P_1 > n + \rho, \tag{109}$$

$$\overline{\sigma} = \sup_{c>0} \sigma(c) < \frac{\beta/P_1 - n}{\beta/P_1 - (n+\rho)}, \quad \text{i.e.} \quad \rho > \left[\frac{\beta}{P_1} - n\right]\left[\frac{\overline{\sigma} - 1}{\overline{\sigma}}\right]. \tag{110}$$

The conditions (109)–(110) ensure the existence – below the isocline $\dot{k} = h(k, c) = 0$ – of a separator, the particular orbit $\Gamma(t) \equiv [k^*(t), c^*(t)]$, depicted in figure 6.2. The existence of this separator is required for persistent (endogenous) per capita growth.

Proof. See Appendix B. □

With regard to Ramsey (optimal) saving, it has been incumbent on us to obtain *sufficient conditions* – applicable to a general GNP-*function*, $y(k)$ and a *general utility* function $u(c)$ – that ensure *persistent* per *capita* growth. The condition (109) is analogous to (44)–(45) with a low ρ taking over the role of a large s. But (109) is not always enough to ensure persistent growth, as (110) is also needed. However, if $u(c)$ always has $\sigma(c) \leq 1$, $\forall c > 0$, then (110) is automatically satisfied[2], irrespective of the size of $\rho > 0$. If $\sigma(c) > 1$, then ρ must be large enough to satisfy (110). Given now the existence of the *separator* $\Gamma(t)$ and hence *persistent* growth, see figure 6.2, the actual selection of the *optimal* path in region I is discussed in Appendix B.

For the class of *isoelastic*[3] utility functions $u(c)$, $\forall c$, $\sigma(c) = \sigma \gtrless 1$ and with the conditions (109)–(110), the *separator* in figure 6.2 is also the *unique optimal orbit* (solution) satisfying (101)–(104), see Appendix B. For optimal saving and persistent growth, the factor shares also behave as in (53), cf. (109).

3.4.2 Ramsey Saving and Endogenous Labor Growth

With the endogenous labor growth, $\dot{L} = nQ_2 = Lnc \equiv L\mathfrak{f}(k, c)$, (55), the function $h(k, c)$, (99), accordingly becomes

$$\dot{k} = h(k, c) = (1/P_1)y(k) - c/p - nck. \tag{111}$$

Theorem 6. *With the Ramsey saving and an endogenous labor growth, a long-run steady state always exists, either in the diversification cone or in the specialization region, depending on the values of the parameters, n, ρ, p. Endogenous per capita growth is therefore impossible, even with the property (109) of $y(k)$.*

Proof. Using the maximum principle and first order conditions similar to those above, (105)–(106), we derive the differential equation for per capita consumption as

2. Hall (1988) estimated that σ is much below unity, $0.1 < \sigma < 0.4$.

3. A common practice is to use: $u(c) = \frac{c^{1-\theta} - 1}{1-\theta}$, $\theta > 0$; $\sigma = 1/\theta$. See Barro and Sala-i-Martin (1995, p. 141).

$$\dot{c} = -\frac{u'(c)}{u''(c)}\left[(1/P_1)y'(k) - \frac{n}{(1+npk)P_2}y(k) - \rho\right] = \eta(k,c). \tag{112}$$

Hence, (111)–(112) define a dynamic system in k and c, and completely determine the behavior of the economic system. It is seen in (112) that $\dot{c}$ is a nonlinear function in k, and

$$\lim_{k\to 0}\dot{c} = \infty, \quad \lim_{k\to\infty}\dot{c} = \frac{u'(c)}{u''(c)}\rho \equiv -c\sigma(c)\rho < 0. \tag{113}$$

Hence, the curve $\dot{c} = 0$, (112), can always be solved for a $k = k^*$, giving a vertical line in the (k, c)-space, irrespective of the properties of $y(k)$. For $\dot{k} = 0$, (111) we have, cf. (109)

$$c = \frac{y(k)}{(1+npk)P_2}, \quad \lim_{k\to\infty} c = \frac{1}{nP_1}\lim_{k\to\infty} y'(k) > 0. \tag{114}$$

The curves for $\dot{k} = 0$ and $\dot{c} = 0$ will then always cross one another in the first quadrant, which establishes Theorem 6. □

Remark 7. The sufficiency of the maximum principle can be proved by using the maximized Hamiltonian. It is known that if (k^*, c^*, π^*) is a solution to (111)–(112), and the maximized Hamiltonian is strictly concave, then (k^*, c^*, π^*) is the unique optimal solution; see Leonard and Long (1992). The Hamiltonian for (111) is now given as, cf. (103),

$$\mathcal{H} = u[c(t)] + \pi(t)\left[(1/P_1)y(k) - c/p - nck\right]. \tag{115}$$

Then, the first order condition gives $\pi^* = u'(c^*)/(nk^* + 1/p)$, and $\dot{k} = 0$, (111), gives $c^* = y(k^*)/[(1 + npk^*)P_2]$. Inserting π^* and c^* into the Hamiltonian, (115), we get the maximized Hamiltonian as $\mathcal{H}^* = u[c^*(t)]$. To prove that the maximized Hamiltonian, $\mathcal{H}^*$, is a concave function, we need to calculate its first and second order derivatives. It gives

$$\frac{d\mathcal{H}^*}{dk^*} = u'(c^*)\frac{dc^*}{dk^*}, \qquad \frac{d^2\mathcal{H}^*}{dk^{*2}} = u''(c^*)\frac{dc^*}{dk^*} + u'(c^*)\frac{d^2c^*}{dk^{*2}}. \tag{116}$$

Assume, $\dot{c} = 0$, (112), then we get

$$\rho = \frac{y'(k^*)}{P_1} - \frac{ny(k^*)}{(1+npk^*)P_2} > 0. \tag{117}$$

Using the positivity property of (117) and c^* to calculate $\frac{dc^*}{dk^*}$ and $\frac{d^2c^*}{dk^{*2}}$, respectively, it is then seen that $\frac{dc^*}{dk^*} > 0$, $\frac{d^2c^*}{dk^{*2}} < 0$, therefore $\frac{d\mathcal{H}^*}{dk^*} > 0$, $\frac{d^2\mathcal{H}^*}{dk^{*2}} < 0$; hence, the concavity of the Hamiltonian is proved and the sufficient condition for the existence of the unique optimal solution is accordingly established. ∇

4. General Equilibrium Dynamics of Large Trading Countries

In all the dynamic systems of a *small* trading economy above, the *terms of trade* were *exogenously* given by the world market conditions, and the actual size (parametric value) of the price ratio was important in forming various *parameter regions* related to the dynamic issues of diversification for a small country. The exports (imports) of a *large* country, however, will affect the worldwide demand/supply conditions and the market clearing (equilibrium) prices. Hence, an *endogenous* explanation of the terms of trade requires supplementary determination of the *momentary international equilibrium* in the commodity markets. Moreover, the dynamic laws governing the factor accumulation of a large country, A, will also be affected by the momentary economic state in country B, representing the "rest of the world". Accordingly, the dimensions of the dynamic system have increased, but under some assumptions, the growth processes of the two interacting countries are tractable, and their properties can be compared with those stated in section 2.

4.1 Factor Accumulation and International Equilibrium Prices

Flexible sector technology has been the standard assumption in dynamic two-factor, two-commodity, two-country trade models. From the very beginning, Oniki and Uzawa (1965) dealt with neoclassical sector technologies; see also Kemp (1969), Bardhan (1970), Woodland (1982), Gandolfo (1994). From this literature, it is well known that the *existence, uniqueness*, and global *stability* of the *international growth equilibrium* are ensured with proportional saving and the consumer good at all times being capital intensive in both countries. This global stability result applies, irrespective of long-run diversification or specialization in one or both countries. But our main object is again to obtain some *global stability conditions* that will preserve *diversification* in *both* countries. In this respect, the analysis in the literature has been inconclusive, in particular when their labor force (growing at the same rate) differ much in relative size, and their saving functions are either proportional or classical, see Takayama (1972, p. 406–409, 433). The issue is hard to solve precisely with general production functions, which makes it difficult to delineate the diversification region for the state variables (factor endowment ratios).

To obtain some clues for a general understanding of the role of *country size* and *sectorial factor intensities* in the growth processes of large trading economies, we study the implications of factor accumulation with *fixed coefficient* technologies in both countries. As we allow for *in-*

ternationally different Leontief *technologies*, our approach accordingly examines the dynamics of a two-factor Ricardian trade model. The consequences of internationally identical technologies appear as a special case.

The terms of trade and international equilibrium in the commodity markets will be determined from the *trade balance* equations of the two countries. By Walras' law, we need only consider market equilibrium for one good. Hence, international equilibrium requires that, cf. (18), (20), (subscripts A and B are used in the former single country symbols)

$$X_{1A} = Y_{1A} - Q_{1A} = -X_{1B} = -(Y_{1B} - Q_{1B}). \qquad (118)$$

Let v_A, v_B represent the country *shares* of world labor (population), i.e.,

$$v_A = L_A/(L_A + L_B), \qquad v_A + v_B = 1. \qquad (119)$$

In terms of the factor endowments ratios, k_A, k_B, and the sectorial factor intensities, k_{1A}, k_{2A}, k_{1B}, k_{2B}, the *region* of *diversification*, $C_k^2 \subset \mathbf{R}_+^2$, analogous to (7), is given by

$$\begin{aligned} C_k^2 = {}&]\min(k_{1A}, k_{2A}),\ \max(k_{1A}, k_{2A})[\\ &\times]\min(k_{1B}, k_{2B}),\ \max(k_{1B}, k_{2B})[, \end{aligned} \qquad (120)$$

with the *boundaries* representing complete *specialization* in *one* of the countries. With Leontief sector technologies, C_k^2, (120), is a *diversification rectangle.*
Under the assumptions of competitive economies and proportional saving in section 1, the trade equilibrium condition (118) will determine the international price ratio, $p = P_1/P_2$.

Lemma 2. *For two large competitive trading economies, with proportional saving and Leontief sector technologies, the international equilibrium terms of trade $p = P_1/P_2$ are given by*

$$p = \frac{y_{2A} v_A s_A (k_{2B} - k_{1B})(k_A - k_{1A}) + y_{2B} v_B s_B (k_{2A} - k_{1A})(k_B - k_{1B})}{y_{1A} v_A (1 - s_A)(k_{2B} - k_{1B})(k_{2A} - k_A) + y_{1B} v_B (1 - s_B)(k_{2A} - k_{1A})(k_{2B} - k_B)}, \qquad (121)$$

and, in the special case of internationally identical technologies,

$$\begin{aligned} p &= \frac{P_1}{P_2} = \frac{y_2}{y_1} \frac{v_A s_A \mathrm{l}_{2A} + v_B s_B \mathrm{l}_{2B}}{v_A (1 - s_A) \mathrm{l}_{1A} + v_B (1 - s_B) \mathrm{l}_{1B}} \\ &= \frac{y_2}{y_1} \frac{(v_A s_A + v_B s_B) k_1 - v_A s_A k_A - v_B s_B k_B}{-[v_A (1 - s_A) + v_B (1 - s_B)] k_2 + v_A (1 - s_A) k_A + v_B (1 - s_B) k_B}. \end{aligned} \qquad (122)$$

The feasible domain of the terms of trade surface (TTS), (121)–(122), $p = p(k_A, k_B)$, satisfying (14)–(15) is the entire diversification rectangle, (120).

Proof. From (24) and (118) we have

$$X_{1A} = L_A[(1 - s_A)y_{1A}\mathtt{l}_{1A} - (s_A/p)y_{2A}\mathtt{l}_{2A}], \qquad (123)$$
$$X_{1B} = L_B[(1 - s_B)y_{1B}\mathtt{l}_{1B} - (s_B/p)y_{2B}\mathtt{l}_{2B}], \qquad (124)$$

where $\mathtt{l}_{1A}$ and $\mathtt{l}_{2A}$ are as (6) with $k = k_A$. Using (118) and solving, $X_{1A} = -X_{1B}$, (123)–(124), for the terms of trade, p, we get the expression (121). □

Remark 8. The *terms* of *trade surfaces* (TTS), (121)–(122), have the equation form,

$$p = \frac{a_0 + a_1 k_A + a_2 k_B}{b_0 + b_1 k_A + b_2 k_B}, \qquad (125)$$

which belongs to the family of quadratics in three variables (conic surfaces). The shape of (125) is a *hyperbolic paraboloid*, upon which hyperbolas appear for fixed k_A or fixed k_B. ∇

Traditionally, the terms of trade are determined by the intersection of reciprocal demand (offer) curves, Oniki and Uzawa (1965). In a growth context, the shifting offer curve technique is rather cumbersome. The same applies to the long-run offer curve methodology, Atsumi (1971).

Remark 9. The shape of the terms of trade surface (TTS) – with flexible sector technologies (internationally identical) and no inferior goods of demand in either country – is similar to (122), i.e., having the monotone partial derivatives

$$k_2 > k_1 : \quad \partial p/\partial k_A > 0, \quad \partial p/\partial k_B > 0, \qquad (126)$$
$$k_1 > k_2 : \quad \partial p/\partial k_A < 0, \quad \partial p/\partial k_B < 0. \qquad (127)$$

A rigorous proof is given in Södersten (1964); cf. Rybczynski (1955), Findlay (1959), Kemp (1969). ∇

4.2 Trade, Growth, and International Equilibrium Dynamics with Proportional Saving and Leontief Sector Technologies

With proportional saving and the same exogenous labor growth, the *dynamics* of the *factor endowment ratios* of two *large trading* economies, is described by, cf. (27),

$$\dot{k}_A = s_A[y_{1A}\mathtt{l}_{1A} + (y_{2A}/p)\mathtt{l}_{2A}] - nk_A = F(k_A, k_B), \qquad (128)$$
$$\dot{k}_B = s_B[y_{1B}\mathtt{l}_{1B} + (y_{2B}/p)\mathtt{l}_{2B}] - nk_B = G(k_A, k_B), \qquad (129)$$

where

$$\mathrm{l}_{1A} = \frac{k_A - k_{2A}}{k_{1A} - k_{2A}}, \qquad \mathrm{l}_{2A} = 1 - \mathrm{l}_{1A}, \tag{130}$$

$$\mathrm{l}_{1B} = \frac{k_B - k_{2B}}{k_{1B} - k_{2B}}, \qquad \mathrm{l}_{2B} = 1 - \mathrm{l}_{1B}, \tag{131}$$

$$p = \frac{y_{2A} v_A s_A (k_{2B} - k_{1B})(k_A - k_{1A}) + y_{2B} v_B s_B (k_{2A} - k_{1A})(k_B - k_{1B})}{y_{1A} v_A (1 - s_A)(k_{2B} - k_{1B})(k_{2A} - k_A) + y_{1B} v_B (1 - s_B)(k_{2A} - k_{1A})(k_{2B} - k_B)}, \tag{132}$$

$$k_A, k_B \in C_k^2. \tag{133}$$

Evidently, (128)–(129) represent a nonlinear and *nonhomogenous* dynamic system in the state variables, k_A and k_B. Moreover, many parameters are involved in a qualitative characterization of the *governing functions*, (F, G). One important objective is to obtain the *parameter restrictions* upon F and G that ensure a unique *international* steady state, which is a global *attractor* located in the *interior* of C_k^2, (120). Economically, we get the global *stability conditions* that allow for long-run *diversification* in *both* countries.

The intersection of the two curves (nullclines), the ($\dot{k}_A = 0$)-curve and the ($\dot{k}_B = 0$)-curve, is the *international equilibrium* (steady state) point, (κ_A, κ_B). As it may be verified from the equations (128)–(132), the *nullclines*, $F(k_A, k_B) = 0$, $G(k_A, k_B) = 0$ are both *quadratic* (conics) in the state variables. In general, $F(k_A, k_B) = 0$, is a hyperbola (including degenerate ones) with a *vertical* asymptote and an *oblique* one, whereas $G(k_A, k_B) = 0$ is a hyperbola (including degenerate ones) with a *horizontal* asymptote and an *oblique* one. These nullclines are shown in figure 7. Further analysis of the vector field (F, G), (128)–(129), gives:

Theorem 7. *For two large, competitive, trading economies, with the same exogenous labor growth, and with different proportional saving and different Leontief technologies, the necessary and sufficient conditions – for the existence, uniqueness, and the global stability of diversified steady state growth in both countries – are given by the joint conditions:*

$$(i) \qquad k_{2A} > k_{1A}, \quad k_{2B} > k_{1B}, \tag{134}$$

$$(ii) \qquad s_A > \frac{nk_{1A}}{y_{1A}}, \quad s_B > \frac{nk_{1B}}{y_{1B}}, \tag{135}$$

$$(iii) \qquad \frac{nk_{2B}}{y_{1A}(1 - s_A)} \geq \frac{v_A}{v_B}, \quad \frac{nk_{2A}}{y_{1B}(1 - s_B)} \geq \frac{v_B}{v_A}. \tag{136}$$

If only (134) is reversed, and (135)–(136) are maintained, then the unique interior steady state is changed from a global attractor to a global repeller, and hence, at least one country will be completely specialized.

With internationally identical technologies, (134)–(136) become

$$(i) \quad k_2 > k_1, \tag{137}$$

$$(ii) \quad s_A > \frac{nk_1}{y_1}, \quad s_B > \frac{nk_1}{y_1}, \tag{138}$$

$$(iii) \quad \frac{nk_2}{y_1(1-s_A)} \geq \frac{v_A}{v_B}, \quad \frac{nk_2}{y_1(1-s_B)} \geq \frac{v_B}{v_A}. \tag{139}$$

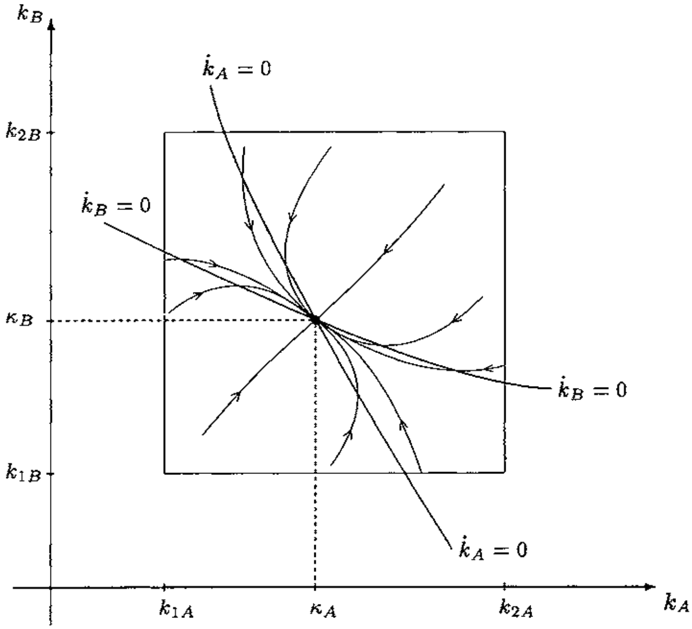


Fig. 7. Phase portrait of (128)–(129) and the conditions (137)–(139)

Proof. The proof is based upon invariance properties of C_k^2. The diversification rectangle, C_k^2, is said to be *positively invariant*, if any *solution* $[k_A(t), k_B(t)]$ to the system (128)–(129) that starts in C_k^2 remains in C_k^2 forever. Likewise, C_k^2 is *negatively invariant* if any solution to (128)–(129), whenever observed in C_k^2, must have been there in the entire past. It can be shown that *positive* invariance of C_k^2 is equivalent to (134)–(136) including equalities in (135). Likewise, C_k^2 is *negatively* invariant iff (135)–(136) are satisfied, including equalities in (135), and both inequalities (134) are *reversed*. Furthermore, if C_k^2 is positively invariant and the inequalities (135) are *strict*, then there is only *one* steady state in C_k^2, and this steady state is a *global* attractor. If only (134) is reversed (negative invariance), then the interior steady state is a global *repeller*. The formal proof is given in Appendix C. □

The necessary and sufficient *parameter conditions* (134)–(136) or (137)–(139) allow opportunities for *free trade* and *diversification* in *both countries* to coexist and be maintained during the *transition* dynamics towards a unique, *diversified*, *international* steady state. As is well known, the one-factor Ricardian trade model does not admit such long-run diversification.

It is observed from (134)–(136) that larger saving rates, s_A, s_B, and larger capital intensities, k_{2A}, k_{2B}, contribute to satisfying the joint conditions. Moreover, if e.g. the share v_A is large, then a large s_A helps to meet the joint conditions. However, the range of (n/y_{1A}) and (n/y_{1B}) has to be restricted. It is seen from parameter conditions (136) that the relative size of the two countries must not differ too much.

In comparing (135)–(136) with (138)–(139), the former conditions are evidently more easily met than the latter, because of greater flexibility in parameter variations. Thus, with accumulating factor endowments and free international trade, it is more difficult to sustain long-run diversification in both countries when the sector technologies are internationally identical. Figure 7 is drawn for the *Heckscher-Ohlin* case of identical technologies, and C_k^2 is a square.

When the inequalities (134), (137) are reversed, it is generally not possible to indicate which country will be specialized and in which good, as any of the four boundaries in figure 7 can be attained, depending on the initial values (factor endowment ratios) of the two countries. The stability conditions (137)–(139) and the TTS surface (122) are closely linked to the Rybczynski theorem.

The dynamic international equilibrium model (128)–(133) of two large countries *trading* in two goods may be seen as an *extension* of the *closed* economy (autarky) growth models with Leontief technologies that *initiated* the *dynamic analysis* of *growing* two-sector economies, see Shinkai (1960), Jones (1965), Corden (1966), Stiglitz and Uzawa (1969), Ramanathan (1973, 1975), Jensen (1994), and Gandolfo (1996).

It may be observed that some *parameter* (sectorial factor intensity) *conditions*, (134), for an attractor in a simple dynamic general equilibrium *trade* model with two *large* countries – in contrast, cf. Remark 1, to a *small* trading country (for which international trade is indispensable for growth) – are *similar* to corresponding parameter conditions of an *aggregated* (global, closed) two-sector economy. The relative size of countries (136) are naturally involved in maintaining diversification within trading subunits of the world economy.

5. Final Comments

We have analyzed *factor accumulation* processes for small and large two-sector economies that *trade* freely in both goods. As to *capital formation* in a small country, the implications of proportional, classical, and optimal saving were examined in combination with assumptions of exogenous or endogenous *labor growth*. Regarding the issues of long-run *diversification* and *persistent* (endogenous) *per capita growth*, a comparison of the theoretical economic results is presented in table 1.

Labor growth Saving behaviour	Exogenous labor growth	Endogenous labor growth
Proportional saving	Diversification: Limited range. Persistent growth: Possible.	Diversification: Limited range. Persistent growth: Not possible.
Classical saving	Diversification: Singularity. Persistent growth: Possible.	Diversification: Limited range. Persistent growth: Not possible.
Optimal saving	Diversification: Singularity. Persistent growth: Possible.	Diversification: Limited range. Persistent growth: Not possible.

Table 1. Dynamics of a small trading economy with constant returns to scale technologies

The unifying mathematical structure of the basic growth models for small trading economies, diversified or specialized, was planar homogenous dynamic systems with labor and capital as the state variables. It encompasses the dynamics of classical, neoclassical, and some newer, endogenous per capita growth models. The regime of *exogenous population* (labor) growth and *high substitution* elasticities between labor and capital only occurs in later (*modern*) stages of economic development.

Of course, we have neglected many aspects of expanding factor endowments and factor reallocations that affect the growth of trading economies. Some extensions may introduce variable returns to scale, human capital, technical progress, presence of non-tradeable goods, disequilibrium dynamics, and uncertainty.

We hope, however, that the theorems and propositions on the general equilibrium *dynamics* of basic *two-dimensional trade* models contribute to building a theoretical framework and a *benchmark* against which the results of extensions and multidimensional dynamic trade models can be appraised.

Appendix A: Endogenous Growth and Inessential Labor

Consider two *homogenous* C^1-class functions, (F, G), of *degree one* as *governing functions* of dynamic systems in the plane

$$\dot{x} = F(x,y) = xF(1, y/x) = xf(r), \tag{A.1}$$
$$\dot{y} = G(x,y) = xG(1, y/x) = xg(r)\dot{r} = h(r) = g(r) - rf(r), \tag{A.2}$$

and

$$\dot{r} = h(r) = g(r) - rf(r). \tag{A.3}$$

Definition A. *Endogenous (persistent) growth of any rapidity:*

$$h(r) > 0, \quad \forall r > r_0. \tag{A.4}$$

From (A.3), we have, cf. (A.1)–(A.2)

$$\begin{aligned} \forall r \geq r_0, \quad h(r) > 0 \;&\Leftrightarrow\; g(r) > rf(r) \\ &\Leftrightarrow\; xG(x,y) > yF(x,y), \quad \forall y \geq r_0 x. \end{aligned} \tag{A.5}$$

Lemma 1A. *A necessary condition for endogenous growth is:*

$$F(0,y) \leq 0, \quad \forall y \geq 0. \tag{A.6}$$

Lemma 1A is implied by the last inequality of (A.5). If Lemma 1A is violated, there will be a root (steady state) in (A.3), which globally *prevents* $r(t) \to \infty$, i.e., endogenous growth.

Lemma 2A. *With a homogenous dynamic system of degree one (A.1)–(A.2), and the first state variable growing exogenously* $\forall x \geq 0$: $F(x,y) = \mathcal{F}(x) = xf(r) = xF(1,0) = xc_0$, $c_0 > 0$, *then a necessary condition for endogenous growth is*, $\forall y > 0$: $G(0,y) > 0$.

Proof. By (A.5) and exogeneity above, we get: $G(x,y) > yc_0$, $\forall y \geq r_0 x$, which requires $G(0,y) > 0$ for $y > 0$, as stated in Lemma 2A. □

Remark A. With exogeneity and homogeneity of degree higher than one, we can have persistent growth, without the condition, $\forall y > 0$: $G(0,y) > 0$. Example: $F(x,y) = 1/2x^2$, $G(x,y) = xy$; $x \geq 0$, $y \geq 0$, we have, $h(r) = r - 1/2r = 1/2r$, which implies $r(t) \to \infty$, even with $G(0,y) = 0$. But with a higher degree than one, the solutions $r(t)$ explode (infinite in finite time), cf. Jensen (1992, p. 190), and we have problems raised by Solow (1994). ∇

Appendix B: Proof of Theorem 5

We consider the dynamic system, cf. (102) and (107),

$$\dot{k} = h(k,c) = y(k)/P_1 - nk - c/p, \tag{B.1}$$

$$\dot{c} = \eta(k,c) = c\,\sigma(c)[y'(k)/P_1 - (n+\rho)], \tag{B.2}$$

in the *closed* first quadrant, $\overline{\mathbf{R}}_+^2$. The constants P_1, p, n, and ρ, are assumed to be positive. The positive intertemporal elasticity of substitution $\sigma(c) = -u'(c)/[cu''(c)]$ depends on the utility function $u(c)$.

Assumption B. *The per capita GNP-function $y(k)$ has the continuity and differentiability properties as follows,*

$$(i) \quad y(k) \in C^0([0,\infty[) \cap C^1(]0,\infty[), \qquad (ii) \quad y(0) \geq 0. \tag{B.3}$$

It is further assumed that

$$(iii) \quad \forall k > 0: \quad y'(k)/P_1 > n + \rho. \tag{B.4}$$

For a concave GNP-function with $y(k) \to \infty$ as $k \to \infty$, (B.4) becomes

$$(iv) \quad \lim_{k\to\infty} y'(k)/P_1 = \beta/P_1 > n + \rho. \tag{B.5}$$

It follows from (B.3)–(B.4) or (B.3) and (B.5) that the system (B.1)–(B.2) has no stationary solutions in $\overline{\mathbf{R}}_+^2$ [except possibly for $(0,0)$], and that the positive k-axis ($c = 0$) is a trajectory (orbit).

Lemma 1B. *If there exists $\delta > 0$ and $k_0 > 0$ such that $\forall k \geq k_0, \forall c > 0$:*

$$\frac{y(k)/P_1}{k} - n - \sigma(c)\left[\frac{y'(k)}{P_1} - n - \rho\right] \geq \delta, \tag{B.6}$$

then there exists to the system (B.1)–(B.2) an orbit–$\Gamma(t) \equiv [k^*(t), c^*(t)]$, $t \in \mathbf{R}$, such that $k^*(t) \to \infty$, $c^*(t) \to \infty$, as $t \to \infty$ – which separates the first quadrant into two regions I and II in figure 6.2. An orbit starting in the lower region I has the same behavior as Γ for $t \to \infty$, whereas an orbit starting in the upper region II eventually meets the c-axis, $k = 0$.

Proof. Consider the region $W_\alpha = \{(k,c) \mid 0 \leq c \leq \alpha k \wedge k \geq k_0\}$, where α is a positive constant chosen such that W_α becomes *positively invariant*, cf. figure B.

Since the vector field (B.1)–(B.2) is directed inward on the line $k = k_0$, and since the positive k-axis is a trajectory, the region W_α is positive

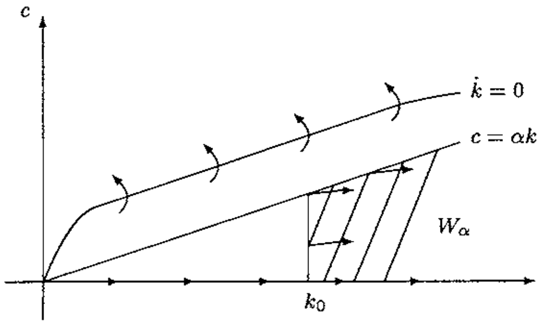


Fig. B. The positive invariant region, W_α, with endogenous (persistent) per capita growth

invariant iff the vector field points inward on the line $c = \alpha k$, $(k > k_0)$. Since the inward pointing normal to the line $c = \alpha k$ is $(\alpha, -1)$, we require

$$\alpha h(k, \alpha k) - \eta(k, \alpha k) > 0, \quad \text{for } k \geq k_0. \tag{B.7}$$

Inserting the expressions for h, (B.1), and η, (B.2), into (B.7), and simplifying, we find the requirement

$$\frac{\alpha}{p} < \frac{y(k)}{P_1 k} - n - \sigma(\alpha k)\left[\frac{y'(k)}{P_1} - (n+\rho)\right] \equiv R(k), \quad \text{for } k \geq k_0. \tag{B.8}$$

A positively invariant region W_α (with some $\alpha > 0$) exists iff $R(k)$ is bounded from below by a positive constant. By (B.6), we have for $k \geq k_0$

$$R(k) \geq \delta > 0. \tag{B.9}$$

Choose α/p to be any positive constant less than δ. Then W_α is positively invariant. For any orbit in the *open* first quadrant, $\mathbf{R}^2_+$, we have by (B.4) that $\dot{c} > 0$. Accordingly, it follows that any orbit starting in W_α must satisfy $k(t) \to \infty$, $c(t) \to \infty$, as $t \to \infty$. Any orbit in $\mathbf{R}^2_+$ must either behave as just characterized (class I) or cross the $\dot{k} = 0$ nullcline (class II). In the latter case, the orbit will meet the c-axis eventually, since otherwise $c(t) \to \infty$ and $k(t) \to k_\varepsilon$ as $t \to \infty$, for some $k_\varepsilon \geq 0$. For t sufficiently large, i.e., k sufficiently small, say $0 < k \leq k^*$, we have from (B.1)–(B.2)

$$-\frac{dc}{dk} = \frac{\sigma(c)[y'(k)/P_1 - n - \rho]}{-(1/c)[y(k)/P_1 - nk] + 1/p} \leq \frac{\bar{\sigma} y'(k)/P_1}{1/(2p)} \equiv a y'(k), \tag{B.10}$$

where $\bar{\sigma}$ is an upper bound for $\sigma(c)$. This immediately rules out $k_\varepsilon > 0$ since then $-\frac{dc}{dk}$ would be bounded above by a constant. If $k_\varepsilon = 0$, we find by integrating from k to k^* for $0 < k < k^*$

$$c(k) \leq c(k^*) + ay(k^*), \quad 0 < k \leq k^*, \tag{B.11}$$

contradicting $c(k) \to \infty$ as $k \to 0^+$.

To get the *separating orbit*, Γ, consider a curve, C, connecting $(k, c) = (1, 0)$ with $(k, c) = (0, 1)$ and intersecting the nullcline $\dot{k} = 0$ once (think of a circle). We can write $C = C_I \cup C_{II} \cup \{(1,0),(0,1)\}$ where C_I and C_{II} consists of the points through which pass orbits of class I and II, respectively. C_{II} must be an *open* and *connected* part of C. Since C_I and C_{II} are both non-empty, $C_I \cup \{(1,0)\}$ must be *closed*. The separating orbit Γ goes through the end point of C_I. □

A powerful and useful extension of Lemma 1B is the simpler separator condition stated in:

Corollary 1B. *With the assumptions of (B.3) and (B.5), the sufficient conditions for existence of the separating orbit, $\Gamma(t)$, cf. Lemma 1B, is given by the restriction*

$$\overline{\sigma} = \sup_{c>0} \sigma(c) < \frac{\beta/P_1 - n}{\beta/P_1 - (n+\rho)}, \tag{B.12}$$

where $\overline{\sigma}$ is the upper bound of the intertemporal substitution elasticity $\sigma(c)$ of $u(c)$ and where β/P_1 is given in (B.5).

Proof. Since $y(k) \to \infty$ as $k \to \infty$, it follows from (B.5) and l'Hospital that $y(k)/k \to \beta$ as $k \to \infty$. Thus for any number $\varepsilon > 0$, there exists a number, k_ε, such that for $k > k_\varepsilon$, we have, cf. (B.6)

$$\begin{aligned}
&\frac{y(k)}{P_1\, k} - n - \sigma(c)\left[\frac{y'(k)}{P_1} - n - \rho\right] \\
&\geq \frac{\beta - \varepsilon}{P_1} - n - \sup_{c>0} \sigma(c)\left[\frac{\beta + \varepsilon}{P_1} - n - \rho\right] \\
&= \frac{\beta}{P_1} - n - \overline{\sigma}\left[\frac{\beta}{P_1} - n - \rho\right] - \frac{\varepsilon}{P_1}(1 + \overline{\sigma}) \equiv \delta.
\end{aligned} \tag{B.13}$$

By assumption (B.5) and (B.12), the sum of the first three terms of δ is positive. Thus, by choosing $\varepsilon > 0$ sufficiently small, also δ is positive. Thus, the requirements of Lemma 1B are satisfied, and hence the separating orbit Γ exists. □

Lemma 2B. *If any selected utility function $u(c)$ is assumed to satisfy*

$$\forall c \geq c_0 \geq 0: \quad 0 \leq u(c) \leq Ac, \tag{B.14}$$

where A is any positive constant, then the convergence of the integral $U = \int_0^\infty u[c(t)]e^{-\rho t}dt$ for a solution $[k(t), c(t)]$ to the system (B.1)–(B.2) is assured, if

$$\overline{\sigma} = \sup_{c>0} \sigma(c) < \frac{\rho}{\beta/P_1 - (n+\rho)}, \tag{B.15}$$

where β is given by (B.5).

Proof. Let $\varepsilon > 0$. Choose k_ε such that $y'(k) < \beta + \varepsilon$ for $k \geq k_\varepsilon$. Choose t_ε such that $k(t) > k_\varepsilon$ for $t > t_\varepsilon$. Then from (B.2), we find

$$\forall t \geq t_\varepsilon : \quad \dot{c} < c \sup_{c>0} \sigma(c)[(\beta+\varepsilon)/P_1 - (n+\rho)] \equiv \alpha c. \tag{B.16}$$

It follows from (B.16) that $c(t) \leq c(t_\varepsilon)e^{\alpha(t-t_\varepsilon)}$, and hence, cf. (B.14)

$$u[c(t)]e^{-\rho t} \leq Ac(t)e^{-\rho t} \leq Ac(t_\varepsilon)e^{-\alpha t_\varepsilon}e^{-(\rho-\alpha)t}. \tag{B.17}$$

Thus the convergence of the integral U is assured, if $\alpha < \rho$, which by (B.16) says

$$\overline{\sigma} = \sup_{c>0} \sigma(c) < \frac{\rho}{(\beta+\varepsilon)/P_1 - (n+\rho)}. \tag{B.18}$$

With $\varepsilon > 0$ chosen sufficiently small, the requirement (B.18) can be satisfied by the condition (B.15). □

Remark B. The condition (B.15) is stronger that (B.12) of Corollary 1B, since by assumption (B.5), we have

$$\frac{\rho}{\beta/P_1 - (n+\rho)} < \frac{\beta/P_1 - n}{\beta/P_1 - (n+\rho)}. \tag{B.19}$$

In short, the existence of separating orbit Γ is assured by $\overline{\sigma} < 1$, but $\overline{\sigma} < 1$ does not itself ensure convergence of U. However, for isoelastic $u(c)$ with $\forall c, \sigma(c) = \sigma$ (constant), it can be verified that the convergence of U is in fact also ensured by the existence condition of the separating orbit, (B.12).
Indeed, with constant intertemporal elasticity of substitution, the separating orbit in figure 6.2 is the optimal solution $[k^*(t), c^*(t)]$ satisfying the transversality condition; see hereto Gandolfo (1996, p. 390).
It remains to be seen how (B.15) may be relaxed for general non-isoelastic $u(c)$ in Ramsey problems. ∇

Appendix C: Proof of Theorem 7

The necessary and sufficient conditions for invariance of the diversification region, C_k^2, (133) and global stability of (128)–(129) are obtained as follows.

C.1 Invariance of C_k^2

To prove *positive invariance*, we first show that the vector field is directed inward on the boundary of C_k^2 iff the inequalities (134)–(136) hold with strict inequalities.
The boundary of C_k^2 consists of four line segments, leading to the following *four requirements* for positive invariance, with $\sigma_A = \text{sign}\,(k_{2A} - k_{1A})$, and $\sigma_B = \text{sign}\,(k_{2B} - k_{1B})$.

$$\sigma_A F(k_{1A}, k_B) > 0 \text{ for all } k_B \in I_B, \tag{C.1}$$

$$\sigma_A F(k_{2A}, k_B) < 0 \text{ for all } k_B \in \overline{I}_B, \tag{C.2}$$

$$\sigma_B G(k_A, k_{1B}) > 0 \text{ for all } k_A \in I_A, \tag{C.3}$$

$$\sigma_B G(k_A, k_{2B}) < 0 \text{ for all } k_A \in \overline{I}_A, \tag{C.4}$$

where I_A is the half-open interval between k_{1A} and k_{2A}, including k_{2A} but not k_{1A}. $\overline{I}_A$ is its closure. I_B and $\overline{I}_B$ are defined similarly.
Next, (C.1) and (C.3) are satisfied iff, cf. (128), (130), (132)

$$\sigma_A(y_{1A}s_A - nk_{1A}) > 0 \text{ and } \sigma_B(y_{1B}s_B - nk_{1B}) > 0, \tag{C.5}$$

which gives (135) with $\sigma_A > 0$, $\sigma_B > 0$, cf. below.
It can be seen that, cf. (128), (132)

$$F(k_{2A}, k_B) = \frac{s_A y_{2A} v_B y_{1B}(1 - s_B)(k_{2B} - k_B) - nk_{2A}D}{D}, \tag{C.6}$$

where the denominator, D, is given by

$$D = v_A s_A y_{2A}(k_{2B} - k_{1B}) + v_B s_B y_{2B}(k_B - k_{1B}). \tag{C.7}$$

Notice that D has the same sign as $k_{2B} - k_{1B}$.
The numerator is a polynomial of degree one in k_B, so it suffices that

$$\sigma_A \sigma_B [s_A y_{2A} v_B y_{1B}(1 - s_B)(k_{2B} - k_B) - nk_{2A}D] < 0, \tag{C.8}$$

for $k_B = k_{1B}$ and for $k_B = k_{2B}$. This gives the following two requirements

$$(i) \quad \sigma_A \sigma_B [-nk_{2A}v_A + v_B y_{1B}(1 - s_B)] s_A y_{2A}(k_{2B} - k_{1B}) < 0, \tag{C.9}$$

$$(ii) \quad \sigma_A \sigma_B (-nk_{2A})(v_A s_A y_{2A} + v_B s_B y_{2B})(k_{2B} - k_{1B}) < 0. \tag{C.10}$$

Using that $\sigma_B(k_{2B} - k_{1B}) > 0$, we reduce (C.9) to

$$\sigma_A[-nk_{2A}v_A + v_B y_{1B}(1 - s_B)] < 0, \tag{C.11}$$

and (C.10) to the simple requirement $\sigma_A > 0$, i.e., $k_{2A} > k_{1A}$. Hence with $\sigma_A > 0$, $\sigma_B > 0$, i.e., (134), then (C.10)-(C.11) give (136).

Requirement (C.4) similarly leads to the two requirements

$$\sigma_B > 0 \text{ and } -nk_{2B}v_B + v_A y_{1A}(1 - s_A) < 0. \tag{C.12}$$

Finally, observe that when $s_A = \frac{nk_{1A}}{y_{1A}}$ (i.e., equality in one of the inequalities of (135) or (C.5)), then $F(k_{1A}, k_B) = 0$ for all k_B. Thus the line $k_A = k_{1A}$ *consists of orbits*. When $s_B = \frac{nk_{1B}}{y_{1B}}$, the line $k_B = k_{1B}$ consists of orbits.

When $\frac{nk_{2A}}{y_{1B}(1-s_B)} = \frac{v_B}{v_A}$ – equality in one of (136) or (C.11) – then the vector field is directed inward on the segment $k_A = k_{2A}$ for $k_B \in]k_{1B}, k_{2B}]$ and is parallel to the segment $k_B = k_{1B}$. An analogous statement holds when $\frac{nk_{2B}}{y_{1A}(1-s_A)} = \frac{v_A}{v_B}$.

We conclude that the system is *positively invariant*, also if one or several of the inequalities (135)–(136) are in fact equalities.

C.2 Global Stability

We show that if C_k^2 is positively invariant, and if the inequalities (135) are strict, then there is one equilibrium in C_k^2, and this equilibrium is a global attractor.

First notice that the nullcline $\dot{k}_A = 0$, i.e., $F(k_A, k_B) = 0$, is a hyperbola. It has the vertical asymptote

$$k_A = k_{1A} + \frac{(k_{2A} - k_{1A})(s_A y_{1A} - nk_{1A})s_B y_{2B}}{ns_B y_{2B}(k_{2A} - k_{1A}) + s_A(y_{2A}y_{1B}(1 - s_B) + y_{1A}y_{2B}s_B)}, \tag{C.13}$$

which lies between $k_A = k_{1A}$ and $k_A = k_{2A}$.

Earlier, we found that $F(k_{2A}, k_B) < 0$ for all $k_B > k_{1B}$. Let k_{B0} be the root of $D = 0$ (D is given in the previous section). Then $k_{B0} < k_{1B}$, and we see that $F(k_{2A}, k_B) > 0$ for k_B close to k_{B0} with $k_B > k_{B0}$. We conclude that one branch of the hyperbola $F(k_A, k_B) = 0$ passes through the line segment $k_A = k_{2A}$, $k_B \leq k_{1B}$. This same branch has the upper half-line given by (C.13) as its asymptote. The other branch of that hyperbola goes through (k_{1A}, k_{1B}) and has the lower half-line given by (C.13) as its asymptote.

Since analogous statements can be made about the nullcline $\dot{k}_B = 0$, the nullclines are located as shown in figure C. We conclude that there is a unique equilibrium in C_k^2. This equilibrium is a global attractor (and a node). When $\frac{nk_{1A}}{y_{1A}}$ approaches s_A from below, the equilibrium moves to the boundary $k_A = k_{1A}$. An analogous statement holds with A replaced by B.

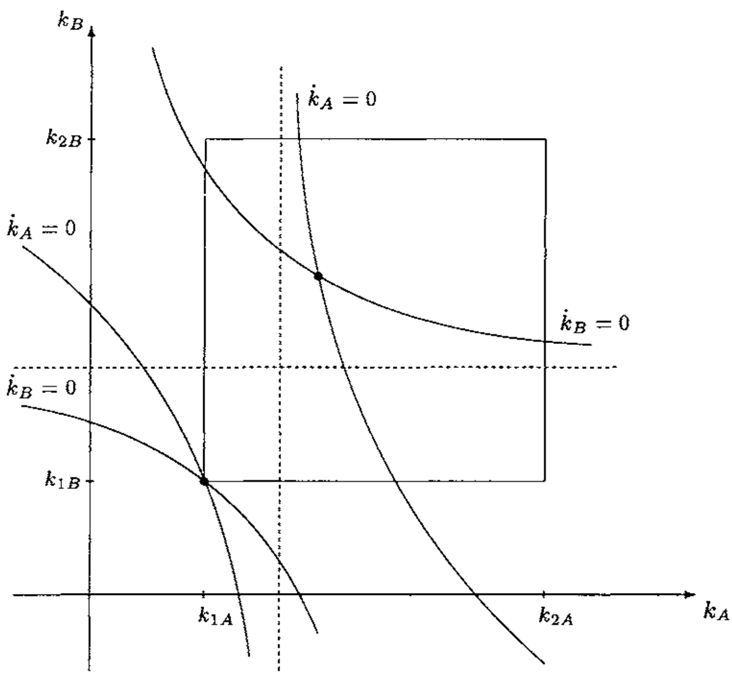


Fig. C. The shape and location of nullclines, $\dot{k}_A = 0$, $\dot{k}_B = 0$, for the dynamic system (128)–(129) of two large trading economies

Remark. Appendices A, B, and C are joint works with Preben Kjeld Alsholm, Technical University of Denmark. ∇

Acknowledgement. Bjarne S. Jensen's research was supported by the Danish Social Sciences Research Council, grant no. 9500740.

We would like to thank Giancarlo Gandolfo, Alan V. Deardorff, and Kar-yiu Wong for their very detailed comments and suggestions. Useful comments were received from Peter Skott, Harald Preissler, Niels Henrik Børjesson, Jimmie Thomsen, and other participants of the Denmark conference.

References

Atsumi, Hiroshi (1971). "The Long-Run Offer Function and a Dynamic Theory of International Trade." *Journal of International Economics* 1, 267–299.

Bardhan, Pranab K. (1970). *Economic Growth, Development, and Foreign Trade.* New York: Wiley-Interscience.

— (1966). "On Factor Accumulation and the Pattern of International Specialisation." *Review of Economics Studies* 33, 39–44.

— (1965a). "Equilibrium Growth in the International Economy." *Quarterly Journal of Economics* 79, 455–464.

— (1965b). "International Differences in Production Functions, Trade and Factor Prices." *Economic Journal* 75, 81–87.

Barro, Robert J., and Xavier Sala-i-Martin (1995). *Economic Growth.* New York: McGraw-Hill.

Becker, Robert A., and Edwin Burmeister (1991). *Growth Theory.* Vol. I–III. Aldershot: Edward Elgar.

Becker, Robert A., and Ciprian Foias (1987). "A Characterization of Ramsey Equilibrium." *Journal of Economic Theory* 41, 173–184. Reprinted in Becker and Burmeister (1991).

Bertola, Giuseppe (1996). "Factor Shares in OLG Models of Growth." *European Economic Review* 40, 1541–1560.

— (1993). "Factor Shares and Savings in Endogenous Growth." *American Economic Review* 83, 1184–1190.

Bhagwati, Jagdish N. (1981) (ed.). *International Trade: Selected Readings.* 2. ed. Cambridge, MA: MIT Press.

Cass, David (1965). "Optimum Growth in an Aggregative Model of Capital Accumulation." *Review of Economic Studies* 32, 233–240. Reprinted in Becker and Burmeister (1991).

Corden, W. Max (1966). "The Two-Sector Growth Model with Fixed Coefficients." *Review of Economic Studies* 33, 253–262.

Deardorff, Alan V. (1994). "Growth and International Investment with Diverging Populations." *Oxford Economic Papers* 46, 477–491.

— (1974). "A Geometry of Growth and Trade." *Canadian Journal of Economics* 2, 295–306.

— (1973). "The Gains from Trade in and out of Steady-State Growth." *Oxford Economic Papers* 25, 173–191.

— (1971). *Growth and Trade in a Two-Sector World.* Unpublished Ph.D. dissertation, Cornell University.

— and James A. Hanson (1978). "Accumulation and a Long-Run Heckscher-Ohlin Theorem." *Economic Inquiry* 16, 288–292.

Deaton, Angus (1992). *Understanding Consumption.* Oxford: Clarendon Press.

Drabicki, John Z., and Akira Takayama (1975). "On the Optimal Growth of the Two-Sector Economy." *Keio Economic Studies* 12, 1–36.

Ethier, Wilfred, and Stephen A. Ross (1971). "International Capital Movements and Long Run Diversification." *Journal of International Economics* 1, 301–314.

Findlay, Ronald (1984). "Growth and Development in Trade Models." In Jones and Kenen (1984).

— (1970). "Factor Proportions and Comparative Advantage in the Long Run." *Journal of Political Economy* 78, 27–34. Reprinted in Bhagwati (1981).

— (1959). "Factor Intensities, Technological Progress and the Terms of Trade." *Oxford Economic Papers* 11, 111–121.

Galor, Oded, and Shoukang Lin (1997). "Dynamic Foundations for the Factor Endowment Model of International Trade." (this volume).

Gandolfo, Giancarlo (1996). *Economic Dynamics: Methods and Models.* 3. ed. Berlin: Springer-Verlag.

— (1994). *International Economics I: The Pure Theory of International Trade.* 2. ed. Berlin: Springer-Verlag.

Grossman, Gene M., and Kenneth Rogoff (1995) (eds.). *Handbook of International Economics.* Vol. III. Amsterdam: Elsevier.

Hall, Robert E. (1988). "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96, 339–357.

Hamada, Koichi (1974). "An Economic Analysis of the Duty-Free Zone." *Journal of International Economics* 4, 225–241.

Jensen, Bjarne S. (1994). *The Dynamic Systems of Basic Economic Growth Models.* Dordrecht: Kluwer Academic.

— and Mogens Esrom Larsen (1987). "Growth and Long-Run Stability." *Acta Applicandae Mathematicae* 9, 219–237. Reprinted in Jensen (1994).

Johnson, Harry G. (1959). "Economic Development and International Trade." *Nationaløkonomisk Tidsskrift* 97, 253–272.

Jones, Larry E., and Rodolfo E. Manuelli (1990). "A Convex Model of Equilibrium Growth: Theory and Policy Implications." *Journal of Political Economy* 98, 1008–1038.

Jones, Ronald W. (1965). "The Structure of Simple General Equilibrium Models." *Journal of Political Economy* 73, 557–572.

— and Peter B. Kenen (1984) (eds.). *Handbook of International Economics.* Vol. I. Amsterdam: North-Holland.

Kemp, Murray C. (1969). *The Pure Theory of International Trade and Investment.* New Jersey: Prentice-Hall.

Klundert, Theo van de, and Sjak Smulders (1992). "Reconstructing Growth Theory: A Survey." *De Economist* 140, 177–203.

Kuznets, Simon (1966). *Modern Economic Growth. Rate, Structure, and Spread.* New Haven: Yale University Press.

Leamer, Edward E., and James Levinsohn (1995) (eds.). "International Trade Theory: The Evidence." Chapter 26 in Grossman and Rogoff (1995).

Leonard, David, and Ngo Van Long (1992). *Optimal Control Theory and Static Optimization in Economics.* Cambridge University Press.

Lucas, Robert E. (1993). "Making a Miracle." *Econometrica* 61, 251–272.

— (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22, 3–42. Reprinted in Becker and Burmeister (1991).

Long, Ngo Van, and Kar-yiu Wong (1997). "Endogenous Growth and International Trade: A Survey." (this volume).

Niehans, Jürgen (1963). "Economic Growth with Two Endogenous Factors." *Quarterly Journal of Economics* 77, 349–371.

Obstfeld, Maurice (1989). "Fiscal Deficits and Relative Prices in a Growing World Economy." *Journal of Monetary Economics* 23, 461–484.

— (1982). "Aggregate Spending and the Terms of Trade: Is There a Laursen-Metzler Effect?" *Quarterly Journal of Economics* 97, 251–270.

— and Kenneth Rogoff (1995). "The Intertemporal Approach to the Current Account." Chapter 34 in Grossman and Rogoff (1995).

Ohlin, Bertil (1933). *Interregional and International Trade.* Cambridge, MA: Harvard University Press.

Oniki, Hajime, and Hirofumi Uzawa (1965). "Patterns of Trade and Investment in a Dynamic Model of International Trade." *Review of Economic Studies* 32, 15–38.

Pitchford, John D. (1960). "Growth and the Elasticity of Factor Substitution." *Economic Record* 36, 491–504.

Rebelo, Sergio (1991). "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy* 99, 500–521.

Ramanathan, R. (1975). "The Elasticity of Substitution and the Speed of Convergence in Growth Models." *Economic Journal* 85, 612–613.

— (1973). "Adjustment Time in the Two-Sector Growth Model with Fixed Coefficients." *Economic Journal* 83, 1236–1244.

Romer, Paul M. (1994). "The Origins of Endogenous Growth." *Journal of Economic Perspectives* 8, 3–22.

— (1986). "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94, 1002–1037. Reprinted in Becker and Burmeister (1991).

Rybczynski, Tad M. (1955). "Factor Endowments and Relative Commodity Prices." *Economica* 22, 336–341.

Saint-Paul, Gilles (1992). "Fiscal Policy in an Endogenous Growth Model." *Quarterly Journal of Economies* 107, 1243–1259.

Samuelson, Paul A. (1978). "The Canonical Classical Model of Political Economy." *Journal of Economic Literature* 16, 1415–1434.

Sala-i-Martin, Xavier (1996). "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence." *European Economic Review* 40, 1325–1352.

Shinkai, Yoichi (1960). "On the Equilibrium Growth of Capital and Labor." *International Economic Review* 1, 107–111.

Simon, Julian L. (1990). "Great and Almost-Great Magnitudes in Economics." *Journal of Economic Perspectives* 4, 149–156.

Smith, M. Alasdair M. (1984). "Capital Theory and Trade Theory." Chapter 6 in Jones and Kenen (1984).

— (1977). "Capital Accumulation in the Open Two-Sector Economy." *Economic Journal* 87, 273–282. Reprinted in Bhagwati (1981, chapter 25).

— (1976). "Trade, Growth and Consumption in Alternative Models of Capital Accumulation." *Journal of International Economic* 6, 371–384.

Södersten, Bo (1964). *A Study of Economic Growth and International Trade.* Stockholm: Almqvist and Wiksell.

Solow, Robert M. (1994). "Perspectives on Growth Theory." *Journal of Economic Perspectives* 8, 45–54.

— (1956). "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70, 65–94. Reprinted in Stiglitz and Uzawa (1969), and in Becker and Burmeister (1991).

— and Paul A. Samuelson (1953). "Balanced Growth under Constant Returns to Scale." *Econometrica* 21, 412–424.

Srinivasan, T.N. (1964). "Optimal Savings in a Two-Sector Model of Growth." *Econometrica* 32, 358–373.

Stiglitz, Joseph E. (1970). "Factor Price Equalization in a Dynamic Economy." *Journal of Political Economy* 78, 456–488.

— and Hirofumi Uzawa (1969) (eds.). *Readings in the Modern Theory of Economic Growth.* Cambridge, MA: MIT Press.

Summers, Robert, and Alan Heston (1991). The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988. *Quarterly Journal of Economics* 106, 327–368.

Swan, Trevor Winchester (1956). "Economic Growth and Capital Accumulation." *Economic Record* 32, 334–361. Reprinted in Stiglitz and Uzawa (1969).

Takayama, Akira (1972). *International Trade.* New York: Holt, Rinehart, and Winston.

Uzawa, Hirofumi (1964). "Optimal Growth in a Two-Sector Model of Capital Accumulation." *Review of Economic Studies* 31, 1–25.

Wan, Henry Y., Jr. (1971). *Economic Growth.* New York: Harcourt Brace Jovanowich.

Wong, Kar-yiu (1995). *International Trade in Goods and Factor Mobility.* Cambridge, MA: MIT Press.

Woodland, Alan D. (1982). *International Trade and Resource Allocation.* Amsterdam: North-Holland.

Yang, Xiaokai, and Jeff Borland (1991). "A Microeconomic Mechanism for Economic Growth." *Journal of Political Economy* 99, 460–482.

Young, Alwyn (1991). Learning by Doing and the Dynamic Effects of International Trade. *Quarterly Journal of Economics* 106, 369–405.

CHAPTER 4

Endogenous Growth, Trade, and Specialization under Variable Returns to Scale: The Case of a Small Open Economy

Ngo Van Long, Kazuo Nishimura, and Koji Shimomura

1. Introduction

The spectacular success of several East and South East Asian economies has sparked a great deal of interest in the search for a better understanding of mechanisms that propel growth. According to the World Bank, the GNP per capita of Hong Kong, adjusted for purchasing power parity, have overtaken that of Canada and Japan. On the other hand, many countries remain unindustrialized and poor. What is it that prevents some countries from industrialization? Is there a poverty trap from which it is difficult to escape? What can national governments and international organizations do to accelerate the growth process of less developed countries?

There are many theories that claim to provide partial answers to the above questions. They range from the culture-based explanations of Harrison (1997) and Lee Kwan Yew to the socio-economic based theory of Murphy, Shleifer, and Vishny (1991), among others. In the book entitled "The Pan-American Dream: Do Latin America's Cultural Values Discourage True Partnership with the United States and Canada?" Harrison argues that Latin America's chronic failure to achieve lasting prosperity is due to an "Ibero-Catholic" culture. Lee Kwan Yew is well known for his view that "Confucian values" constitute the main driving force behind the East and South East Asian miracles. Eisuke Sakakibara's book (1993), "Beyond Capitalism", advances the view that values other than capitalistic profit-seeking ones contribute much to economic growth in East Asia.

Murphy, Shleifer, and Vishny (1991, p. 505) argue that "the allocation of talents to the rent-seeking sectors might be the reason for stagnation in much of Africa and Latin America, for slow growth in the United States, and for success of newly industrializing countries where these sectors are smaller." According to this view, bribes, taxes, and fees in a rent-seeking society constitute a tax on the profit of the productive sector; the higher the tax, the lower the incentive to invest. These authors suggest that a measure of this "tax rate" might be the size of government consumption, thus providing a plausible explanation of Barro's (1991) finding that countries with smaller government consumption relative GDP grow faster. The recent empirical work of Mauro (1995) lends support to this socio-economic approach. Using a newly assembled data set consisting of indices of corruption, red tape, and efficiency of the judicial system for about 70 countries, Mauro finds that countries with higher corruption tend to have a lower ratio of investment to GNP, and therefore slower growth.

Another stream of thoughts relates economic performance to more traditional concepts in economics such as returns to scale, externalities, and complementarity in demands and supplies. If there are impediments to world trade, small countries cannot take advantage of increasing returns to scale. For large economies, such as India, complementarity is a key factor for potential development. As an example of complementarity, it is often stated that the industrialization of one sector enlarges the size of market of other sectors. This process can be self-reinforcing, due to spillover effects, and backward and forward linkages. The idea of coordinated investments is at the heart of the theory of the "big push" associated with Rosenstein-Rodan (1943) and others. This line of arguments has been further developed by economists such as Nurkse (1953), Scitovsky (1954), and Flemming (1955). Their theories can be interpreted in terms of the concept of multiple equilibria in general equilibrium theory. This interpretive effort is most evident in another paper by Murphy, Shleifer, and Vishny (1989), where a set of models are presented, in all of which the central role is given to pecuniary externalities generated by imperfect competition with large fixed costs. An economy may remain settled in a low income equilibrium due to coordination failure. Active government intervention might be needed to move the economy to a high income equilibrium. However, by restricting attention to a two-period framework, these models, while offering a great deal of insight, lack much in dynamics. To talk about poverty traps, it is essential to address the issue of stability of equilibria. This has been formalized by Durlauf (1993) in a stochastic growth with many heterogeneous industries employing non-convex technologies that incorporate spillover effects from the history of production decisions to the productivity of the economy at the current time.

Another source of multiple equilibria can be traced to the role of wealth distribution. Galor and Zeira (1993) cite empirical work that show a positive correlation between the degree of equality in the distribution of wealth and the rate growth of GDP. They build an overlapping generations model that exhibits multiple equilibria, and show how the initial distribution of wealth affects aggregate output and investment.[1]

This chapter also explores multiple equilibria and poverty traps, but from a different perspective. Our model differs from the above literature in several important respects. Firstly, we adopt the infinite horizon optimizing approach, where the optimizing entity is a social planner. This enable us to demonstrate that attainment of a low level equilibrium may be due to the conscious choice of a social planner, given the initial capital stock.[2] This approach allows us to focus on the interplay between the properties of the planner's rate of discount and the properties of the GNP function. This contrasts sharply with models which rely on constant saving rates out of capital and wage incomes, or on the assumption that only workers save, and capitalists always dissave, an assumption commonly made in overlapping generations models. Since these models are by now well understood [see Galor (1996) for an exposition], our model serves as a counterpoint, and sheds light on an alternative mechanism of development.

The second distinguishing characteristic of our model is that we do not postulate the shape of the function relating aggregate per capita output (in value) to the capital labor ratio. It would be easy to generate multiple equilibria and low level trap from such a postulate [see Barro and Sala-i-Martin (1995) for an exposition]. Instead, we derive the properties of this function (called the GNP function in the international trade literature) from the properties of sectoral production functions, using the fact that capital and labor must be allocated to the two sectors to maximize current national income at given world prices. Our task of characterizing this function is complicated, because we allow for variable returns to scale in each sector.[3]

The third prominent feature of our model is the role of international trade in the growth process. With the exception of the article by Majumdar and Mitra (1995) which will comment on at a later stage, in models of poverty traps it is typically assumed that the economy is closed. By allowing for trade, we are able to identify another source of low level equilibrium. Under the assumptions made in our model (increasing returns to scale at low levels of output, low marginal product of capital when the capital labor ratio is near zero, no physical depre-

1. For further reviews of the literature on development traps, see Azariadis and Drazen (1990) and Azariadis (1996).

2. Whether he/she is benevolent and the objective function reflects the preferences of the consumers or not, is a distinct issue, to be discussed later.

3. Increasing returns to scale at low output levels may be attributed to factors such as set-up costs of various types; we assume constant returns to scale at high output levels.

ciation of capital) it is possible to show that in the absence of trade, the economy's stock of capital can only grow or stay constant, while with trade, the planner may choose to run down the capital stock by exchanging its capital stock for consumption goods at the given world price ratio. He will choose to do this if his rate of discount is sufficiently great and the initial capital stock is small. On the other hand, if the initial stock is sufficiently large, the economy will be able to take off, and perpetual growth in per capita output and consumption is possible. Thus, the opening of trade may be favorable or unfavorable to growth, depending on whether the initial capital stock exceeds or fall short of a certain threshold level.

Among the policy implications of our model are the effects of foreign aid and direct foreign investment on economic growth. In the traditional Solow growth model, an international donation of capital to a low saving economy will result in a temporary burst of output, investment, and consumption, but eventually the economy will return to the old steady state. So foreign aid has no long-lasting effects on the economy. By contrast, in our model, a foreign injection of capital to raise the domestic capital stock above a threshold level will permanently increase the economy's income and welfare. Such an injection may come about by foreign aid or direct foreign investment that gives the domestic government a share in ownership. In this connection it is interesting to note that a large proportion of foreign investment in China takes the form of joint-ventures, in contrast to the prevailing mode in Russia and Eastern Europe.

Another policy implication relates to the planner's rate of discount. Throughout our formal analysis, we take this rate of discount as given. This does not prevent one from peering beyond the model and ask questions about possible changes in the rate of discount. This rate reflects both the degree of impatience and uncertainty of tenure of the decision maker. For example, if the decision maker is the ruling capitalist class, any uncertainty concerning war, revolution, expropriation, will increase the discount rate, and result in the choice of a time path of declining capital and income. Conversely, political stability is conductive to growth. Therefore, any package of foreign aid should be reinforced by efforts to promote democracy, as in the long run this is the only form of government compatible with political stability. History shows that wars are typically initiated by totalitarian governments.

In our model, perpetual growth is a possible outcome only if the relative price of the capital intensive good in terms of the consumption good is sufficiently high. If a country exports the capital intensive good, and it is not a consumption good, then removal of foreign tariffs on this good will help growth. If the capital intensive good is the consumption good, and the limiting marginal product of capital (as capital tends to infinity) in that sector is low, then perpetual growth is not possible, unless there is technological innovation that raises the limiting marginal product of capital.

Given that we use the infinite horizon optimizing approach, it might be argued that even if trade leads to negative growth, this is beneficial from the point of view of the social planner. While this argument is correct, it does not detract from the importance of the possible negative effect of trade on growth. This is so for several reasons. First, a fall in per capita income in one country may have adverse impacts on other countries, when one takes into account illegal immigration induced by the growing income gap between rich and poor economies. Second, a social planner may represent a dominant interest group, possibly operating under uncertainty of tenure, and its discount rate (as well as its utility function) may not reflect an appropriately defined social rate of discount. Third, even if the social planner is truly representative of all individuals of the current generation, the objective function may fail to satisfy certain ethical criteria regarding intergenerational equity. We do not intend to address these issues here, as they are beyond the scope of the chapter. It suffices to point out that the social planner set-up does not necessarily mean that the chosen path is recommended.

Our optimizing approach and the small open economy setting of the chapter make it very close in spirit to the article by Majumdar and Mitra (1995). However, our results are different from theirs. Majumdar and Mitra show that there exists a poverty trap for their closed economy, and that when the country is opened to trade, it will overcome the poverty trap and succeed in securing ever rising consumption per capita, regardless of the initial condition.

The main reason for the difference in the results is that Majumdar and Mitra adopt the following assumptions: (a) capital is the only factor to be allocated between the consumption goods and the capital goods sectors, (b) output of the consumption goods sector is linear in capital input, and (c) only the capital goods sector has the S-shape production function. It follows that their GNP function for the small open economy is bounded below by a linear and increasing function of the form $Y = AK$. This is not the case in our model. We assume that both capital and labor are to be allocated between the two sectors, in keeping with the Heckscher-Ohlin-Uzawa tradition; and both sectors have the S-shape production function. Our GNP function exhibits zero marginal products of capital when the stock of capital approaches zero. It follows that for any positive rate of discount, the open economy will choose to decumulate the stock of capital (by exchanging capital goods for consumption goods), provided that the initial stock is sufficiently small. Unlike the Majumdar-Mitra model, the opening of trade does not ensure growth.

For any initial capital stock that exceeds a threshold level, perpetual growth will take place provided the rate of discount is not too great. This is due to the assumption that the marginal product of capital in the capital-intensive sector approaches a positive value when capital approaches infinity. This assumption corresponds to a common feature of all endogenous growth model. Whether capital is merely physical capital, or embodied in new designs, or human capital, for perpetual growth there must exist a positive lower bound on the social marginal product of at least one capital stock, or an aggregate of several capital stocks when this variable becomes arbitrarily large. Whether this arises from externalities or not is simply a matter of details.[4]

Our analysis in the remaining sections makes precise our intuitive reasoning. The first section states the main assumptions. The second section describes the optimal growth model of a small open economy. The third section proves the existence of a unique threshold. The fourth section gives some concluding remarks.

2. The Assumptions

Consider a small country dynamic model producing good 1 and good 2 using two factors of production, capital and labor. The price ratio is exogenous and constant over time. There is no difference between new and existing capital goods, and both consumption- and capital-goods have their international markets in which households face given international prices. It is immaterial whether we identify good 1 with the consumption good or with the capital good. On the other hand, both labor- and capital-services are internationally non-traded.

Let us specify production technology. The production function of good j is assumed to be a homothetic function of labor (L_j) and capital (K_j)

$$y_j = G[F^j(K_j, L_j)] = G(z_j), \quad j = 1, 2, \tag{1}$$

where the two functions, $F^j(K_j, L_j)$, $j = 1, 2$, are assumed to satisfy all properties which the neoclassical production functions in the standard Heckscher-Ohlin model of international trade would have to satisfy. Moreover, we make the following assumption concerning the two functions, $F^j(K_j, L_j)$, $j = 1, 2$.

4. See the survey of Long and Wong (1996). See also Jensen and Larsen (1987).

Assumption 1. *For any $\omega > 0$, if $\omega = F_k^j/F_l^j$, $j = 1, 2$, then $K_1/L_1 < K_2/L_2$.*

In a word, as far as the F-functions are concerned, good 1 is more labor-intensive than good 2.

Next, let us assume that $G(\cdot)$ has the following properties.

Assumption 2. *(i) $G(z)$ is increasing in $z > 0$;*

(ii) $G(z)$ is continuous in $z \geq 0$ and $G(0) = 0$;

(iii) There exists a Z^* such that $G(z) = z$ for $z \in [Z^*;\ \infty)$ and $G(z) < z$ for $z \in (0;\ Z^*)$;

(iv) $G(z)$ has continuous second derivative on $[0;\ Z^*)$ and $G'(0) = 0$;

(v) There exists a unique $\bar{Z}$ on $[0;\ Z^*)$ such that $G''(\bar{Z}) = 0$.

We may think of z_j as "fictitious" output of good j, and $G(z_j)$ as the true output of good j. Assumption 2(iii) implies that for $z_j \geq Z^*$, the fictitious output and the true output are equal, and for $z_j < Z^*$, the fictitious output is greater than the true output. Figure 1 shows that the graph of $G(z)$ is S-shaped.

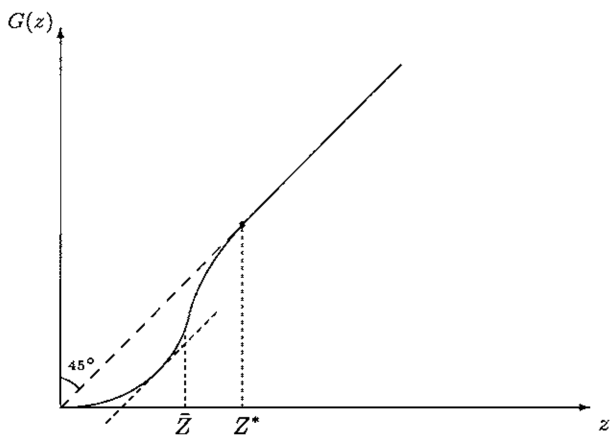


Fig. 1. The graph of $G(z)$

Lemma 1. *Under Assumption 2, $G(z)/z$ is increasing in $z \in [0;\ Z^*)$.*

Proof. Differentiating $G(z)/z$ with respect to z, we have

$$\frac{d}{dz}\left[\frac{G(z)}{z}\right] = \frac{zG'(z) - G(z)}{z^2} \equiv \frac{\theta(z)}{z^2}.$$

Considering Assumption 2(ii), $\theta(0) = 0$. Since $d\theta(z)/dz = zG''(z) > 0$ for $z \in (0; \bar{Z})$, $\theta(z) > 0$ for $z \in [0; \bar{Z})$. Suppose that there is $z_0 \in (\bar{Z};\ Z^*)$ such that $\theta(z_0) < 0$. Then, for any $z \in [z_0;\ Z^*)$ $\theta(z) < 0$, since $d\theta(z)/dz = zG''(z) < 0$ in the interval, which means that $G(z)$ cannot catch up with the 45^0-line at Z^*. □

Given that $G(z)$ is S-shaped, we may wonder under what conditions $G[F^j(K, \bar{L})]$, when drawn against K for a given $\bar{L} > 0$, also has a similar S-shape.

Assumption 3. *(i)* $\lim_{K\to 0} G'[F^j(K, L_j)]F^j_K(K, L_j) = 0$ *for all* $L_j > 0$*;*
(ii) For a given $\bar{L}$*, there is a unique* K_j^{**} *such that*

$$G''[F^j(K_j^{**}, \bar{L})][F^j_K(K_j^{**}, \bar{L})]^2 + G'[F^j(K_j^{**}, \bar{L})]F^j_{KK}(K_j^{**}, \bar{L}) = 0. \quad (2)$$

K_j^{**} *is an inflection point for* $G[F^j(K, \bar{L})]$ *when it is treated as a function of* K*;*
(iii) $\lim_{K\to\infty} F^j_K(K, L_j) > 0$ *for all* $L_j > 0$.

Lemma 2. *Under Assumption 3, the production function of good* j $G[F^j(K, \bar{L})]$ *has the properties*
(i) $G[F^j(0, \bar{L})] = 0$*;*
(ii) $\partial G[F^j(K, \bar{L})]/\partial K = 0$ *if* $K = 0$*, and* > 0 *if* $K > 0$*;*
(iii) $\partial G[F^j(K, \bar{L})]/\partial K$ *is continuous in* $K > 0$*;*
(iv) $G[F^j(K, \bar{L})] \equiv F^j(K, \bar{L})$ *for any* $K \geq K_j^*$*, where* K_j^* *is defined as a unique solution to* $Z^* = G[F^j(K, \bar{L})]$. K_j^* *is the point where increasing returns to scale are exhausted;*
(v) $G[F^j(K, \bar{L})] < F^j(K, \bar{L})$ *for any* $0 < K < K_j^*$*;*
(vi) $\partial^2 G[F^j(K, \bar{L})]/\partial K^2 > 0$ *for* $K \in (0;\ K_j^{**})$ *and* < 0 *for* $K \in (K_j^{**};\ K_j^*)$.

Proof. The proof is elementary, and is left to the reader. □

Figure 2 depicts the graph of $G[F^j(K, \bar{L})]$, which is equal to $F^j(K, \bar{L})$ if $K \geq K_j^*$, but less than that otherwise. One implication of the assumption that production technologies in both sectors would be completely Heckscher-Ohlin if $G(\cdot) \equiv 1$ is that, letting p be any given

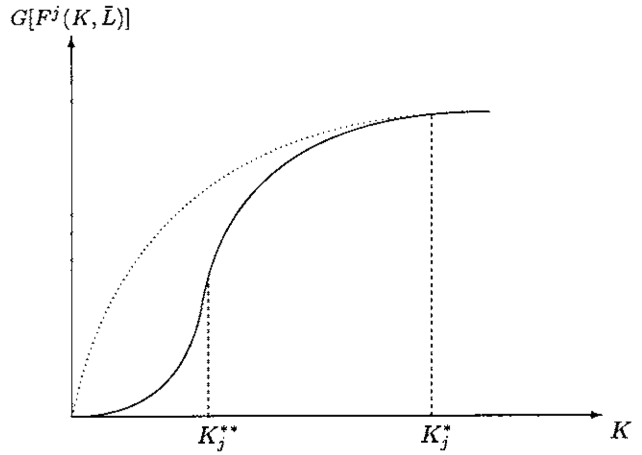


Fig. 2. The production function of good j

international relative price of good 1 measured by good 2, the system of equations

$$\begin{aligned} &p[F^1(k_1,\bar{L}) - k_1 F_K^1(k_1,\bar{L})] = F^2(k_2,\bar{L}) - k_2 F_K^2(k_2,\bar{L}), \\ &pF_K^1(k_1,\bar{L}) = F_K^2(k_2,\bar{L}), \end{aligned} \tag{3}$$

where the unknowns are k_j, $j = 1, 2$, has a solution, $\hat{K}_j$, $j = 1, 2$, such as depicted in figure 3. Note that when the capital endowment K satisfies $\hat{K}_1 < K < \hat{K}_2$ (resp. $0 < K \leq \hat{K}_1$, $\hat{K}_2 \leq K$) production would be incompletely specialized (resp. completely specialized to good 1, completely specialized to good 2) if the F functions were true production functions.

Assumption 4. $\hat{K}_1 > \max\{K_1^*;\ K_2^*\}$.

Now let us define the GNP function as follows

$$\begin{aligned} g(K) \equiv \max\ &pG[F^1(K_1,L_1)] + G[F^2(K_2,L_2)] \\ &\text{s.t.}\ \ K \geq K_1 + K_2, \quad L \geq L_1 + L_2. \end{aligned} \tag{4}$$

We have the following lemma.

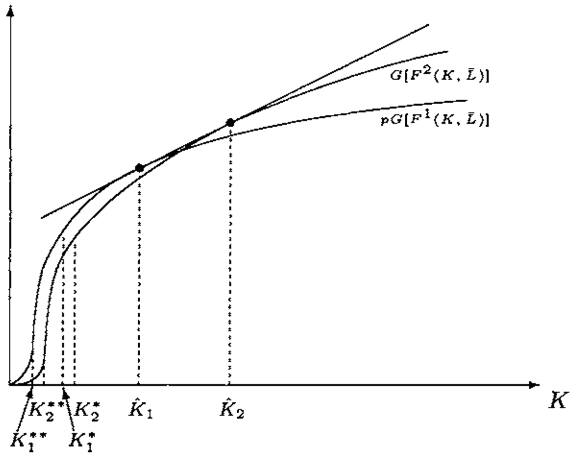


Fig. 3. The graphs of $pG[F^1(K,\bar{L})]$ and $G[F^2(K,\bar{L})]$

Lemma 3. *Under the foregoing assumptions, $g(K)$ satisfies (a)–(c):*
(a) There are two values, $\hat{K}_1^*$ and $\hat{K}_2^*$, such that $g(K) = pG[F^1(K,\bar{L})]$ for $K \in (0;\ \hat{K}_1^*]$ and $g(K) = G[F^1(K,\bar{L})]$ for any $K \in [\hat{K}_2^*;\ \infty)$, i.e., production is completely specialized to either good. Incomplete specialization takes place for $K \in (\hat{K}_1^*;\ \hat{K}_2^*)$;
(b) $\lim_{K\to\hat{K}_1^*+} g'(K) > \partial pG[F^1(K,\bar{L})]/\partial K|_{\hat{K}_1^*}$ *and*
$\lim_{K\to\hat{K}_2^*-} g'(K) < \partial G[F^2(K,\bar{L})]/\partial K|_{\hat{K}_2^*}$;
(c) There are two values, K_1^0 and K_2^0, $\hat{K}_1^* < K_1^0 < K_2^0 < \hat{K}_2^*$, such that

$$\lim_{K\to K_1^0} g'(K) = \lim_{K\to K_2^0} g'(K) = \frac{g(\hat{K}_2) - g(\hat{K}_1)}{\hat{K}_2 - \hat{K}_1},$$

and for any $K \in (K_1^0;\ K_2^0)$

$$g(K) = g(\hat{K}_1) + g'(\hat{K}_1)(K - \hat{K}_1).$$

Proof. See the Appendix. □

Figure 4 depicts the graph of $g(K)$. As is shown in the Appendix, when the capital stock is at the point $\hat{K}_1^*$ or $\hat{K}_2^*$, there is discrete jump in the output of both goods and the economy switches from incomplete specialization to complete specialization. This is obvious because it is not efficient to produce goods at small scale of operation.

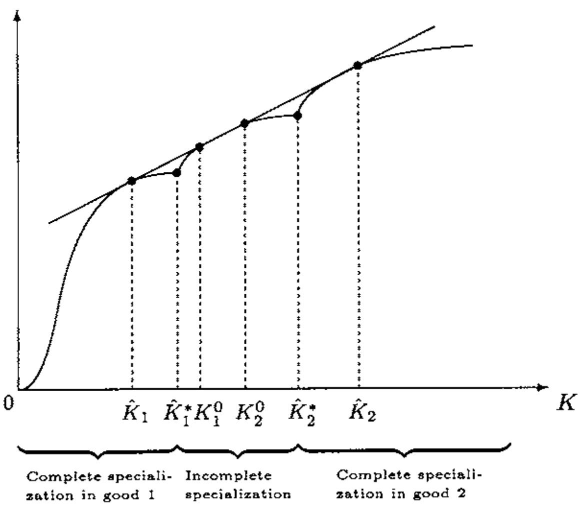


Fig. 4. The GNP function

3. The Optimal Growth Model

Using the GNP function $g(K)$, (4), we can formulate the optimal growth model as the solution of the dynamic problem

$$\begin{aligned} \max & \int_0^\infty u(c)e^{-\rho t}dt \\ \text{s.t.} \ & \dot{K} = g(K) - c, \\ & K \geq 0, \quad c \geq 0, \\ & K(0) = K_0 \ \text{given}, \end{aligned} \tag{P}$$

where the utility function $u(c)$ satisfies the following conditions:

$$\begin{aligned} &u'(c) > 0, \quad u''(c) < 0, \quad \forall c > 0, \\ &u'(0) = \infty, \quad u'(\infty) = 0. \end{aligned}$$

Concerning the time-discount rate ρ, we assume that it is quite small.

Assumption 5. *ρ is so small that there is a unique $\tilde{K} \in (0; K_1^{**})$ such that $g(\tilde{K}) = \rho\tilde{K}$ and $\lim_{X \to K^+} g'(X) > \rho$ for any $K \geq \tilde{K}$.*

Since $g(K)$ is strictly convex for $K \in (0;\ K_1^{**})$, it follows from Assumption 5 that there is $\breve{K} \in (0;\ K_1^{**})$ such that $g'(\breve{K}) = \rho$.

Associated with the problem (P) is the Hamiltonian

$$H \equiv u(c) + \pi[g(K) - c].$$

Denoting by $(K(t), c(t))$ a path as a feasible solution without specifying an initial condition and by $(K(t, K_0), c(t, K_0))$ a path from the initial point K_0, the necessary condition for optimality is that there is a costate variable $\pi(t)$ such that

$$\frac{\partial H}{\partial c} = u'(c(t)) - \pi(t) = 0, \tag{5.1}$$

$$\dot{\pi}(t) = \pi(t)[\rho - g'(K(t))], \text{ where } g'(K) \text{ exists}, \tag{5.2}$$

$$\dot{K}(t) = g(K(t)) - c(t). \tag{5.3}$$

Using the first equation we can rewrite the second one as

$$\dot{c}(t) = \frac{c(t)}{\sigma(t)}\,[g'(K(t)) - \rho], \quad \sigma(t) \equiv -\frac{c(t)u''(c(t))}{u'(c(t))}. \tag{6}$$

4. The Existence of Threshold

A continuous time optimal growth problem with a convex-concave production function in an infinite time horizon was studied by Skiba (1978). However his characterization was not complete. Below we follow methods in Leonard and Long (1992, Chap. 9) and Deckert and Nishimura (1983) and provide a characterization of optimal paths in the non-concave problem (P), where the Hamiltonian is not concave in the state variable. We first state the following result due to Michel (1982).

Lemma 4. *A necessary condition for $(K(t), c(t))$ to be an optimal solution to (P) is that there exists a costate variable $\pi(t)$ that is continuous with respect to t for $0 < t < \infty$.*

An optimal path of the control variable $c(t)$ satisfies $\pi(t) = u'(c(t))$. Hence it is also continuous with respect to t by Lemma 4. We say that a path $(K(t), \pi(t))$ with $\pi(t) = u'(c(t))$ is optimal if a path $(K(t), c(t))$ is optimal.

Using Lemma 4, we can prove that the optimal path of capital stock is monotone. Suppose that $K(t, K_0)$ is not monotone with respect to t. Then, there exists $t_1 > 0$ and $T > 0$ such that $K(t_1, K_0) \neq K(t_1 + T, K_0)$, and we can construct an alternative path in the following way

$$\breve{K}(t, K_0) = \begin{cases} K(t, K_0), & 0 \leq t \leq t_1 + T, \\ K(t - T, K_0), & t_1 + T < t, \end{cases}$$

$$\breve{\pi}(t, K_0) = \begin{cases} \pi(t, K_0), & 0 \le t \le t_1 + T, \\ \pi(t, K_0), & t_1 + T < t. \end{cases}$$

By the autonomous nature of the problem (i.e., time appears explicitly only in the exponential discount term $e^{-\rho t}$), the path $(\breve{K}(t, K_0), \breve{\pi}(t, K_0))$ must also be optimal. But $\breve{\pi}(t, K_0)$ is discontinuous, which contradicts Lemma 4. The optimal path of state variables must be therefore monotone.

Lemma 5. *The optimal path of capital stock $K(t)$ is monotone.*

Based on the foregoing lemma, we shall prove the main propositions.

Proposition 1. *There exists $K^{\#}\,(< \breve{K})$ such that $K(t, K_0)$ converges to zero $\forall K_0 \in (0;\ K^{\#})$.*

Proof. Suppose not. That is, suppose that for any $K^{\#} \in (0;\ \breve{K})$ there is $K_0^0 \in (0;\ K^{\#})$ such that $K(t, K_0^0)$ does not converge to zero. Since the system (5.3)–(6) has no stationary state other than $(K, c)\ [\equiv (K, g'(K))]$ in the interval $(0;\ \breve{K})$ and we can verify that $(\breve{K}, \breve{c})$ is locally unstable, Lemma 4 implies that $K(t, K_0^0)$ must reach $\breve{K}$ at some finite time T. Then we have

$$\begin{aligned} K(T, K_0^0) &= \breve{K}, \\ \dot{K}(T, K_0^0) &= \dot{K}(0, \breve{K}), \\ c(T, K_0^0) &= c(0, K_0^0). \end{aligned}$$

The third equality is due to Lemma 4, and the second one is due to the first and third equations.

From (6), $c(t, K_0^0)$ is decreasing for $0 < t < T$. Then

$$g(K_0^0) - \dot{K}(0, K_0^0) = c(0, K_0^0) > c(T, K_0^0) = c(0, \breve{K}).$$

Hence $g(K_0^0) > c(0, \breve{K})$. However, K_0^0 can be chosen to be arbitrarily small, we have $g(K_0^0) \le c(0, \breve{K})$ by an appropriate choice of K_0^0. This contradiction establishes the validity of Proposition 1. □

Proposition 2. *For any initial condition $K_0 \in [\tilde{K};\ \infty)$, $K(t, K_0)$ must diverge to infinity.*

Proof. Suppose not. Then there exists $K_0^0 \in [\tilde{K};\ \infty)$ such that either

(i) $K(t, K_0^0)$ is monotonously decreasing;

or

(ii) $K(t, K_0^0)$ is monotonously increasing and converges to a finite value.

However, (ii) is impossible, because of Assumption 5.[5] Thus let us concentrate on (i). Since $K(t, K_0^0) < K_0^0$ for all $t > 0$,

$$g(K(t, K_0^0)) - \rho K(t, K_0^0) < g(K_0^0) - \rho K_0^0$$

for all $t \geq 0$. Integrating both sides of the inequality, we have

$$\int_0^\infty [g(K(t, K_0^0)) - \rho K(t, K_0^0)] e^{-\rho t} dt < \frac{1}{\rho}[g(K_0^0) - \rho K_0^0]. \tag{7}$$

Considering

$$\int_0^\infty \dot{K} e^{-\rho t} dt = \left[K e^{-\rho t}\right]_0^\infty + \int_0^\infty K e^{-\rho t} dt,$$

equation (7) becomes

$$\int_0^\infty [g(K(t, K_0^0)) - \dot{K}(t, K_0^0)] e^{-\rho t} dt - K_0^0 < \frac{1}{\rho}\,[g(K_0^0) - \rho K_0^0],$$

or

$$\rho \int_0^\infty c(t, K_0^0) e^{-\rho t} dt < g(K_0^0).$$

Let $\bar{c} \equiv \rho \int_0^\infty c(t, K_0^0) e^{-\rho t} dt$. Then, by Jensen's inequality,

$$u(\bar{c}) > \rho \int_0^\infty u(c(t, K_0^0)) e^{-\rho t} dt,$$

or

$$\int_0^\infty u(\bar{c}) e^{-\rho t} dt > \int_0^\infty u(c(t, K_0^0)) e^{-\rho t} dt.$$

That is, the constant-consumption path $c(t) \equiv \bar{c}$ gives a higher utility than $\int_0^\infty u(c(t, K_0^0)) e^{-\rho t} dt$ and is feasible because $\bar{c} < g(K_0^0)$, a contradiction. □

Lastly, let us concentrate on the interval $[K^\#;\ \tilde{K})$. First, if $K(t, K_0)$ diverges (resp. converges) to infinity (resp. zero), so does $K(t, K_0^0)$ for any $K_0^0 >$ (resp. $<$) K_0. Second, as we already stated, there is no stationary state in $[\tilde{K};\ \infty)$. Therefore there must be a threshold $K^\Xi \in [K^\#;\ \tilde{K})$. We arrive at the main theorem.

Theorem. *There is a threshold K^Ξ in $[\breve{K};\ \tilde{K})$ such that if $K_0 > K^\Xi$, then $K(t, K_0)$ diverges to infinity, and if $K_0 < K^\Xi$, then $K(t, K_0)$ converges to zero.*

5. Note that there is no stationary state in $[\tilde{K};\ \infty)$.

5. Concluding Remarks

This chapter has constructed an optimal growth model of a small open economy whose production structure is Heckscher-Ohlin except that each production function is homothetic and has a increasing-returns-to-scale portion, and has shown that if the time-discount rate is sufficiently low, then there exists the "poverty trap".

This result is quite different from Majumdar and Mitra (1995) who show that the "poverty trap" which appears in the autarchic economy can disappear and a persistent growth is possible once the economy becomes a small open economy. On the other hand, our result suggests that trade does not always provide the necessary engine for growth.

The source of the difference between the results of Majumdar and Mitra and ours lies in the formulation of production structure. They assume that the capital productivity of the consumption good is constant, which means that if the time-discount rate is smaller than the constant capital-productivity, then the rate of consumption steadily increases and it is necessary to accumulate capital stock in order to meet the increase in the rate of consumption.

On the other hand, we assumed that *both* industries have a portion of increasing-returns-to-scale, which means that the marginal productivity of capital is smaller than the time-discount rate when the initial stock of capital is sufficiently small. Then, consumption steadily decreases. If consumption steadily decreases along an optimal growth path, the capital stock must also decrease steadily. Trade can be a cause of the poverty trap.

Lastly, let us make a remark on trade pattern. Our theorem suggests that the pattern of international trade depends on whether the initial stock of capital is greater than the threshold level: If the former is greater than the latter, then the small open economy produces only the more capital-intensive good in the long run. If the former is smaller than the latter, then the economy produces only the more labor-intensive good. Furthermore, when the economy switches from complete specialization to incomplete specialization, there are discrete jumps in the output of both goods.

Appendix: The GNP Function $g(K)$

Let us derive the GNP function $g(K)$.

For given p and a fixed workforce $\bar{L}$, there exists a unique line in the output space that is consistent with both full employment of labor and diversification if the true production functions are F^1 and F^2. This is the

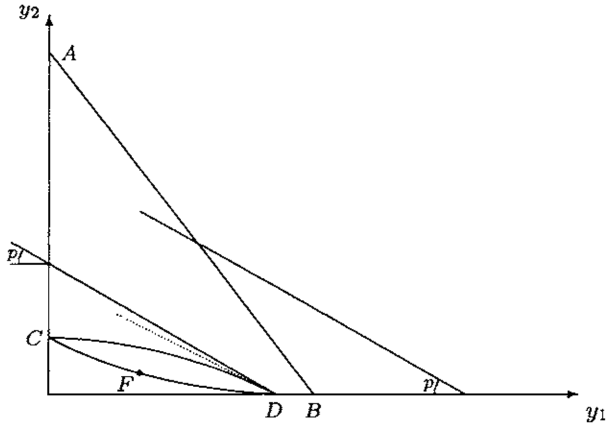


Fig. A.1. The Rybczynski line corresponding to F^1 and F^2

line AB. It is the graphical representation of the full labor employment equation

$$a_{L_1}(\omega)z_1 + a_{L_2}(\omega)z_2 = \bar{L},$$

where ω is the factor price ratio that uniquely corresponds to the goods price ratio p if both goods are produced, and a_{L_j}, $j = 1, 2$, is the amount of labor per unit of output of z_j. There is similar equation for the full employment of capital consistent with diversification.

$$a_{K_1}(\omega)z_1 + a_{K_2}(\omega)z_2 = K.$$

Clearly, for the two lines to intersect in the positive quadrant, K must be within the range $[\hat{K}_1;\ \hat{K}_2]$, where $\hat{K}_j$; $j = 1, 2$, is defined by

$$\frac{\bar{L}}{\hat{K}_j} = \frac{a_{L_j}(\omega)}{a_{K_j}(\omega)}.$$

An alternative definition of $\hat{K}_j$, $j = 1, 2$, is that they are the solution to the system of equation (4) in the text. When $K = \hat{K}_1$ (resp. $\hat{K}_2$), the fictitious production possibility curve for z_1 and z_2 lies everywhere below (resp. above) the line AB, except at B (resp. A), and its slope at B (resp. A) is $-p$. Assumption 4 in the text implies that $F^j(\hat{K}_j, \bar{L}) > Z^*$, $j = 1, 2$, and therefore $G[F^j(\hat{K}_j, \bar{L})] = F^j(\hat{K}_j, \bar{L})$. Hence OA and OB depict the output levels of good 1 and good 2, respectively.

Take any K which is smaller than $\hat{K}_1$. The concave curve CD in figure A.1 is the fictitious production possibility curve which corresponds to this K. Then, since the true production function of each good $G[F^j(\cdot)]$ is always less efficient than the fictitious production function $F^j(\cdot)$, the true production possibility curve CFD should be inside the fictitious one. Clearly, D is the optimal point. We arrive at the first proposition.

Proposition A.1. *If $K \leq \hat{K}_1$, then production is completely specialized to good 1.*

A parallel argument brings us to the second proposition.

Proposition A.2. *If $K \geq \hat{K}_2$, then production is completely specialized to good 2.*

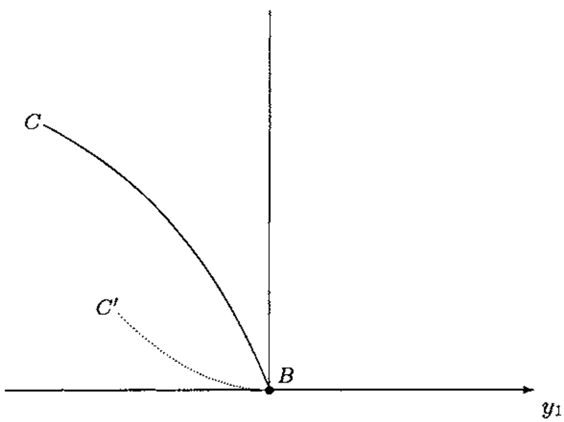


Fig. A.2. A neighborhood of point *B* in figure A.1

When $K = \hat{K}_1$, we have figure A.2 which depicts a neighborhood of point B in figure A.1, where BC is the fictitious production possibility curve which is defined by

$$\begin{aligned} Y_2 = T(Y_1, \hat{K}_1) \equiv \max_{K_j, L_j} \ & F^2(K_2, L_2) \\ \text{s.t.} \ & \hat{K}_1 \geq K_1 + K_2, \\ & \bar{L} \ \geq L_1 + L_2, \\ & Y_1 \ \leq F^1(K_1, L_1). \end{aligned}$$

BC' is the true production possibility curve which is defined by the system of equations

$$z_2 = G[T(G^{-1}(z_1),\ \hat{K}_1)] \equiv \tilde{T}(z_1, \hat{K}_1).$$

Note that the true production possibility curve is below the fictitious production possibility curve, since the true production functions $G[F^j(K_j, L_j)]$ are less efficient than the fictitious production functions. Moreover, differentiating $\tilde{T}(Y_1, \hat{K}_1)$ with respect to Y_1 and evaluating the partial derivative at B in figure A.2, we see that

$$\frac{\partial z_2}{\partial z_1} = \frac{G'[T(z_1, \hat{K}_1)]}{G'(z_1)} \frac{\partial T(z_1, \hat{K}_1)}{\partial z_1} = 0,$$

because $G'[T(z_1, \hat{K}_1)] = 0$ at B. Therefore, the production possibility curve $z_2 = \tilde{T}(z_1, \hat{K}_1)$ never touches the price line starting from the intersection of $z_2 = \tilde{T}(z_1, \hat{K}_1)$ and the horizontal axis of coordinates just except for that intersection. Clearly this fact still holds even if K increases slightly from $\hat{K}_1$. We have the following proposition.

Proposition A.3. *In a neighborhood of $K = \hat{K}_1$, production is completely specialized to good 1.*

A parallel argument ensures us that the following proposition also holds.

Proposition A.4. *In a neighborhood of $K = \hat{K}_2$, production is completely specialized to good 2.*

Next let us focus on the open interval $(\hat{K}_1;\ \hat{K}_2)$. At first, let us check the shape of an isoquant curve $\bar{g} = pG(z_1) + G(z_2)$ such that $\bar{g}$ is greater than $(1+p)Z^*$. First, for a given $\bar{g}$, define the corresponding $\bar{Z}_1$ and $\bar{Z}_2$ by

$$\bar{Z}_1 \equiv \frac{\bar{g} - Z^*}{p} > Z^*, \qquad \bar{Z}_2 \equiv \bar{g} - pZ^* > Z^*.$$

For z_2 in the interval $(Z^*;\ \bar{Z}_2)$, any small decrease in z_2, say $|\Delta z_2|$, must be compensated for by a small increase in z_1 so that GNP remains constant. However, when $z_2 = Z^*$, any decrease in z_2 by $|\Delta z_2|$, where $|\Delta z_2|$ is any finite amount such that $0 < |\Delta z_2| < Z^*$, implies that $G(z_2)$ falls by more than $|\Delta z_2|$. To maintain GNP at $\bar{g}$, there must be a compensating increase in z_1: $|\Delta z_1| = |\Delta G(z_2)|/p = |\Delta z_2|/p$. This explains why the isoquant curve for $\bar{g}$ lies *above* the dotted line $c\bar{Z}_1^f$, which has the slope $-p$.

Taking

$$\left.\frac{dz_2}{dz_1}\right|_{\bar{g}=pG(z_1)+G(z_2)} = -\frac{pG'(z_1)}{G'(z_2)},$$

$$\left.\frac{d^2z_2}{dz_1^2}\right|_{\bar{g}=pG(z_1)+G(z_2)} = -\frac{p[G''(z_1)(G'(z_2))^2 + pG''(z_2)(G'(z_1))^2]}{(G'(z_2))^3},$$

we have

Lemma A.1. *Let us denote an isoquant curve of $\bar{g} = pG(z_1) + G(z_2)$ by $z_2 = \Phi(z_1;\bar{g})$. Then for any $\bar{g}$ which is greater than $(p+1)G(Z^*)$, $\Phi(z_1;\bar{g})$ has the following properties:*

(i) $d\Phi/dz_1 < 0$ for any $z_1 \in (0;\ \bar{g}/p)$;

(ii) $\lim_{z_1\to 0} d\Phi/dz_1 = 0$ and $\lim_{z_1\to\bar{g}/p} d\Phi/dz_1 = -\infty$;

(iii) There are two values of Z_1, Z' and Z'' where $0 < Z' < Z^*$ and $\bar{Z}_1 < Z'' < \bar{g}/p$, such that

$$\frac{d^2\Phi}{dz_1^2}\left\{\begin{array}{c} < \\ = \\ > \end{array}\right\} 0, \ \textit{if}\ z_1 \in \left\{\begin{array}{c} (0;\ Z')\cup(Z'';\ \bar{g}/p), \\ [Z^*;\ \bar{Z}_1]\cup\{Z'\}\cup\{Z''\}, \\ (Z';\ Z^*)\cup(\bar{Z}_1;\ Z''); \end{array}\right.$$

(iv) $d\Phi/dz_1 = -p$ if $z_1 \in [Z^*;\ \bar{Z}_1]$.

Proof. The proof is elementary, and is left to the reader. □

The curve that depicts $\bar{g} = pG(z_1) + G(z_2)$ is *continuous*, and on that curve, when $z_2 = 0$ then $\bar{g} = pG(z_1)$, which has the solution $z_1 = G^{-1}(\bar{g}/p) = \bar{g}/p$, where the second equality follows from the assumption $\bar{g} > (p+1)G(Z^*)$. On the other hand, when $z_1 = 0$ then $\bar{g} = G(z_2)$, which has the solution $z_2 = G^{-1}(\bar{g}) = \bar{g}$, where the second equality again follows from the above assumption. Hence, if $\bar{g} > (p+1)G(Z^*)$, Z_1^f coincides with Z_1^0 in figure A.3 and the isoquant curve must be depicted like in figure A.4.

Based on the foregoing argument, we can prove that under Assumption 1 in the text there is a unique interval of K in which incomplete specialization.

First, the existence of such an interval. Since Assumption 1 implies that the Rybczynski line corresponding to the fictitious production possibility locus passes above the point (Z^*, Z^*) like $AA'EB'B$ in figure A.5, we have incomplete specialization at any point on a portion of the Rybczynski line, AB.

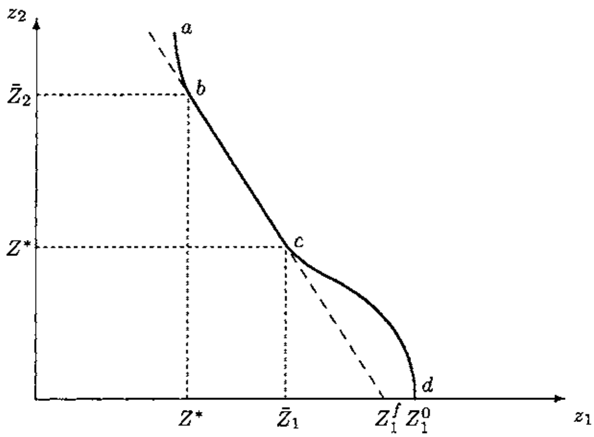


Fig. A.3. The curve $\bar{g} = pG(z_1) + G(z_2)$

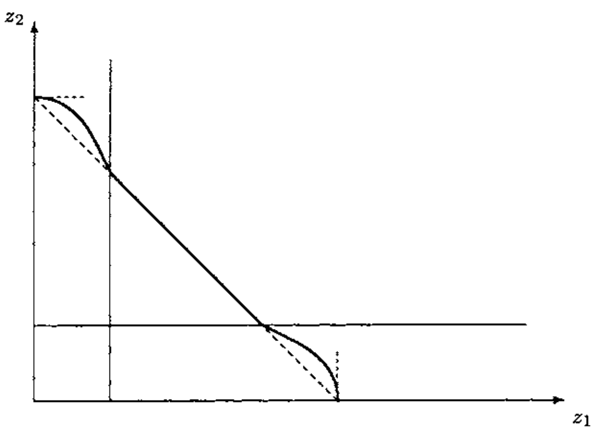


Fig. A.4. The curve $\bar{g} = pG(z_1) + G(z_2)$

[Note that the GNP function is derived by solving the optimization problem

$$g(K) \equiv \max_{z_1, z_2} \; pG(z_1) + G(z_2) \qquad \text{(P}'\text{)}$$
$$\text{s.t.} \;\; z_2 = \tilde{T}(z_1, K).$$

Incomplete specialization takes place iff the solution implies positive outputs for both goods.]

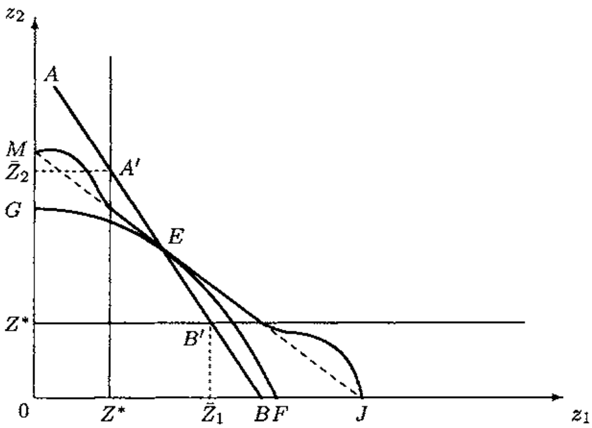


Fig. A.5. The optimal solution to (P′)

Let K_2^0 be defined by

$$\bar{Z}_2 = \tilde{T}(Z^*, K_2^0),$$

where $(Z^*, \bar{Z}_2)$ represents the coordinates of point A' of figure A.5. Then, for all K in $[K_1^0;\, K_2^0]$, production is incompletely specialized and output of each good will be at least equal to Z^*.

We now show that there exists a *unique* $\hat{K}_1^*$ in $(\hat{K}_1;\, K_1^0)$ at which production switches from specialization in good 1 to incomplete specialization, and that at the switch, output of good 1 (resp. good 2) makes a discrete jump downwards (resp. upwards). Figure A.6 exhibits the situation at the switching point.

Let us obtain, at first, the slopes of the GNP function $g(K)$ and the production function $pG[F^1(K, \bar{L})]$. Making use of the first-order condi-

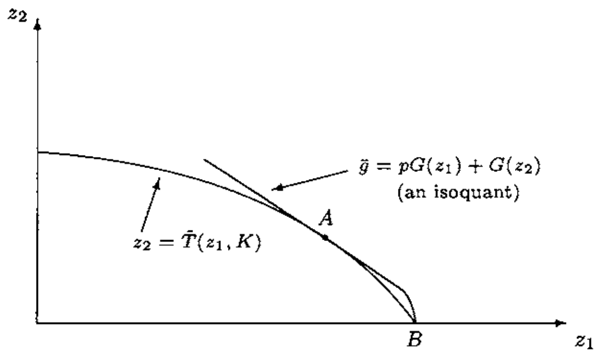


Fig. A.6. The switching point

tion of the problem (P′) and the envelope theorem, we have

$$\begin{aligned}\frac{dg(K)}{dK} &= G'[\tilde{T}(z_1, K)]\tilde{T}_K(z_1, K)\\ &= -\frac{pG'(z_1)\tilde{T}_K(z_1, K)}{\tilde{T}_K(z_1, K)}\\ &= pG'[F^1(K_1, L_1)]F_K^1(K_1, L_1),\end{aligned}$$

where (K_1, L_1) corresponds to point A in figure A.6. On the other hand

$$\frac{dpG[F^1(K, \bar{L})]}{dK} = pG'[F^1(K, \bar{L})]F_K^1(K, \bar{L}).$$

Since the Rybczynski line is above the point (Z^*, Z^*), it is clear from figure A.5 that both $G'[F^1(k_1, l_1)]$ and $G'[F^1(K, \bar{L})]$ are 1 at A and B. Thus $dg(K)/dK = pF_K^1(K_1, L_1)$ and $dpG[F^1(K, \bar{L})]/dK = pF_K^1(K, \bar{L})$. Since $F_K^1(K_1, L_1)$ must be greater than $F_K^1(K, \bar{L})$, due to the Stolper-Samuelson Theorem[6], at the switching point between complete specialization in good 1 and incomplete specialization we always have

$$\frac{dg(K)}{dK} > \frac{dpG[F^1(K, \bar{L})]}{dK}.$$

6. Note that good 1 is assumed to be labor-intensive.

By a parallel argument, we can show that at the switching point between incomplete specialization and complete specialization in good 2 we always have

$$\frac{dg(K)}{dK} < \frac{dpG[F^1(K,\bar{L})]}{dK}.$$

Therefore, the above results imply that incomplete specialization occurs in a unique and connected interval, $[\hat{K}_1^*;\ \hat{K}_2^*]$, and that the intersection of the curves $pG[F^1(K,\bar{L})]$ and $G[F^2(K,\bar{L})]$ must be interior of the interval.

Proposition A.5. *If the Rybczynski line corresponding to the "imaginary" production functions $F^1(\cdot)$ and $F^2(\cdot)$ passes above the point (Z^*, Z^*), then the interval of incomplete specialization uniquely exists within the interval $(\hat{K}_1^*;\ \hat{K}_2^*)$ and its interior contains the intersection of $pG[F^1(K,\bar{L})]$ and $G[F^2(K,\bar{L})]$.*

Based on the foregoing argument, we can depict the graphs of $g(K)$ like in figure 4.

Acknowledgement. We thank Oded Galor, Bjarne S. Jensen, Peter Skott, Ping Wang, and Kar-yiu Wong for useful comments.

References

Azariadis, Costas (1996). "The Economics of Poverty Traps, Part One: Complete Markets," *Journal of Economic Growth* 1, 449–486.

—— and Allan Drazen (1990). "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105, 501–526.

Barro, Robert J. (1991). "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics* 106, 407–444.

—— and Xavier Sala-i-Martin (1995). *Economic Growth*, New York: McGraw-Hill.

Deckert, W. Davis, and Kazuo Nishimura (1983). "A Complete Characterization of Optimal Growth Paths in an Aggregative Model with a Non-concave Production Function," *Journal of Economic Theory* 31, 332–354.

Durlauf, Steven N. (1993). "Non-ergodic Economic Growth," *Review of Economic Studies* 50, 352–356.

Flemming, John M. (1955). "External Economies and the Doctrine of Balanced Growth, *Economic Journal* 65, 241–256.

Galor, Oded (1996). "Convergence? Inferences from Theoretical Models," *Economic Journal* 106, 1056–1069.

—— and Joseph Zeira (1993). "Income Distribution and Macroeconomics," *Review of Economic Studies* 60, 25–52.

Harrison, Lawrence (1997). *The Pan-American Dream: Do Latin America's Cultural Values Discourage True Partnership with the United States and Canada?* New York: Basic Books.

Jensen, Bjarne S., and Mogens Esrom Larsen (1987). "Growth and Long-run Stability," *Acta Applicandae Mathematicae* 9, 219–237.

Leonard, Daniel, and Ngo Vang Long (1992). *Optimal Control Theory and Static Optimization in Economics*, Cambridge University Press.

Long, Ngo Van, and Kar-yiu Wong (1997). "Endogenous Growth and International Trade: A Survey." (this volume).

Majumdar, Mukul, and Tapan Mitra (1995). "Patterns of Trade and Growth under Increasing Returns: Escape from the Poverty Trap," *Japanese Economic Review* 46, 206–225.

Matsuyama, Kiminori (1991). "Increasing Returns, Industrialization, and the Indeterminacy of Equilibrium," *Quarterly Journal of Economics* 106, 517–550.

Mauro, Paolo (1995). "Corruption and Growth," *Quarterly Journal of Economics* 110, 681–712.

Michel, Phillip (1982). "On the Transversality Condition in Infinite Horizon Optimal Problem," *Econometrica* 50, 975–985.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1989). "Industrialization and the Big Push," *Journal of Political Economy* 97, 1003–1026.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1991). "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics* 108, 503–530.

Nurkse, Ragar (1953). *Problems of Capital Formation in Underdeveloped Countries*, New York: Oxford University Press.

Rosenstein-Rodan, Paul N. (1943), "Problems of Industrialization of Eastern and South-eastern Europe," *Economic Journal* 53, 202–211.

Sakakibara, Eisuke (1993). *Beyond Capitalism: The Japanese Model of Market Economics*, Lanham: University Press of America.

Scitovsky, Tibor (1954). "Two Concepts of External Economies," *Journal of Political Economy* 62, 143–151.

Skiba, A.K. (1978). "Optimal Growth with a Convex-Concave Production Function," *Econometrica* 46, 527–539.

CHAPTER 5

Dynamic Foundations for the Factor Endowment Model of International Trade

Oded Galor and Shoukang Lin

1. Introduction

This chapter establishes dynamic microeconomic foundations for the fundamental propositions of the influential model of international trade theory – the Heckscher-Ohlin model. It analyzes the long-run trade patterns and their implications for factor returns within a comprehensive dynamic general equilibrium model characterized by a two-country two-sector overlapping-generations world where countries differ in their rates of time preference. The chapter develops a two-country, two-sector overlapping-generations model along the lines of the traditional two-sector growth model (e.g., Uzawa (1964), Srinivasan (1964), Oniki and Uzawa (1965), and Shell (1967); see also Jensen and Wang (1997)), and two-sector overlapping-generations model, Galor (1992).

The analysis demonstrates that in a two-country two-sector overlapping-generations world in which countries differ (slightly) in their rates of time preference and the investment good is capital intensive the higher the rate of time preference, the lower the steady-state level of the capital-labor ratio and the lower the steady-state relative price of the capital intensive good. Thus:

a. The low time preference country exports the capital intensive good in the steady-state equilibrium, whereas the high time preference country exports the labor-intensive good.

b. International trade increases the steady-state real return to labor and decreases the steady-state real return to capital in the high time preference country, whereas in the low time preference country the real return to capital increases and the real return to labor decreases.

c. International trade equalizes factor prices across countries along the transition path from the autarkic to the trade steady-state equilibrium as well as in the steady-state trade equilibrium, (for wide range of parameters that generates diversification in production).[1]

1. This result is documented in our earlier working paper, Galor and Lin (1989), and is subsequently established by others.

It should be noted that the capital intensity of the investment good sector is required in order to assure that the perfect-foresight equilibrium is well defined. As was established by Galor (1992), as long as the investment good is capital intensive the dynamical system (under autarky) is characterized by a unique perfect foresight equilibrium. Furthermore, if a non-trivial steady-state equilibrium exists and is unique than the it is a saddle. However, if the consumption good is capital intensive the equilibrium paths in indeterminate (i.e., there exists a continuum of equilibria from a given initial condition), and the model is naturally not well specified. The difference between the stability requirement in the growth model and the overlapping-generations model are partly due to the fact that in the overlapping-generations model saving is a function wage income whereas in the growth model it is a function of aggregate income.[2]

Several attempts to provide dynamic microeconomic foundations for the Heckscher-Ohlin model have been conducted in the literature.[3] Oniki and Uzawa (1965) and Bardhan (1970) extended the two-sector growth model to a two-country world, demonstrating that in a world in which the propensities to save differ across countries, the country with the higher propensity to save exports the capital intensive good in the long run. Stiglitz (1970) demonstrates that, in a two-country two-sector infinite horizon world where the (constant) rates of time preference differ across countries, factor price equalization does not hold in the long run. Findlay (1970) establishes the relationship between trade patterns of a small three-sector economy, and saving propensities and rates of population growth, and Matsuyama (1988) considers the trade patterns of a small three-sector economy in a life cycle model.

In contrast to Findlay (1970) and Matsuyama (1988), the current study considers large countries, permitting a comprehensive general equilibrium analysis in which the terms of trade dynamics are endogenously determined. Unlike Oniki and Uzawa (1965), Bardhan (1970), and Findlay (1970), individuals' savings are the outcome of an intertemporal optimization. As opposed to Stiglitz (1970)'s infinite horizon model, where the long-run equilibrium is characterized by the equalization of the rate of return to capital and the rate of time preference, the choice of an

2. See Galor (1996) for a related discussion.

3. Dynamic microeconomic foundations for various characteristics of international economics have been established in the literature. Buiter (1981) establishes dynamic foundations for the patterns of international lending and borrowing, within a framework of two Diamond overlapping-generations economies which differ in their rates of time preference. Galor (1986) establishes dynamic foundations for the patterns of international labor migration within a similar setting, and Eaton (1987) provides the foundations for the specific-factors model.

overlapping-generations model, in which individuals are finitely lived, allows for factor price equalization, despite differences in time preference across countries.[4]

2. The Autarky

Consider a world where economic activity extends over infinite discrete time and is conducted under perfect competition and certainty. In every period, a perishable consumption good and an investment good are produced, using two factors, capital and labor, in the production process. Capital is fully depreciated after a single period and there is no population growth.[5] That is, the endowment of labor at time t, L_t, is exogenously given and is invariant over time. Thus, $L_t = L$, $\forall t$. The stock of capital at time $t+1$, K_{t+1}, is equal to the output of the investment good produced at t, Y_t. Thus $K_{t+1} = Y_t$, where K_0 is exogenously given.[6]

2.1 Production

Production technologies employed in both the consumption good sector and the investment good sector exhibit constant returns to scale. The output of the consumption good and the output of the investment good produced at time t, X_t and Y_t, respectively, are

$$X_t = F_x(K_t^x, L_t^x) = L_t^x F_x(K_t^x/L_t^x, 1) \equiv L_t^x f_x(k_t^x), \tag{1}$$
$$Y_t = F_y(K_t^y, L_t^y) = L_t^y F_y(K_t^y/L_t^y, 1) \equiv L_t^y f_y(k_t^y), \tag{2}$$

where $k_t^j \equiv K_t^j/L_t^j$ is the capital-labor ratio in sector j at time t, $j = x, y$. The production function $f_j : \mathbf{R}_+ \to \mathbf{R}_+$ is twice continuously differentiable, positive, increasing, and strictly concave. That is, $f_j(k_t^j) > 0$, $f_j'(k_t^j) > 0$, $f_j''(k_t^j) < 0$, $\forall k_t^j > 0$. In addition, it satisfies the Inada conditions $\lim_{k^j \to \infty} f_j'(k^j) = 0$ and $\lim_{k^j \to 0} f_j'(k^j) = \infty$.

4. As is discussed by Stiglitz (1970), if the rate of time preference is not constant factor price equalization is feasible in the optimal growth model as well.

5. The analysis is perfectly applicable under any feasible rates of capital depreciation and population growth.

6. The production side (section 2.1) follows the traditional two-sector growth model (e.g., Uzawa (1964), Srinivasan (1964), and Shell (1967)). The consumption and savings (section 2.2) and the dynamic equilibrium (section 2.3) differ, however, due to the overlapping-generations structure and the finiteness of lifetime. Furthermore, given rational expectations, the determinacy of the dynamic equilibrium requires the existence of a saddle path stable steady-state equilibrium.

Furthermore, the investment good is capital intensive:[7]

$$k^x(\omega_t) < k^y(\omega_t), \quad \forall \omega_t > 0, \tag{3}$$

where ω_t is the wage-rental ratio at time t.

Suppose that both goods are produced and that labor and capital are perfectly mobile across sectors.[8] The demands for labor and capital are therefore characterized by the first-order conditions for profit maximization

$$r_t = p_t f'_x(k_t^x) = f'_y(k_t^y), \tag{4}$$

$$w_t = p_t[f_x(k_t^x) - f'_x(k_t^x)k_t^x] = f_y(k_t^y) - f'_y(k_t^y)k_t^y, \tag{5}$$

where p_t is the price of the consumption good, r_t is the return to capital and w_t is the wage rate, at time t. The investment good is the numeraire.

The wage-rental ratio, $\omega_t \equiv w_t/r_t$, is therefore,

$$\omega_t = \varphi^j(k_t^j), \quad j = x, y,$$

where $d\omega_t/dk_t^j > 0$, $j = x, y$. Hence, $\varphi^j(k_t^j)$, which is strictly increasing in k_t^j, is invertible, and $k_t^j = (\varphi^j)^{-1}(\omega_t) \equiv k^j(\omega_t)$, $j = x, y$.
The price of the consumption good at time t, p_t, is

$$p_t = p(\omega_t) \equiv \frac{f'_y(k^y[\omega_t])}{f'_x(k^x[\omega_t])}, \tag{6}$$

where $p'(\omega_t) < 0$. Furthermore, there exists a single valued function $\omega\colon \mathbf{R}_+ \to \mathbf{R}_+$ such that $\omega_t = \omega(p_t)$.

Thus, given the price of the consumption good at time t, p_t, the capital-labor ratios in both sectors, k_t^y, and k_t^x, the wage rate, w_t, and the interest rate, r_t, are uniquely determined.

$$w_t = f_y(k^y[\omega(p_t)]) - f'_y(k^y[\omega(p_t)])k^y[\omega(p_t)] \equiv w(p_t), \tag{7}$$

$$r_t = f'_y(k^y[\omega(p_t)]) \equiv r(p_t). \tag{8}$$

Furthermore, given the per worker capital stock, k_t, where $k_t \equiv K_t/L$, the per-worker production of the consumption good, x_t, and the investment good, y_t, is uniquely determined.

$$x_t = \frac{k_t - k_t^y}{k_t^x - k_t^y} f_x(k_t^x) \equiv x(p_t, k_t), \tag{9}$$

$$y_t = \frac{k_t^x - k_t}{k_t^x - k_t^y} f_y(k_t^y) \equiv y(p_t, k_t). \tag{10}$$

7. If the consumption good is capital intensive, the dynamic equilibrium is indeterminate as was established by Galor (1992).

8. The boundary conditions on preferences and technologies guarantee that in autarky both goods are produced in every period.

Lemma 1. *(Stolper-Samuelson theorem). Suppose that* $x_t > 0$ *and* $y_t > 0$. *If* $k_t^x(\omega_t) < k_t^y(\omega_t)$, $\forall \omega_t > 0$, *then* $\frac{d\omega_t}{dp_t} > 0$, *and* $\frac{dw_t}{dp_t}\frac{p_t}{w_t} > 1$, $\frac{dr_t}{dp_t} < 0$.

Thus, as long as both goods are produced, if the investment good is capital intensive, an increase in the relative price of the consumption good raises the real wage rate and lowers the real interest rate.

Lemma 2. *(Rybczynski theorem). Suppose that* $x_t > 0$ *and* $y_t > 0$. *If* $k_t^x(\omega_t) < k_t^y(\omega_t)$, $\forall \omega_t > 0$, *then* $\frac{\partial x_t}{\partial k_t} < 0$, $\frac{\partial y_t}{\partial k_t} > 0$, *and* $\frac{\partial y_t}{\partial k_t}\frac{k_t}{y_t} > 1$.

Thus, as long as both goods are produced, if the investment good is capital intensive, a marginal increase in the capital-labor ratio, given goods' prices, decreases the production of the consumption good and increases the production of the investment good.

2.2 Consumption and Savings

In every period t, L individuals are born. Individuals are identical within as well as across time. Individuals live two periods. In the first period, they work and earn the competitive market wage, w_t, and in the second period they are retired. Individuals born at t are characterized by their intertemporal utility function

$$U(c_t^t, c_{t+1}^t) = u(c_t^t) + \frac{1}{1+\rho}\, u(c_{t+1}^t), \tag{11}$$

defined over non-negative consumption during the first and the second periods of their life. The rate of time preference, $\rho \geq 0$.

The intertemporal utility function is twice continuously differentiable, monotonically increasing, and strictly quasi-concave, over the interior of consumption set. Furthermore, $\lim_{c_t^t \to 0} u'(c_t^t) = \lim_{c_{t+1}^t \to 0} u'(c_{t+1}^t) = \infty$.

During the first period of their lifetime individuals born at time t supply their unit-endowment labor inelastically. The resulting wage income, w_t, is allocated between first period consumption, c_t^t, and savings s_t.

$$s_t = w_t - p_t\, c_t^t, \tag{12}$$

where p_t is the price of the consumption good at time t.

Individuals save by purchasing the investment good which is the only store of value in the economy. Savings earn the given *gross* rate of return, r_{t+1} (i.e., the marginal productivity of capital at time $t+1$) in the following period and enable individuals to consume during retirement. Second period consumption of an individual of generation t, c_{t+1}^t, is therefore

$$c_{t+1}^t = r_{t+1} s_t / p_{t+1}. \tag{13}$$

The level of savings is chosen so as to maximize the intertemporal utility function.

$$\begin{aligned} s_t &= s(w_t, p_t, p_{t+1}/r_{t+1}; \rho) \\ &= \operatorname{argmax} \left\{ u[(w_t - s_t)/p_t] + \frac{1}{1+\rho} u[(r_{t+1}/p_{t+1}) s_t] \right\} \\ &\quad \text{s.t. } 0 \le s_t \le w_t, \end{aligned} \tag{14}$$

where r_{t+1} and p_{t+1} are the rationally anticipated return to capital and the price level in period $t+1$.

Given (p_t, p_{t+1}) and consequently $w_t = w(p_t)$ and $r_{t+1} = r(p_{t+1})$, the properties of the utility function imply that $s(w_t, p_t, p_{t+1}/r_{t+1}; \rho)$ exists and is unique.

$$s(w_t, p_t, p_{t+1}/r_{t+1}; \rho) \equiv S(p_t, p_{t+1}; \rho). \tag{15}$$

Given the time separability of the utility function, first and second period consumption are normal goods and consequently savings are an increasing function of the wage rate, a non-decreasing function of the interest rate and a decreasing function of the time preference. Furthermore, it is assumed that savings are a non-decreasing function of the rental rate.[9] That is,

$$\frac{\partial s_t}{\partial w_t} > 0, \quad \frac{\partial s_t}{\partial r_{t+1}} \ge 0, \quad \frac{\partial s_t}{\partial \rho} < 0. \tag{16}$$

Lemma 3. *Let $S(p_t, p_{t+1}; \rho) \equiv s(w(p_t), p_t, r(p_{t+1})/p_{t+1}; \rho)$. If $k_t^x(\omega_t) < k_t^y(\omega_t)$, $\forall \omega_t > 0$, and $\frac{\partial s_t}{\partial r_{t+1}} > 0$, then $\frac{\partial S(p_t, p_{t+1})}{\partial p_t} > 0$ and $\frac{\partial S(p_t, p_{t+1})}{\partial p_{t+1}} < 0$.*

Proof. See Galor (1992). □

2.3 Dynamic Equilibrium

The evolution of the capital stock is governed by the production of the investment good.

$$k_{t+1} = y(p_t, k_t). \tag{17}$$

9. $\frac{\partial s_t}{\partial r_{t+1}} \ge 0$ if and only if $u'(c_t)u'(c_{t+1}) \ge -[u''(c_{t+1})u'(c_t)s_t r_{t+1}]/p_{t+1}$.

Furthermore, the clearance of the goods' markets in every period t requires (recalling Walras' law) that the demand for the capital good (i.e., savings) will equal the production of the capital good. Namely,

$$S(p_t, p_{t+1}; \rho) = y(p_t, k_t). \tag{18}$$

Suppose that $\partial s_t/\partial r_{t+1} > 0$.[10] Noting Lemma 3, $\partial S/\partial p_{t+1} < 0$, and consequently it follows from (18) that

$$p_{t+1} = \phi(p_t, k_t; \rho). \tag{19}$$

Lemma 4. *If $k_t^x(\omega_t) < k_t^y(\omega_t)$, $\forall \omega_t > 0$, then $\frac{\partial \phi}{\partial \rho}(p_t, k_t; \rho) < 0$, and $\frac{\partial \phi}{\partial k_t}(p_t, k_t; \rho) < 0$.*

Proof. See Appendix. □

Definition 1. *An autarkic dynamic equilibrium is a sequence $\{p_t, k_t\}_{t=0}^{\infty}$ under which*

$$\begin{aligned} k_{t+1} &= y(p_t, k_t), \\ p_{t+1} &= \phi(p_t, k_t; \rho), \end{aligned}$$

where k_0 is exogenously given.[11]

Definition 2. *An autarkic steady state equilibrium is a pair $\{\overline{p}, \overline{k}\}$ under which*

$$\begin{aligned} \overline{k} &= y(\overline{p}, \overline{k}), \\ \overline{p} &= \phi(\overline{p}, \overline{k}; \rho). \end{aligned}$$

Remark 1. The boundary conditions on preferences and technologies guarantee a strictly positive production of both goods at time t, as long as $k_t > 0$. ∇

Lemma 5. *If the investment good is capital intensive and there exists a unique noon-trivial steady-state equilibrium then it is a saddle.*[12]

10. The case in which $\partial s_t/\partial r_{t+1} = 0$, as implied by log-linear preferences is analyzed in section 4. This separation is necessary since under this condition the dynamics of the system are characterized by a single non-linear difference equation, rather than a system of two non-linear difference equations.

11. Definition 1 does not apply to the case in which $\partial S/\partial r_{t+1} = 0$, (e.g., log-linear utility functions). Savings in this case are a constant proportion of the wage, w_t, regardless of future prices and interest rates. Consequently, (19) does not hold and the system is characterized by a single nonlinear difference equation. Section 4 analyzes the case in which $\partial S/\partial r_{t+1} = 0$ and demonstrates that the qualitative results follow.

12. Sufficient conditions for the existence of a unique non-trivial steady-state equilibrium are provided by Galor (1992).

Proof. See Galor (1992). □

Lemma 6. *Let*

$$A \equiv \begin{pmatrix} \frac{\partial y}{\partial k}(\overline{p},\overline{k}) & \frac{\partial y}{\partial p}(\overline{p},\overline{k}) \\ \frac{\partial \phi}{\partial k}(\overline{p},\overline{k}) & \frac{\partial \phi}{\partial p}(\overline{p},\overline{k}) \end{pmatrix} \text{ and } I \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The steady-state equilibrium is a saddle if and only if $\det(I - A) < 0$.

Proof. See Appendix. □

Remark 2. If the conditions for the saddle-path stability are satisfied around the steady-state equilibrium $(\bar{p}, \bar{k})$, then for $\epsilon > 0$ sufficiently small and $\forall k_0 \in B_\epsilon(\bar{k})$, the dynamic equilibrium $\{k_t, p_t\}_{t=0}^\infty$ is uniquely determined.[13] ∇

Proposition 1. *Consider a locally saddle-path stable steady-state autarkic equilibrium. If* $k_t^x(\omega_t) < k_t^y(\omega_t)$, $\forall \omega_t > 0$, *then* $\frac{d\overline{k}}{d\rho} < 0$ *and* $\frac{d\overline{p}}{d\rho} < 0$.

Proof. Differentiating the steady-state equilibrium conditions with respect to ρ

$$\begin{pmatrix} \frac{d\overline{k}}{d\rho} \\ \frac{d\overline{p}}{d\rho} \end{pmatrix} = \begin{pmatrix} \frac{\partial \overline{y}}{\partial k} & \frac{\partial \overline{y}}{\partial p} \\ \frac{\partial \overline{\phi}}{\partial k} & \frac{\partial \overline{\phi}}{\partial p} \end{pmatrix} \begin{pmatrix} \frac{d\overline{k}}{d\rho} \\ \frac{d\overline{p}}{d\rho} \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\partial \overline{\phi}}{\partial \rho} \end{pmatrix}. \tag{20}$$

Then

$$\begin{aligned} \begin{pmatrix} \frac{d\overline{k}}{d\rho} \\ \frac{d\overline{p}}{d\rho} \end{pmatrix} &= (I - A)^{-1} \begin{pmatrix} 0 \\ \frac{\partial \overline{\phi}}{\partial \rho} \end{pmatrix} \\ &= \frac{1}{\det(I - A)} \begin{pmatrix} 1 - \frac{\partial \overline{\phi}}{\partial p} & \frac{\partial \overline{y}}{\partial p} \\ \frac{\partial \overline{\phi}}{\partial k} & 1 - \frac{\partial \overline{y}}{\partial k} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{\partial \overline{\phi}}{\partial \rho} \end{pmatrix}. \end{aligned} \tag{21}$$

Thus,

$$\frac{d\overline{k}}{d\rho} = \frac{1}{\det(I - A)} \frac{\partial \overline{y}}{\partial p} \frac{\partial \overline{\phi}}{\partial \rho}, \tag{22}$$

$$\frac{d\overline{p}}{d\rho} = \frac{1}{\det(I - A)} (1 - \frac{\partial \overline{y}}{\partial k}) \frac{\partial \phi}{\partial \rho}. \tag{23}$$

13. This observation requires a global analysis of the dynamic system. As is shown in Galor (1992) if $\frac{\partial S(p_t, p_t; \rho)}{\partial p_t} + \frac{\partial S(p_t, p_t; \rho)}{\partial p_{t+1}} - \frac{\partial y(p_t, k_t)}{\partial p_t} > 0$, all steady-state equilibria lie along a unique dynamic path and thus Remark 2 follows. The system, however, may be characterized by multiple, saddle-path stable, non-trivial steady-state equilibria.

Since $\frac{\partial \bar{y}}{\partial k} > 1$, as follows from Lemma 2, and $\frac{\partial y}{\partial p} < 0$ (a movement along the production possibility frontier resulting from changes in relative prices), the proposition follows from Lemma 4 and Lemma 6. □

Thus, if the investment good is capital intensive, the higher is the rate of time preference, the lower is the steady-state level of the capital-labor ratio and the lower is the price of the consumption good.

3. Trade Equilibrium

Consider a world that consists of two countries, $i = A, B$, which are identical in all respects except for their rates of time preference. The countries are engaged in free-trade in goods. International labor migration and international lending and borrowing, however, are prohibited.[14] A trade equilibrium requires the clearance of the world market for the capital good. Domestic investment, however, must equal domestic savings in each country.

Definition 3. *A dynamic trade equilibrium (under diversification in production in each country) is a sequence $\{p_t, k_t^A, k_t^B\}_{t=0}^{\infty}$ under which*

$$\begin{aligned} k_{t+1}^A &= S(p_t, p_{t+1}; \rho^A), \\ k_{t+1}^B &= S(p_t, p_{t+1}; \rho^B), \\ y(p_t, k_t^A) + y(p_t, k_t^B) &= S(p_t, p_{t+1}; \rho^A) + S(p_t, p_{t+1}; \rho^B), \end{aligned}$$

where k_0^A, k_0^B and p_0 are exogenously given.[15]

Definition 4. *A steady-state trade equilibrium (under diversification in production in each country) is a triplet $\{(k^A)^*, (k^B)^*, p^*\}$ under which*

$$\begin{aligned} (k^A)^* &= S(p^*; \rho^A), \\ (k^B)^* &= S(p^*; \rho^B), \\ y(p^*, (k^A)^*) + y(p^*, (k^B)^*) &= S(p^*; \rho^A) + S(p^*; \rho^B). \end{aligned}$$

3.1 The Patterns of Trade

Theorem 1. *(Trade Patterns). Consider a steady-state trade equilibrium of a two-country world, in which countries are identical in every respect except for a small difference in their rates of time preference.*

14. In accordance with the traditional literature concerning trade pattern, trade in goods is permitted and international lending and borrowing is prohibited.
15. If international lending and borrowing is permitted, then the aggregate capital accumulation in the world economy equals the aggregate world savings.

Then, the low time preference country exports the capital intensive good whereas the high time preference country exports the labor intensive good.

Proof. Without loss of generality, let $\rho^A > \rho^B$, (i.e., country A is the high time preference country). Let $\rho^A = \rho^B + d\rho$, where $d\rho > 0$ is a sufficiently small constant. Then as follows from (16), in a steady-state free-trade equilibrium, country A saves less than country B, i.e.,

$$S(p^*;\rho^A) < S(p^*;\rho^B). \tag{24}$$

Following Definition 4 and (31)

$$(k^B)^* = (k^A)^* + dk, \quad dk > 0. \tag{25}$$

$dk > 0$ can be chosen to be sufficiently small by an appropriate choice of $d\rho$. Thus, the clearance of the world market for the investment good implies therefore that

$$y(p^*,(k^A)^*) + y(p^*,(k^A)^* + dk) = 2(k^A)^* + dk. \tag{26}$$

A Taylor expansion of $y(p^*,(k^A)^* + dk)$ yields

$$y(p^*,(k^A)^* + dk) = y(p^*,(k^A)^*) + \frac{\partial y(p^*,\tilde{k})}{\partial k}\,dk, \tag{27}$$

for some $\tilde{k} \in [(k^A)^*,\ (k^A)^* + dk]$. Consequently, using (26)–(27)

$$2[y(p^*,(k^A)^*) - (k^A)^*] + \left[\frac{\partial y(p^*,\tilde{k})}{\partial k} - 1\right] dk = 0. \tag{28}$$

To establish the proposition it is sufficient to show that if $k_t^x < k_t^y$ (i.e., the investment good is capital intensive) then the high time preference country (country A) imports the (capital intensive) investment good, i.e.,

$$y(p^*,(k^A)^*) < S(p^*,\rho^A) = (k^A)^*. \tag{29}$$

Namely, it is sufficient to show that the steady-state domestic demand for capital good (savings) is higher than the domestic production of the capital good. Noting that $\frac{\partial y(p^*,\tilde{k})}{\partial k} > 1$,[16] (29) follows from (28) and the proposition follows. □

16. By Lemma 2, at the steady-state equilibrium $\frac{\partial y(\bar{p},\bar{k})}{\partial k} > 1$. Therefore, continuity guarantees that $\frac{\partial y}{\partial k} > 1$ in a small neighborhood of the steady-state equilibrium.

3.2 Trade and Factor Returns

Proposition 2. *Consider a two-country world in which countries are identical in every respect except for a small difference in their rates of time preference. If the autarkic steady-state equilibria are locally saddle-path stable and the investment good is capital intensive, then the autarkic steady-state equilibrium in the high time preference country is characterized by a lower capital-labor ratio and a lower relative price for the consumption good.*

Proof. The proposition is a corollary of Proposition 1. □

Lemma 7. $\frac{\partial y}{\partial p^i}(\overline{p}^i, \overline{k}^i) + \left[\frac{\partial y}{\partial k^i}(\overline{p}^i, \overline{k}^i) - 1\right] \frac{\partial S}{\partial p^i}(\overline{p}^i; \rho^i) < 0,\ \forall i,\ i = A, B.$

Proof. See Appendix. □

Proposition 3. *Consider a two-country world in which countries are identical in all respects except for a small difference in their rates of time preference. If the autarkic and the trade steady-state equilibria are locally saddle-path stable, then the relative price of the consumption good in the steady-state trade equilibrium lies between the relative prices of the two economies in the autarkic steady-state equilibrium.*

Proof. Without loss of generality let $\rho^A > \rho^B$. It follows from Proposition 2 that $\overline{p}^A < \overline{p}^B$ and $\overline{k}^A < \overline{k}^B$.

Suppose that the proposition does not hold. In particular, suppose that $\overline{p}^A < \overline{p}^B < p^*$. Then $\exists d\overline{p}^i > 0, \forall i, i = A, B$, such that

$$p^* = \overline{p}^i + d\overline{p}^i, \quad \forall i,\ i = A, B. \tag{30}$$

Following the definition of a steady-state trade equilibrium

$$\begin{gathered}(k^i)^* = S(\overline{p}^i + d\overline{p}^i; \rho^i),\ \ \forall i,\ \ i = A, B, \\ \sum_{i=A,B} y(\overline{p}^i + d\overline{p}^i, (k^i)^*) - \sum_{i=A,B} S(\overline{p}^i + d\overline{p}^i; \rho^i) = 0.\end{gathered} \tag{31}$$

A Taylor expansion of $S(\overline{p}^i + d\overline{p}^i; \rho^i)$ yields

$$(k^i)^* = S(\overline{p}^i + d\overline{p}^i; \rho^i) = S(\overline{p}^i; \rho^i) + \frac{\partial S}{\partial p^i}(\tilde{p}^i; \rho^i)d\overline{p}^i, \tag{32}$$

for some $\tilde{p}^i \in [\overline{p}^i,\ \overline{p}^i + d\overline{p}^i]$. Using the definition of an autarkic steady-state equilibrium, it follows from (33) that

$$(k^i)^* = \overline{k}^i + \frac{\partial S}{\partial p^i}(\tilde{p}^i; \rho^i)d\overline{p}^i. \tag{33}$$

A Taylor expansion of $y(\overline{p}^i + d\overline{p}^i, (k^i)^*)$ noting (34), yields

$$\begin{aligned} y(\overline{p}^i + d\overline{p}^i, (\overline{k}^i)^*) = y(\overline{p}^i, \overline{k}^i) \; &+ \; \frac{\partial y}{\partial p^i}(\hat{p}^i, \hat{k}^i) d\overline{p}^i \\ &+ \; \frac{\partial y}{\partial k^i}(\hat{p}^i, \hat{k}^i) \frac{\partial S}{\partial p^i}(\tilde{p}^i, \rho^i) d\overline{p}^i, \end{aligned} \qquad (34)$$

where $\hat{p}^i \in [\overline{p}^i, \; \overline{p}^i + d\overline{p}^i]$ and $\hat{k}^i \in [\overline{k}^i, \; \overline{k}^i + \frac{\partial S}{\partial p^i}(\tilde{p}^i; \rho^i) d\overline{p}^i]$. Thus, noting that $\overline{k}^i = y(\overline{p}^i, \overline{k}^i)$, it follows from (33)–(35) and Lemma 7 that for all i, $i = A, B$

$$\begin{aligned} y(p^*, (k^i)^*) - S(p^*, \rho^i) \; = \; & \frac{\partial y}{\partial p^i}(\hat{p}^i, \hat{k}^i) + \\ & \left[\frac{\partial y}{\partial k^i}(\hat{p}^i, \hat{k}^i) - 1 \right] \frac{\partial S}{\partial p^i}(\tilde{p}^i; \rho^i) d\overline{p}^i < 0, \end{aligned} \qquad (35)$$

in contradiction to the steady-state trade equilibrium condition (32). Note that Lemma 7 is applicable since small differences in ρ are considered.

Similarly if $p^* < \overline{p}^A < \overline{p}^B$, or $\overline{p}^A = p^*$, or $\overline{p}^B = p^*$ a contradiction to (32) can be established. Thus, $\overline{p}^B < p^* < \overline{p}^A$. □

Theorem 2. *(Trade and Factor Returns). Consider a two-country world, in which countries are identical in all respects, except for a small difference in the rates of time preference. If the autarkic steady-state equilibria are locally saddle-path stable, then trade raises the steady-state real wage and lowers the steady-state real return to capital in the high time preference country, and lowers the steady-state real wage and raises the steady-state real return to capital in the low time preference country.*

Proof. If $k_t^x < k_t^y$ and $\rho^A > \rho^B$, then according to Proposition 3, $\overline{p}^A < p^* < \overline{p}^B$. Thus, noting Lemma 1 the theorem follows. □

Theorem 3. *(Factor Price Equalization). Consider a two-country world in which countries are identical in all respects except for a small difference in their rates of time preference. Then, if both goods are produced in each country, trade equalizes factor prices across countries.*

Proof. Following (7)–(8), as long as both goods are produced in country i, $w_t^i = w(p^i)$ and $r_t^i = r(p^i)$, $\forall i$, $i = A, B$, where $w : \mathbf{R}_+ \to \mathbf{R}_+$ and $r : \mathbf{R}_+ \to \mathbf{R}_+$ are single valued functions. Thus, the equalization of good prices in every period results in the equalization of factor prices. □

Remark 3. Unlike the infinite horizon model (with constant rate of time preference) in which trade leads in the long run to specialization in production of at least one of the countries, diversification in production is feasible in a two-sector overlapping-generations economy. Consequently, long-run factor price equalization which fails to exist in an infinite horizon model holds in the overlapping-generations model. ∇

4. An Extension

The dynamic system of the described economy does not apply when savings are not a function of interest rates; an important case that includes log-linear utility functions. The purpose of this section is to demonstrate that the entire results follow under this case as well.

If $\frac{\partial S_t}{\partial r_{t+1}} = 0$, then $s_t = S(p_t; \rho)$. Thus, the dynamics of the economy are characterized by the system

$$y(p_t, k_t) = S(p_t; \rho), \tag{36}$$

$$k_{t+1} = y(p_t, k_t). \tag{37}$$

It follows from (37) that

$$p_t = \xi(k_t; \rho). \tag{38}$$

Substituting (39) into (38), the evolution of the economy is governed by the first order difference equation

$$k_{t+1} = \psi(k_t; \rho) = y(\xi(k_t; \rho), k_t). \tag{39}$$

Definition 5. *Consider an overlapping generations economy in which savings do not depend on interest rates, an autarkic equilibrium is a sequence $\{k_t\}_{t=0}^{\infty}$ under which*

$$k_{t+1} = \psi(k_t; \rho),$$

where k_0 is exogenously given.

Definition 6. *Consider an overlapping generations economy in which savings do not depend on interest rates, an autarkic steady state equilibrium is $\bar{k}$ under which*

$$\bar{k} = \psi(\bar{k}; \rho).$$

Lemma 8. *A nontrivial autarkic steady-state equilibrium exists if*

$$\lim_{k_t \to 0} \frac{\partial \psi(k_t;\rho)}{\partial k_t} > 1 \quad \textit{and} \quad \lim_{k_t \to \infty} \frac{\partial \psi(k_t;\rho)}{\partial k_t} < 1.$$

Proof. Noting that $\psi(0;\rho) = 0$, the lemma follows from Definition 6.[17] □

Remark 4. If production technologies are of the Cobb-Douglas type, $\psi(k_t;\rho)$ is an increasing, strictly concave function of k_t that satisfies the above conditions. Thus, a nontrivial equilibrium exists and is globally stable. See Galor (1992). ∇

Proposition 4. *Let the steady state equilibrium of the dynamical system (40) be locally stable, i.e.,* $-1 < \frac{\partial \psi(\bar{k};\rho)}{\partial k} < 1$. *If* $k^x < k^y$, *then*

$$\frac{d\bar{k}}{d\rho} < 0, \quad \frac{d\bar{p}}{d\rho} < 0.$$

Proof. Differentiating (40) with respect to k_t around the steady state equilibrium,

$$\frac{\partial \psi(\bar{k};\rho)}{\partial k} = \frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial p} \frac{\partial \xi(\bar{k};\rho)}{\partial k} + \frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial k}. \tag{40}$$

Noting that

$$\frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial p} < 0, \quad \frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial k} > 1, \quad \text{and} \quad \frac{\partial \psi(\bar{k};\rho)}{\partial k} < 1. \tag{41}$$

It follows that $\frac{\partial \xi(\bar{k};\rho)}{\partial k} > 0$. Using (37) and the implicit function theorem

$$\frac{\partial \xi(\bar{k};\rho)}{\partial k} = \frac{\frac{\partial y(\bar{p};\bar{k})}{\partial k}}{\frac{\partial S(\bar{p};\rho)}{\partial p} - \frac{\partial y(\bar{p},\bar{k})}{\partial p}}. \tag{42}$$

Therefore,

$$\frac{\partial S(\bar{p};\rho)}{\partial p} - \frac{\partial y(\bar{p};\bar{k})}{\partial p} > 0. \tag{43}$$

Furthermore, it follows from (37) that

$$\frac{\partial \xi(\bar{k};\rho)}{\partial \rho} = -\frac{\frac{\partial S(\bar{p};\rho)}{\partial \rho}}{\frac{\partial S(\bar{p};\rho)}{\partial p} - \frac{\partial y(\bar{p},\bar{k})}{\partial p}} > 0. \tag{44}$$

17. See Galor and Ryder (1989) for the existence of a nontrivial steady-state equilibrium in a one-good overlapping-generations world.

Totally differentiating (40), it follows from (45) and the stability condition that

$$\frac{d\bar{k}}{d\rho} = \frac{\frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial p}\frac{\partial \xi(\bar{k};\rho)}{\partial \rho}}{1 - \frac{\partial \psi(\bar{k};\rho)}{\partial k}} < 0. \tag{45}$$

Totally differentiate (39), noting (41), (43), and (46)

$$\frac{d\bar{p}}{d\rho} = \frac{\partial \xi(\bar{k};\rho)}{\partial \rho}\frac{1 - \frac{\partial y(\xi(\bar{k};\rho),\bar{k})}{\partial k}}{1 - \frac{\partial \psi(\bar{k};\rho)}{\partial k}} < 0. \tag{46}$$

□

Thus, for $\partial S/\partial r_{t+1} = 0$, Proposition 1 holds. Since in section 3 no assumptions are made about the functional form of the saving function, the entire results follow.

5. Concluding Remarks

This chapter develops a two-country, two-sector overlapping-generations model along the lines of the traditional two-sector growth model. The chapter establishes dynamic microeconomic foundations for the fundamental propositions of international trade theory. The analysis is conducted within a dynamic general equilibrium model of a two-country two-sector overlapping-generations world where countries differ in their rates of time preference. The study demonstrates that in the long run the low time preference country exports the capital intensive good whereas the high time preference country exports the labor intensive good. Furthermore, international trade increases the long-run real wage and decreases the long-run real interest rate in the high time preference country, and equalizes (under some configurations) factor prices across countries.

Appendix

Proof of Lemma 4. Totally differentiating (18), it follows that

$$\frac{\partial p_{t+1}}{\partial \rho} = \frac{\partial \phi}{\partial \rho} = -\frac{\partial S/\partial \rho}{\partial S/\partial p_{t+1}}, \tag{A.1}$$

$$\frac{\partial p_{t+1}}{\partial k_t} = \frac{\partial \phi}{\partial k_t} = \frac{\partial y/\partial k_t}{\partial S/\partial p_{t+1}}. \tag{A.2}$$

Thus, noting (16), Lemmas 2–3, the lemma follows. □

Proof of Lemma 6. The proof of Lemma 6 requires the establishment of the following two lemmas.

Lemma A1. *The linear operator A has two distinct real eigenvalues.*

Proof of Lemma A1. Let $c(\lambda)$ be the characteristic polynomial of A,

$$c(\lambda) = \det(A - \lambda I) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A. \tag{A.3}$$

The eigenvalues of A are therefore

$$\lambda_{1,2} = [\operatorname{tr} A \pm \sqrt{\Delta}]/2, \tag{A.4}$$

where $\Delta \equiv \operatorname{tr} A^2 - 4\det A$. Following the definition of A and rearranging terms,

$$\Delta = \left[\frac{\partial y(\bar{p},\bar{k})}{\partial k} - \frac{\partial \phi(\bar{p},\bar{k})}{\partial p}\right]^2 + 4\,\frac{\partial y(\bar{p},\bar{k})}{\partial p}\frac{\partial \phi(\bar{p},\bar{k})}{\partial k}. \tag{A.5}$$

Noting that $\frac{\partial y(\bar{p},\bar{k})}{\partial p} < 0$, it follows from Lemma 3 that $\Delta > 0$ and consequently the linear operator A has two distinct real eigenvalues. □

Lemma A2. *The dynamical system is non-oscillatory around the steady-state equilibrium.*

Proof of Lemma A2. The dynamical system is non-oscillatory if both eigenvalues are non-negative. Following (A4)

$$\lambda_1 + \lambda_2 = \operatorname{tr} A, \tag{A.6}$$

$$\lambda_1 \cdot \lambda_2 = \det A, \tag{A.7}$$

where as follows from the proof of Lemma 4 and the definition of A

$$\operatorname{tr} A = \frac{\partial y(\bar{p},\bar{k})}{\partial k} + \frac{\frac{\partial y(\bar{p},\bar{k})}{\partial p} - \frac{\partial S(\bar{p},\bar{p};\rho)}{\partial p_t}}{\frac{\partial S(\bar{p},\bar{p};\rho)}{\partial p_{t+1}}}, \tag{A.8}$$

$$\det A = -\frac{\frac{\partial y(\bar{p},\bar{k})}{\partial k}\frac{\partial S(\bar{p},\bar{p};\rho)}{\partial p_t}}{\frac{\partial S(\bar{p},\bar{p};\rho)}{\partial p_{t+1}}}. \tag{A.9}$$

Thus, as follows from Lemmas 2–3, $\operatorname{tr} A > 0$, and $\det A > 0$, and consequently $\lambda_1 > 0$ and $\lambda_2 > 0$. □

Following Lemmas A1–A2, the steady-state equilibrium is saddle-path stable if and only if $\lambda_1 > 1 > \lambda_2$. Using (A4), $\lambda_1 > 1 > \lambda_2$ if and only if $1 - \operatorname{tr} A + \det A = \det(I - A) < 0$. Thus, Lemma 6 follows. □

Proof of Lemma 7. Totally differentiating (17)–(18) evaluated at the steady-state equilibrium of country i, $(\overline{p}^i, \overline{k}^i)$, it follows that

$$d\overline{k}^i = \frac{\partial S}{\partial p^i}(\overline{p}^i; \rho^i) d\overline{p}^i + \frac{\partial S}{\partial \rho^i}(\overline{p}^i; \rho^i) d\rho^i, \tag{A.10}$$

$$\frac{\partial y}{\partial p^i}(\overline{p}^i, \overline{k}^i) d\overline{p}^i + \frac{\partial y}{\partial k^i}(\overline{p}^i, \overline{k}^i) d\overline{k}^i = \frac{\partial S}{\partial p^i}(\overline{p}^i; \rho^i) d\overline{p}^i + \frac{\partial S}{\partial \rho^i}(\overline{p}^i; \rho^i) d\rho^i. \tag{A.11}$$

Thus,

$$\frac{d\overline{p}^i}{d\rho^i} = \frac{\left[1 - \frac{\partial y}{\partial k^i}(\overline{p}^i, \overline{k}^i)\right] \frac{\partial S}{\partial \rho^i}(\overline{p}^i; \rho^i)}{\frac{\partial y}{\partial p^i}(\overline{p}^i, \overline{k}^i) + \left[\frac{\partial y}{\partial k^i}(\overline{p}^i, \overline{k}^i) - 1\right] \frac{\partial S}{\partial p^i}(\overline{p}^i; \rho^i)}. \tag{A.12}$$

Following Proposition 1, $\frac{d\overline{p}^i}{d\rho^i} < 0$. Thus, noting (16) and that $\frac{\partial y(\overline{p}^i, \overline{k}^i)}{\partial k^i}$ > 1, the lemma follows. □

Acknowledgement. The authors would like to thank an anonymous reviewer, Leonard Cheng, Bjarne S. Jensen, Alasdair Smith, Kar-yiu Wong, and participants of the Denmark conference for useful comments.

References

Bardhan, Pranab K. (1970). *Economic Growth, Development, and Foreign Trade.* New York: Wiley-Interscience.

Buiter, Willem H. (1981). "Time Preference and International Lending and Borrowing in an Overlapping-Generations Model." *Journal of Political Economy* 89, 769–797.

Eaton, Jonathan (1987). "A Dynamic Specific-Factors Model of International Trade." *Review of Economic Studies* 54, 325–338.

Findlay, Ronald (1970). "Factor Proportions and Comparative Advantage in the Long Run." *Journal of Political Economy* 78, 27–34.

— (1984). "Growth and Development in Trade Models." In R.W. Jones and P.B. Kenen (eds.), *Handbook of International Economics,* pages 185–236. Amsterdam: North-Holland.

Galor, Oded (1986). "Time Preference and International Labor Migration." *Journal of Economic Theory* 38, 1–20.

— (1992). "A Two-Sector Overlapping-Generations Model: A Global Characterization of the Dynamical System." *Econometrica* 60, 351–386.

— (1996). "Convergence? Inferences from Theoretical Models." *Economic Journal* 106, 1056–1069.

— and Shoukang Lin (1989). "Trade Patterns in a Two Sector Overlapping-Generations Economy." Working Paper No. 89–18, Brown University.

— and Harl E. Ryder (1989). "Existence, Uniqueness and Stability of Equilibrium in an Overlapping-Generations Model with Productive Capital." *Journal of Economic Theory* 49, 360–375.

Jensen, Bjarne S., and Chunyan Wang (1997). "General Equilibrium Dynamics of Basic Trade Models for Growing Economies." (this volume).

Matsuyama, Kiminori (1988). "Life-Cycle Saving and Comparative Advantage in the Long Run." *Economics Letters* 28, 375–379.

Oniki, Hajime, and Hirofumi Uzawa (1965). "Patterns of Trade and Investment in a Dynamic Model of International Trade." *Review of Economic Studies* 32, 15–38.

Shell, Karl (1967). "Optimal Programs of Capital Accumulation for an Economy in which there is Exogenous Technical Change." In K. Shell (ed.), *Essays on the Theory of Optimal Economic Growth.* Cambridge, MA: MIT Press.

Srinivasan, T.N. (1964). "Optimal Savings in a Two-Sector Model of Growth." *Econometrica* 32, 358–373.

Stiglitz, Joseph E. (1970). "Factor Price Equalization in a Dynamic Economy." *Journal of Political Economy* 78, 456–488.

Uzawa, Hirofumi (1964). "Optimal Growth in a Two-Sector Model of Capital Accumulation." *Review of Economic Studies* 31, 1–24.

Part III

Growth and Trade with Endogenous Accumulation of Human or Public Capital

CHAPTER 6

Public and Private Capital in an Endogenously Growing Open Economy

Stephen J. Turnovsky

1. Introduction

The impact of public investment on productive capacity and macroeconomic performance has recently begun to attract the attention of economists. Much of this research was stimulated by Aschauer's (1989a, 1989b) striking findings suggesting that in the United States public capital has a powerful impact on the productivity of private capital. Aschauer's results were controversial and generated both empirical and theoretical research into the role of public investment. While the evidence is mixed, there seems to be a consensus generally supporting the productivity of public investment, although its impact is viewed as being somewhat weaker than that originally suggested by Aschauer.[1]

The theoretical analysis of the productivity of public investment has revolved around analyzing its impact on the growth of private capital and output in the economy. Government expenditure has been introduced as an argument in the production function to reflect an externality in production. In doing this, several strands of literature can be identified. First, there is a substantial literature examining productive government expenditure using Ramsey type models; that is, models that converge either to a stationary state, in which all real variables, including the capital stock, remain constant, or to a growth path along which they grow at some exogenously determined rate. Within this framework, two approaches to incorporating government expenditure can be identified. Most of the existing literature treats the current *flows* of government expenditure as the sources of contributions to productive capacity; see e.g. Aschauer and Greenwood (1985), Aschauer (1988), Barro (1989), Turnovsky and Fisher (1995) and Lee (1995). While the flow specification has the virtue of tractability, it is open to the criticism that insofar as productive government expenditures are intended to represent public infrastructure, such as roads and education, it is the accumulated *stock*, rather than the current flow, that is relevant.

1. A comprehensive review of recent empirical literature is given by Gramlich (1994).

Despite this, within the Ramsey framework, relatively few authors have adopted the alternative approach of modelling productive government expenditure as a stock. Arrow and Kurz (1970) were the first authors to model government expenditure as a form of investment. More recently, Baxter and King (1993) study the macroeconomic implications of increases in the stocks of public goods. They derive the transitional dynamic response of output, investment, consumption, employment, and interest rates to such policies by calibrating a real business cycle model.

The Ramsey model suffers from the drawback that its steady-state growth rate is determined by such factors as the rates of population growth and technological change, and is therefore independent of the usual macroeconomic policy instruments. By contrast, the more recent AK endogenous growth literature has emphasized fiscal policy – and in particular government expenditure policy – as important determinants of long-run growth and growth differentials.[2] Authors such as Barro (1990), Turnovsky (1996a, 1996c) have introduced productive government expenditure, although as a flow. These studies are therefore subject to the shortcomings noted above.

This chapter develops an endogenous growth model of an open economy in which output depends upon the stocks of both private and public capital and which is free to accumulate traded bonds in a perfect world financial market. The dynamic evolution of the economy therefore depends upon the time paths of both capital goods and is thus characterized by transitional dynamics. There are several significant reasons for considering the role of public investment an open economy. The first is that public capital is often a larger component of total capital stock in small economies than in larger economies, less exposed to international trade. Traditionally government investment has played a more significant role in the development of smaller countries like New Zealand and Australia, for example, than in larger economies like the United States.[3] A second reason for choosing such an economy is that it offers the strategic advantage of preserving expositional simplicity. Under the assumption of a perfectly competitive financial market, the accumulation of capital on the one hand, and the growth of consumption on the other, proceed largely independently, enabling us to analyze the dynamic interaction between the two types of capital in a more transparent way. Third, with

2. The effects of government expenditure on growth is emphasized by Barro (1990), Barro and Sala-i-Martin (1992a), Turnovsky (1996a, 1996b). Other authors focus on tax policy; see e.g. Jones and Manuelli (1990), King and Rebelo (1990), Rebelo (1991), Easterly and Rebelo (1993), Jones, Manuelli, and Rossi (1993), Pecorino (1993).

3. Rodrik (1996) has presented empirical evidence supporting the proposition that open economies have larger governments.

the opening of world capital markets, the process of growth in an open economy, and the role played by the government in this process, is obviously of importance in its own right.

Two key features characterize the model and are important determinants of its growth path. The first is that as in standard intertemporal models of small economies, capital accumulation incurs adjustment costs. Second, and less familiar, public capital is subject to congestion. The few existing models that do introduce public capital, treat it as a pure public good, thus failing to take account of the congestion typically associated with public capital. Yet, as Barro and Sala-i-Martin (1995) have argued, virtually all public services are characterized by some degree of congestion, making this a more appropriate assumption. Even national defense, sometimes cited as the purest of public goods, is subject to congestion.[4]

In contrast to the usual specification of congestion in macro growth models, which is typically to normalize aggregate government expenditure by the size of the economy, we allow for a more general parameterization of the degree of congestion, using a form of congestion function from the public goods literature.[5] This is important since the degree of congestion turns out to play a significant role in both determining the effectiveness of public investment on the performance of the economy, as well as in the determination of optimal tax policy. Thus for these various reasons, the explicit consideration of congestion is important in analyzing the role of public investment.

The literature introducing both private and public capital into growth models is sparse. Three recent papers to do so include Futagami, Morita, and Shibata (1993), Glomm and Ravikumar (1994), and Turnovsky (1997).[6] But not only do these papers deal with a closed economy, they abstract from one of the two key aspects being introduced here, namely adjustment costs, treating investment as being residually deter-

4. For example, Thompson (1976) argues that Y and K represent prizes to potential foreign aggressors. If these increase while expenditure remains unchanged, foreigners become more threatening. Accordingly, the government has to raise G in proportion to Y and K if a given state of national security is to be maintained. In this sense national defense is subject to congestion in a similar way as are domestic government services.

5. See e.g. Edwards (1990). More detailed specifications of the microeconomic aspects of congestion are provided by Oakland (1972), Ebrill and Slutsky (1982), and Cornes and Sandler (1986).

6. There is a substantial literature of two-sector endogenous growth models in which the two capital goods are human and nonhuman capital; see e.g. Lucas (1988), Mulligan and Sala-i-Martin (1993), and Pecorino (1993). The present analysis shares some of the characteristics of these models. Clarida and Findlay (1992) present a small international model in which there is only government owned capital.

mined. But adjustment costs in the investment process are important in differentiating between the two types of capital goods, and as shown by Turnovsky (1996b), they also play a fundamental role in determining the nature of the long-run dynamics.[7] Glomm and Ravikumar (1994) and Turnovsky (1997) both emphasize congestion. But private capital in the Glomm-Ravikumar model fully depreciates each period, rather than being subject to at most gradual (or possibly zero) depreciation. This enables the dynamics of the system to be represented by a single state variable alone, so that the system behaves much more like the Barro model in which government expenditure is introduced as a flow. In particular, under constant returns to scale in the reproducible factors, there are no transitional dynamics and the economy is always on a balanced growth path.

The rest of the chapter proceeds as follows. After setting out the analytical framework in section 2, section 3 determines the equilibrium in a centrally planned small open economy, in which the government controls all quantities directly. First-best optimal government expenditure policies are discussed. We begin with the case where the rate of government investment is (arbitrarily) determined as a fixed share of output. The effects of government expenditure on growth, as represented by an increase in this share, are discussed. The optimal rate of government investment is also determined. We show further that the proposition obtained by Barro (1990) in the case where government expenditure impacts on production as a flow – that the growth-maximizing rate of government expenditure coincides with the welfare-maximizing rate – does not extend to the present context.

Section 4 derives a decentralized equilibrium in which the government controls resources only indirectly, through taxation. The effects of various forms of distortionary taxes on the equilibrium are discussed. Section 5 discusses optimal tax policy, in which the decentralized economy attains the first-best equilibrium of the central planner. In order to achieve this, both the steady-state equilibrium and the transitional adjustment path must be replicated. This requires the introduction of a more flexible tax scheme than in the case where the economy is always on its balanced growth path, when, for example, a *fixed* income tax in conjunction together with a *fixed* consumption tax – the latter essentially acting as a lump-sum tax – can replicate the first-best optimum [see e.g. Turnovsky (1996a)]. In the present context the income tax must be *time-varying*, so as to generate the appropriate transitional adjustment

7. Ortigueira and Santos (1997) emphasize the role of adjustment costs in generating plausible speeds of convergence in the two-sector Lucas (1988) model of endogenous growth.

path. An important aspect of the analysis is the characterization of the optimal tax structure, showing its role in correcting for two potential sources of externalities. Section 6 reviews our findings.

2. The Analytical Framework

We consider a small open economy populated by identical representative agents who consume and produce a single traded commodity. Output of this good, y, produced by the typical domestic representative agent is determined by his privately owned capital stock, k, and the services, K_g^s, derived by the firm from its use of public (government) capital stock, in accordance with the constant returns to scale technology:

$$y = F(k, K_g^s) \equiv f\left(\frac{K_g^s}{k}\right) k, \qquad f' > 0, \quad f'' < 0. \tag{1.1}$$

Equation (1.1) embodies the assumption that the services of public capital enhance the productivity of private capital, though at a diminishing rate. The model abstracts from labor so that private capital should be interpreted broadly to include human as well as physical capital; see Rebelo (1991).[8]

The productive services derived by the agent from government capital are represented by

$$K_g^s = K_g(k/K)^{1-\sigma}, \qquad 0 \leq \sigma \leq 1, \tag{1.2}$$

where K_g denotes the aggregate stock of public capital and K denotes the aggregate stock of private stock. Equation (1.2) incorporates the possibility that the public capital may be associated with congestion.[9] The specification in (1.2) characterizes what one can call relative congestion, in that the productive services derived by an individual from a given stock of public capital depends upon his individual capital stock relative to the aggregate.[10] This encourages the use of private capital and is important in the determination of the optimal tax rate.[11] Equation (1.2)

8. It would be straightforward, but tedious, to extend this analysis to include human and nonhuman private capital, as well as public capital.

9. The function (1.2) is the standard specification in the median voter model of congestion; see e.g. Edwards (1990). It implies decreasing marginal congestion provided $\sigma < 1$.

10. A natural alternative specification of congestion is to assume that it is of the absolute form $K_g^s = K_g K^{\sigma-1}$. However, this formulation is in general inconsistent with an equilibrium of ongoing endogenous growth.

11. Previous studies to analyze the effects of congestion on optimal tax policy include Barro and Sala-i-Martin (1992a) and Turnovsky (1996a).

also implies that in order for the level of public capital services, K_g^s, available to the individual firm to remain constant over time, given its individual capital stock, k, the growth rate of K_g must be related to that of K in accordance with $\dot{K}_g/K_g = (1-\sigma)\dot{K}/K$ so that σ parameterizes the degree of (relative) congestion associated with the public good.

The case $\sigma = 1$ corresponds to a non-rival, non-excludable public capital good that is available equally to each firm, independent of the size of the economy; there is no congestion. There are few examples of such pure public goods, so that this case should be viewed largely as a benchmark. At the other extreme, if $\sigma = 0$, then only if K_g increases in direct proportion to the aggregate capital stock, K, does the level of the public service available to the individual firm remain fixed. We shall refer to this case as being one of *proportional* congestion, meaning that the congestion grows in direct proportion to the size of the economy.[12] Road services and infrastructure that play a productive role in facilitating the distribution of the firm's output may serve as examples of public goods subject to this type of congestion. In between, $0 < \sigma < 1$, describes partial congestion, where K_g can increase at a slower rate than does K and still maintain a fixed level of public services to the firm.[13]

The specification of government services by (1.2) implies that the use of public capital is congested only by the use of private capital. Other formulations are also possible. For example, public services might be congested by output or employment. But with labor fixed inelastically, (1.2) is an appropriate specification, especially since our focus is on the interaction of public and private capital accumulation.[14]

Substituting (1.2) into (1.1), the individual firm's production function can be expressed as

$$y = f\left(\frac{K_g}{k}\left[\frac{k}{K}\right]^{1-\sigma}\right)k = f\left(\frac{K_g}{K}\left[\frac{K}{k}\right]^{\sigma}\right)k. \tag{1.1'}$$

As long as $\sigma \neq 1$, so that the public good is associated with some congestion, aggregate capital is introduced into the production function of the individual firm in an analogous way to Romer (1986). With all agents being identical, the relationship aggregate and individual capital stocks are related by $K = Nk$, where N is the number of representative agents. Thus in equilibrium, the individual output y and aggregate

12. In the case $\sigma = 0$ the good is like a private good in that the median voter receives his proportionate share.

13. The case $\sigma < 0$ can be interpreted as describing an extreme situation where the congestion of the public good is faster than the growth of the economy. While we do not discuss it, one can easily interpret our results in that case.

14. See Glomm and Ravikumar (1994) for alternative formulations of congestion.

output $Y = Ny$ may be expressed as

$$y = f\left(\frac{K_g}{K}N^{\sigma}\right)k, \quad Y = f\left(\frac{K_g}{K}N^{\sigma}\right)K. \tag{1.1''}$$

The critical difference between the perception of the world as seen by the representative firm and as seen by the central planner is as follows. The representative firm treats the aggregate capital stock K as given, with the relationship $K = Nk$, as employed in (1.1″) holding as an equilibrium one. The central planner, on the other hand, takes this relationship into account when determining his decisions. For expositional convenience we shall set the number of agents $N = 1$, enabling us to drop the distinction between aggregate and individual quantities in equilibrium. While for our purposes this normalization suffices, it is not entirely innocuous either, since the effects of congestion (as parameterized by σ) do depend upon the number of agents.[15]

The agent consumes this good at the rate C, yielding intertemporal utility over an infinite time horizon represented by the intertemporal isoelastic utility function:

$$\Omega \equiv \int_0^{\infty} \frac{1}{\gamma} C^{\gamma} e^{-\rho t} dt, \qquad -\infty < \gamma < 1. \tag{2}$$

Private capital, K, depreciates at the constant rate δ_k, so that letting I denote the rate of gross private investment, net private capital accumulates at the rate

$$\dot{K} = I - \delta_k K. \tag{3.1}$$

Likewise, public capital, K_g, depreciates at the constant rate δ_g, so that letting G denote the rate of gross public investment, the rate of net public capital accumulation follows

$$\dot{K}_g = G - \delta_g K_g. \tag{3.2}$$

New output may be transformed to either type of capital. In either case this process involves adjustment costs (installation costs) that we incorporate in the quadratic (convex) functions[16]

$$\Omega(I, K) \equiv I\left(1 + \frac{h_1}{2}\frac{I}{K}\right), \tag{4.1}$$

15. The dependence of the growth rate upon the population size is emphasized by Glomm and Ravikumar (1994). The economy is thus one in which growth is subject to "scale effects", the empirical relevance of which has recently been questioned; see Jones (1995). The normalization $N = 1$ can easily be relaxed if one wishes.

16. We shall that the depreciation rate is sufficiently large to ensure that gross investment is always positive.

$$\Psi(G, K_g) \equiv G\left(1 + \frac{h_2}{2}\frac{G}{K_g}\right). \tag{4.2}$$

Equations (4.1)–(4.2) are applications of the familiar Hayashi (1982) cost of adjustment framework, where we assume that for both types of capital the adjustment costs are proportional to the *rate* of investment per unit of installed capital, rather than to its absolute level. The linear homogeneity of this function is necessary if a steady-state equilibrium having ongoing growth is to be sustained.

In addition, the economy also accumulates net foreign bonds, b, that pay an exogenously given world interest rate, r. Thus, the accumulation and consumption decisions facing the economy are constrained by the economy-wide resource constraint:

$$\dot{b} = rb + f\left(\frac{K_g}{K}\right)K - C - I\left(1 + \frac{h_1}{2}\frac{I}{K}\right) - G\left(1 + \frac{h_2}{2}\frac{G}{K_g}\right). \tag{5}$$

This equation asserts that current account balance for the small open economy, given by the right hand side of (5), consists of domestic output less the output used up in consumption and in the accumulation and installation of the two types of capital, plus the interest earned (or owed) on its holdings of traded bonds.

In order for an equilibrium with steady ongoing growth to be sustained, the current flow of government expenditure, G, must be linked to the size of the economy. While there are several ways this might be accomplished, a natural case to consider, if one wishes to parameterize expenditure policy explicitly, is to specify[17]

$$G = gf\left(\frac{K_g}{K}\right)K. \tag{6}$$

As long as g remains fixed, the government is claiming a fixed share of the growing output for gross investment, so that an increase in the share, g, parameterizes an expansionary expenditure policy in a growing economy.[18] In section 3.4 below we also discuss the case where government expenditure is set optimally along with private expenditures. As will be seen, the optimal expenditure policy will require the fraction g to be *time-varying*, continuously adapting to the changing aggregate

17. Other rules determining government expenditure are also possible. For example, (6′) below postulates expenditure to be related to total GNP, rather than to current output.

18. Barro (1990) and Rebelo (1991) in effect parameterize government expenditure in this fashion by assuming that all income tax revenues are spent; i.e. $G = \tau Y$.

stocks of public and private capital. This optimum serves as an important benchmark in explaining the effects of changes in government expenditure away from the optimum.

3. Equilibrium in the Centrally Planned Economy

Taking g to be an arbitrarily set fraction, the central planner's problem is to choose the rate of consumption, C, the rates of investment, I and G, and the rates of asset accumulation, $\dot{b}$, $\dot{K}$, and $\dot{K}_g$ to maximize (2), subject to (3.1)–(3.2) and (5)–(6). With the simplifying assumption $N = 1$, the present value Hamiltonian for this optimization is given by

$$\begin{aligned} H \equiv\ & \frac{1}{\gamma}C^{\gamma}e^{-\rho t} + \eta e^{-\rho t} \\ & \cdot \left[f\left(\frac{K_g}{K}\right)K + rb - C - I\left(1 + \frac{h_1}{2}\frac{I}{K}\right) - G\left(1 + \frac{h_2}{2}\frac{G}{K_g}\right) - \dot{b}\right] \\ & + q_1' e^{-\rho t}[I - \delta_k K - \dot{K}] + q_2' e^{-\rho t}[G - \delta_g K_g - \dot{K}_g] \\ & + \nu' e^{-\rho t}\left[gf\left(\frac{K_g}{K}\right)K - G\right], \end{aligned} \tag{7}$$

where η is the shadow value (marginal utility) of wealth in the form of internationally traded bonds (or new output); q_1', q_2' are the shadow values of the private and public capital stocks; ν' is the shadow value of devoting a marginal unit of output to the government. Analysis of the model is simplified by using the shadow value of wealth as numeraire. Consequently $q_1 \equiv q_1'/\eta$ is the (market) value of private capital, $q_2 \equiv q_2'/\eta$ is the imputed value of public capital, and $\nu \equiv \nu'/\eta$ is the shadow value of allocating a marginal unit of output to the government, all measured in terms of the (unitary) price of foreign bonds. As we will see presently, in equilibrium $\nu \gtrless 0$, depending upon the size of government expenditure relative to the optimum.

The optimality conditions with respect to C, I, and G are respectively

$$C^{\gamma-1} = \eta, \tag{8.1}$$

$$\left(1 + h_1\frac{I}{K}\right) = q_1, \tag{8.2}$$

$$\left(1 + h_2\frac{G}{K_g}\right) + \nu = q_2. \tag{8.3}$$

Equation (8.1) equates the marginal utility of consumption to the shadow value of wealth. Equation (8.2) equates the marginal cost of an extra unit of private investment, inclusive of its installation cost, $h_1 I/K$ to the shadow value of private capital, q_1. Likewise, (8.3) equates the marginal benefits of an extra unit of public investment to the shadow value of the public capital stock, q_2. As well as the resource cost, measured by the term in parentheses in (8.3), the marginal benefits also include the shadow value of having a larger public sector, measured by ν. Equations (8.2)–(8.3) may be immediately solved to yield the following expressions for the rates of capital accumulation (ϕ_k, ϕ_g):

$$\tfrac{I}{K} = \tfrac{q_1-1}{h_1}, \qquad \tfrac{\dot{K}}{K} = \tfrac{q_1-1}{h_1} - \delta_k \quad \equiv \phi_k, \tag{9.1}$$

$$\tfrac{G}{K_g} = \tfrac{q_2-\nu-1}{h_2} = g\tfrac{f(K_g/K)K}{K_g}, \quad \tfrac{\dot{K}_g}{K_g} = \tfrac{q_2-\nu-1}{h_2} - \delta_g \equiv \phi_g. \tag{9.2}$$

The optimality condition with respect to traded bonds is given by the arbitrage condition:

$$\rho - \frac{\dot{\eta}}{\eta} = r. \tag{10.1}$$

Equation (10.1) is the standard Keynes-Ramsey consumption rule, equating the marginal return on consumption to the rate of return on holding a foreign bond. With ρ and r both constant, it implies that the marginal utility η grows at the constant rate $(\rho - r)$. Taking the time derivative of (8.1) and combining with (10.1), we see that in equilibrium consumption grows at the constant rate

$$\frac{\dot{C}}{C} = \frac{r-\rho}{1-\gamma} \equiv \psi. \tag{11}$$

The equilibrium consumption growth rate of a small open economy facing perfect financial markets depends upon the given world interest rate and preference parameters; it is independent of domestic production conditions.[19] The level of consumption at time t is

$$C(t) = C(0)e^{\psi t}, \tag{12}$$

where the initial level of consumption $C(0)$ is yet to be determined.

The optimality conditions with respect to the two types of capital, K, and K_g, are described by the arbitrage conditions

$$(1+\nu g)\frac{[f(z) - f'(z)z]}{q_1} + \frac{\dot{q}_1}{q_1} + \frac{(q_1-1)^2}{2h_1 q_1} - \delta_k = r, \tag{10.2}$$

19. This is the endogenous growth analogue to "consumption smoothing." In Ramsey-type models of small open economies, the restriction $\rho = r$ is typically imposed in order to ensure that the economy converges to a stationary state. This would imply that consumption remains constant over time.

$$(1+\nu g)\frac{f'(z)}{q_2}+\frac{\dot{q}_2}{q_2}+\frac{(q_2+\nu-1)^2}{2h_2q_2}-\delta_g=r, \tag{10.3}$$

where $z \equiv K_g/K$ denotes the ratio of public to private capital. Thus, equation (10.2) equates the social rate of return on private capital, net of physical depreciation to the given rate of return on the traded bond. The former consists of the following components. First, is the net marginal *social* product of capital per unit of installed capital, valued at the price q_1. This measure incorporates the fact that when the government ties its expenditures to output as in (6), an increase in private capital, by increasing output, will also induce an increase in the size of the government, the social contribution of which is valued at the shadow price ν. The second component is the rate of capital gain. The third element, which is less familiar, is equal to $(q_1 I - \Omega)/q_1 K$. This component reflects the fact that an additional source of benefits of higher capital stock is to reduce the installation costs (which depend upon I/K) associated with new investment. The interpretation of (10.3) is analogous.

In order to ensure that the intertemporal resource constraint is met, the following transversality conditions must hold:

$$\lim_{t\to\infty}\eta b e^{-\rho t}=0,\quad \lim_{t\to\infty}q_1' K e^{-\rho t}=0,\quad \lim_{t\to\infty}q_2' K_g e^{-\rho t}=0. \tag{13}$$

3.1 Equilibrium Dynamics: Private and Public Capital

A consequence of the perfect world capital market is that the equilibrium dynamics of the economy dichotomize. Equations (9.1)–(9.2) and (10.2)–(10.3) determine the evolution of the two types of capital stocks. Having determined these, (12) in conjunction with the aggregate resource constraint and the transversality condition on traded bonds, determines the evolution of traded bonds and the current account, consistent with the intertemporal solvency of the economy.

The equilibrium describing the accumulation of the two types of capital is represented by the following system

$$\frac{q_2-\nu-1}{h_2}=g\frac{f(z)}{z}, \tag{14.1}$$

$$\dot{q}_1=(r+\delta_k)q_1-\frac{(q_1-1)^2}{2h_1}-(1+\nu g)[f(z)-f'(z)z], \tag{14.2}$$

$$\dot{q}_2=(r+\delta_g)q_2-\frac{(q_2-\nu-1)^2}{2h_2}-(1+\nu g)f'(z), \tag{14.3}$$

$$\frac{\dot{z}}{z}=\left[\frac{q_2-\nu-1}{h_2}-\delta_g\right]-\left[\frac{q_1-1}{h_1}-\delta_k\right]\quad(\equiv \phi_g-\phi_k). \tag{14.4}$$

Equations (14.1)–(14.3) repeat (9.2) and (10.2)–(10.3) respectively, noting the definition of the ratio of the two types of capital stock. Equation (14.1) can be viewed as determining the shadow value of investment in terms of q_2 and z; i.e. determining $\nu = \nu(q_2, z)$. The evolution of the shadow values is described by (14.2)–(14.3). The fourth equation determines the proportionate rate of change of the ratio of capital stocks and is obtained by substituting (9.1)–(9.2) into the relationship $\dot{z}/z = \dot{K}_g/K_g - \dot{K}/K$.

The critical determinants of the growth rate of private capital include the market price of installed capital, q_1,and the relative stock of capital, z, the paths of which are determined by (14.2) and (14.4). The short-run dynamics will be discussed in section 3.5. In order for the capital stocks K and K_g ultimately to follow paths of steady growth, the stationary solution to this system, attained when $\dot{q}_1 = \dot{q}_2 = \dot{k} = 0$, must have at least one *real* solution.

The costs of adjustment associated with the accumulation of both types of capital introduce nonlinearities into the dynamic system (14.1)–(14.4), leading to potential existence and nonuniqueness problems of equilibrium. This issue has been discussed in a simplified version of this small open economy model, having only private capital.[20] The intuition of the argument, which applies here as well, is simply that with adjustment costs, the returns to capital due to valuation differences between installed capital and the resources they embody [the third element in (10.2)] may be sufficiently large so as to cause the overall returns to capital to dominate sufficiently the returns to traded bonds, so that irrespective of the price of capital no long-run balanced growth equilibrium can exist in which the returns to the two assets are brought into equality. Further technical consideration of the existence problem are provided in the Appendix. Henceforth, our discussion proceeds under the assumption that a steady-state equilibrium does indeed exist.

Since the dynamic system (14.1)–(14.4) is nonlinear, we proceed by considering the linearized dynamics of the two types of capital about steady state. In describing these dynamics we first substitute $\nu = \nu(q_2, z)$ into (14.2)–(14.4), yielding an autonomous system in q_1, q_2, and z. This system is then linearized around steady state, while imposing the appropriate transversality conditions. Performing the linearization, the dynamics can be represented by the system (where tildes denote steady states)

20. See Turnovsky (1996b).

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{z} \end{bmatrix} = A \begin{bmatrix} q_1 - \tilde{q}_1 \\ q_2 - \tilde{q}_2 \\ z - \tilde{z} \end{bmatrix}, \tag{15}$$

with

$$A = \begin{bmatrix} (r+\delta_k) - \frac{\tilde{q}_1 - 1}{h_1} & -g(f - \tilde{z}f') & (1+\tilde{\nu}g)f''\tilde{z} - \frac{g^2 h_2}{\tilde{z}^2}(f - \tilde{z}f')^2 \\ 0 & (r+\delta_g) - gf' & -(1+\tilde{\nu}g)f'' + \frac{g^2 h_2}{\tilde{z}^3}(f - \tilde{z}f')^2 \\ -\frac{\tilde{z}}{h_1} & 0 & -\frac{g}{\tilde{z}}(f - \tilde{z}f') \end{bmatrix}.$$

The second and third of the transversality conditions (13) are relevant for determining the dynamics of the capital stocks. We focus on the second condition and show that this condition will be met if and only if[21]

$$r + \delta_k > \frac{\tilde{q}_1 - 1}{h_1}. \tag{16.1}$$

Likewise, the third condition will be met if and only if

$$r + \delta_g > \frac{\tilde{q}_2 - \tilde{\nu} - 1}{h_2} = g\frac{f}{\tilde{z}} > gf'. \tag{16.2}$$

These conditions assert that the transversality condition will hold if and only if the respective net growth rate of capital is less than the given world rate of interest; $r > \tilde{\phi}_i$, $i = k, g$. In addition, we impose the condition $1 + \tilde{\nu}g > 0$; that is, the marginal physical product of either form of capital, inclusive of the induced effect through its impact on the induced size of government, is positive.[22]

Under these conditions, the determinant of the matrix in (15), $A < 0$, implying that the linearized system has two unstable roots, and one stable root, $\lambda < 0$. We assume that the two shadow values q_1 and q_2 can respond instantaneously to new information, while since both types of

21. This condition can be established as follows. Using the definition of $q_1' \equiv q_1\eta$, equation (9.1) and (10.1), we have $\lim_{t\to\infty} q_1'(t)K(t)\exp\{-\rho t\} = \lim_{t\to\infty} q_1(t)\eta(0)K_0 \exp\{\int_0^t \phi_k(s)ds - rt\}$. The transversality condition will be met as long as $r > \bar{\phi}_k$; that is if $r + \delta_k > (\tilde{q}_1 - 1)/h_1$. Now consider the steady state to (14.2), namely $\tilde{q}_1^2 - 2[1 + h_1(r+\delta_k)] + 1 + 2h_1(1+\tilde{v}g)[f(\tilde{z}) - f'(\tilde{z})\tilde{z}] = 0$. Treating this as a quadratic equation in $\tilde{q}_1$, the transversality condition holds if one takes the negative root to this equation, thus ruling out the positive root; see also Turnovsky (1996b). The same observation applies to (16.2).

22. If $\tilde{v} < 0$ so that the government exceeds its optimal size, the restriction $1 + \tilde{v}g > 0$ imposes an upper limit on the size of the government.

capital involve adjustment costs and are therefore constrained to evolve gradually over time, their ratio, z, is also restricted to continuous adjustments. The linearized system is therefore a saddlepoint. Starting from an initially given ratio, z_0, in the neighborhood of the equilibrium steady-state growth path, the stable adjustment path of the linearized system evolves as follows[23]:

$$z(t) - \tilde{z} = (z_0 - \tilde{z})e^{\lambda t}, \tag{17.1}$$

$$q_1(t) - \tilde{q}_1 = \frac{\tilde{z}\zeta}{[(r+\delta_g) - gf'] - \lambda}(z - \tilde{z}), \tag{17.2}$$

$$q_2(t) - \tilde{q}_2 = \frac{\zeta}{\lambda - [(r+\delta_g) - gf']}(z - \tilde{z}), \tag{17.3}$$

where $\zeta \equiv -(1+\tilde{\nu}g)f'' + \frac{g^2 h_2}{\tilde{z}^3}[f - \tilde{z}f']^2 > 0$. Corresponding to the monotonic adjustment of the relative capital stocks described in (17.1), the positive relationship (17.2) and the negative relationship (17.3) describe the respective stable adjustments in the shadow values of private and public capital. The signs of these relationships reflect the fact that as the ratio of public to private capital increases (i.e. the relative scarcity of private to public capital increases) the shadow value of private capital rises, while that of public capital falls.

It is straightforward and more to the point of our discussion to express the transitional dynamics in terms of growth rates, rather than shadow values. In steady-state equilibrium the ratio of public to private capital remains constant, so that both types of capital grow asymptotically at the same rate. Denoting this common rate by $\tilde{\phi}_k = \tilde{\phi}_g \equiv \tilde{\phi}$, the linearized transitional paths followed by the respective growth rates are[24]:

$$\phi_k - \tilde{\phi} = \frac{\tilde{z}\zeta/h_1}{[(r+\delta_g) - gf'] - \lambda}(z - \tilde{z}), \tag{18.1}$$

$$\phi_g - \tilde{\phi} = -\frac{g}{\tilde{z}^2}[f - \tilde{z}f'](z - \tilde{z}). \tag{18.2}$$

These are illustrated in figure 1, where the positively sloped locus XX corresponds to the stable transitional adjustment path in the growth rate of private capital and the negatively sloped locus YY corresponds to the stable adjustment in the growth rate of public capital. The striking feature of the adjustment is that during any transition the growth

23. In deriving (17.2), we have made use of the condition that in steady-state equilibrium $(\tilde{q}_1 - 1)/h_1 - \delta_k = (\tilde{q}_2 - \tilde{v} - 1)/h_2 - \delta_g$; see (20.4) below.

24. The solution for the transitional path for the growth rate ϕ_g, for constant g, is obtained by linearizing the relationship $\phi_g = gf(z)/z$.

rates of the two forms of capital are moving in opposite directions. This figure forms the basis for the analysis of the dynamic effects of a fiscal expansion.

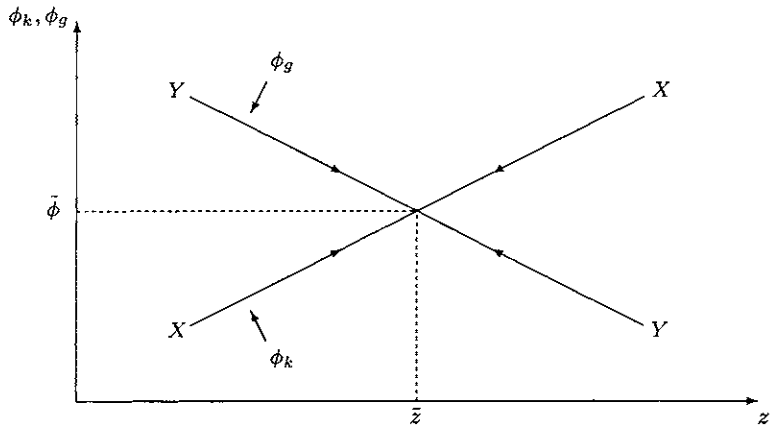


Fig. 1. Stable adjustment paths for growth rates of public and private capital

3.2 Equilibrium Dynamics: Current Account

To obtain the time path for the current account we proceed as follows. First, linearize the production function in (5). Next, substitute the solutions for $K(t)$, $K_g(t)$, and $C(t)$ from (9.1)–(9.2) and (12) into (5). This leads to the following linear approximation describing the rate of accumulation of traded bonds:

$$\dot{b} = rb + \Gamma_0 e^{\int_0^t \phi_k(s)ds} + \Gamma_{g,0} e^{\int_0^t \phi_g(s)ds} - C(0)e^{\psi t},$$

where $\Gamma_0 \equiv f(K_{g,0}/K_0)K_0 - \Omega_0$, $\Gamma_{g,0} = -\Psi_0$, reflect the initial impacts of the private and public capital stocks on the economy's net output. Solving this equation and invoking the transversality condition on the traded bond, we can show that

$$C(0) = (r - \psi)\{b_0 + \Gamma_0 \int_0^\infty e^{\int_0^s \phi_k(\tau)d\tau - rs} ds + \Gamma_{g,0} \int_0^\infty e^{\int_0^s \phi_g(\tau)d\tau - rs} ds\}. \qquad (19.1)$$

This equation determines the initial level of consumption $C(0)$, consistent with the intertemporal solvency of the small open economy. The term in parentheses can be interpreted as the present discounted value of wealth, allowing for the fact that both types of capital grow at the respective net rates indicated by (9.1)–(9.2). Substituting this initial condition into the general solution for $b(t)$, we find that the stock of traded bonds follows the growth path

$$b(t) = -e^{rt}\left[\Gamma_0 \int_t^\infty e^{\int_0^s \phi_k(\tau)d\tau - rs} ds \right. \\ \left. + \Gamma_{g,0}\int_t^\infty e^{\int_0^s \phi_g(\tau)d\tau - rs} ds - \frac{C(0)}{r-\psi} e^{(\psi - r)t}\right]. \qquad (19.2)$$

Holdings of traded bonds are subject to transitional dynamics, in the sense that their growth rate $\dot{b}/b$ would vary through time, even if the growth rates of capital ϕ_k, ϕ_g were constant. Asymptotically the growth rate converges to $\max[\psi, \tilde{\phi}]$ and which it will be depends critically upon the size of the consumer rate of time preference relative to the rates of return on investment opportunities. For example, if domestic agents are sufficiently patient (i.e. ρ is sufficiently small) one can show that $\psi > \tilde{\phi}$. In the long run domestic consumption will grow at a faster rate than does either form of domestic capital or domestic output. By being patient, the agents choose to consume a small fraction of their wealth. This enables them to accumulate foreign assets, running up a current account surplus and generating a positively growing stock of foreign assets. It is the income from these assets that enables the small economy to sustain a long-run growth rate of consumption in excess of the growth rate of domestic productive capacity. The opposite applies if $\psi < \tilde{\phi}$. In the long run, the country accumulates an ever increasing foreign *debt* and is unable to maintain a consumption growth rate equal to that of domestic output.[25]

25. Thus a feature of this equilibrium is that it sustains differential growth rates of consumption and domestic output This is a consequence of the economy being small and open. It is in contrast to a closed economy in which, constrained by the growth of its own resources, all real variables, including consumption and output, ultimately have to grow at the same rate. Thus in the small open economy, the consumption-domestic output ratio will either tend to zero or infinity. However, this is not of particular concern insofar as a sustainable equilibrium is concerned, since consumption is determined by wealth as in the right hand side of (19.1).

We shall assume that the country is sufficiently small so that it can maintain a growth rate which is unrelated to that in the rest of the world. Ultimately, this requirement imposes a constraint on the growth rate of the economy. If it grows faster than the rest of the world, at some point it will cease to be small. While we do not attempt to resolve this issue here, we should note that the issue of convergence in

3.3 Optimal Government Expenditure

So far, the equilibrium has been derived on the assumption that the government claims an *arbitrary* share of output g. The *optimal* share of government expenditure is determined by setting $\partial H/\partial g = 0$ in (7). This leads to setting $\nu = 0$ in the equilibrium (14.1)–(14.4). The dynamic equations thus determine an autonomous system in the three dynamic variables q_1, q_2, and z, having the same saddlepath property as before. Now the corresponding optimal share of government expenditure, $\hat{g}(t)$, is determined from

$$\hat{g}(t)\frac{f(z)}{z} = \frac{q_2 - 1}{h_2}. \tag{14.1$'$}$$

A key feature of the optimal policy, $\hat{g}(t)$, is that it is *time-varying.* Differentiating (14.1′) with respect to t, the optimal rate of change, $\dot{\hat{g}}(t)$, may be conveniently expressed in the form

$$\dot{\hat{g}}(t) = \frac{1}{f(z)/z}\left[\frac{\dot{q}_2}{h_2} - \hat{g}(t)\frac{\dot{f}/z}{f/z}\right]. \tag{14.1$''$}$$

From this equation the optimal fraction of output claimed by the government is subject to two offsetting influences. First, as the ratio of public to private capital, z, increases over time, the average productivity of public capital, f/z, declines, causing $\hat{g}$ to rise. At the same time, the increase in z reduces the shadow value of public capital, q_2, reducing the incentive for the government to further invest in public capital.

As $z \to \tilde{z}$ and $q_2 \to \tilde{q}_2$, (14.1′) implies that $\hat{g}(t) \to \tilde{\hat{g}}$. One can further establish that at the steady-state $(\partial\tilde{\nu}/\partial g)_{\nu=0} < 0$. That is, in the neighborhood of the steady-state optimum, the shadow value of increasing the size of government is positive, if the size of the government is less than the long-run optimum, and it is negative if the size of the government exceeds the optimum; i.e. $\tilde{\nu} \gtrless 0$ according as $g \lessgtr \tilde{\hat{g}}$.

3.4 Long-Run Effects of Fiscal Expansion

Because of the forward-looking nature of the shadow values, the transitional dynamics are determined in part by the steady-state equilibrium. It is therefore convenient to begin with the latter, which is characterized by:

$$\frac{\tilde{q}_2 - \tilde{\nu} - 1}{h_2} = g\frac{f(\tilde{z})}{\tilde{z}}, \tag{20.1}$$

international growth rates is receiving attention in the literature; see e.g. Barro and Sala-i-Martin (1992b).

$$(1+\tilde{\nu}g)[f(\tilde{z})-f'(\tilde{z})\tilde{z}]+\frac{(\tilde{q}_1-1)^2}{2h_1}-(r+\delta_k)\tilde{q}_1=0, \quad (20.2)$$

$$(1+\tilde{\nu}g)f'(\tilde{z})+\frac{(\tilde{q}_2-\tilde{\nu}-1)^2}{2h_2}-(r+\delta_g)\tilde{q}_2=0, \quad (20.3)$$

$$\left[\frac{\tilde{q}_2-\tilde{\nu}-1}{h_2}-\delta_g\right]=\left[\frac{\tilde{q}_1-1}{h_1}-\delta_k\right]\equiv\tilde{\phi}. \quad (20.4)$$

These relations determine the steady-state shadow values, $\tilde{q}_1$, $\tilde{q}_2$, and $\tilde{\nu}$; the ratio of the two capital stocks, $\tilde{z}$; and the common long-run growth rate $\tilde{\phi}$. We shall focus on the latter two effects. The equilibrium growth rate may be either positive or negative. But if it is positive, the transversality conditions (16.1)–(16.2) impose the restriction that it cannot exceed the rate of return on traded bonds; i.e. $r > \tilde{\phi}$.

Differentiating (20.1)–(20.4) with respect to g yields:

$$\frac{d\tilde{\phi}}{dg}=\frac{1}{\tau h_1}\left((r-\tilde{\phi})(1+\tilde{\nu}g)ff''\tilde{z}-(r+\delta_g)\frac{\tilde{\nu}g}{\tilde{z}}[f-\tilde{z}f']^2\right), \quad (21.1)$$

$$\begin{aligned}\frac{d\tilde{k}}{dg}&=-\frac{1}{\tau}\Big(f[(r+\delta_g)-gf'](r-\tilde{\phi})\\&\quad+\frac{\tilde{z}}{h_1}[f-\tilde{z}f']\left[\frac{f}{\tilde{z}}gh_2(r-\tilde{\phi})-(r+\delta_g)\tilde{\nu}\right]\Big). \quad (21.2)\end{aligned}$$

where $\tau < 0$, is the Jacobian of (20.1)–(20.4). Equation (21.1) indicates that an increase in g has two effects on the common steady-state growth rate of the two types of capital. First, to the extent that the net social marginal physical product of capital, taking into account the value of the induced effect through the size of government, is positive, $(1+\tilde{\nu}g > 0)$ it is *growth-enhancing*. This effect is described by the first term in parentheses. If, in addition, $\tilde{\nu} > 0$, so that the initial amount of output devoted to the government is below the optimum, then increasing g toward the optimum will further enhance the growth rate. However, if $\tilde{\nu} < 0$, so that initially too much current output is being absorbed by the government, then this second effect will be *growth-reducing*. In this case, the net result of an increase in government expenditure depends upon which effect dominates and there is an optimal steady-state growth maximizing rate of government expenditure at which these two effects are precisely balanced.

Since $(d\tilde{\phi}/dg)_{\nu=0} > 0$, the long-run growth-maximizing level of g, $\bar{g}$ say, *exceeds* the long-run welfare maximizing level, $\hat{g}$. This is in contrast to Barro (1990), who, introducing government expenditure as a flow in the production function, finds that the welfare-maximizing and growth-maximizing shares of government expenditure coincide. The difference is accounted for by the fact that when government expenditure influences production as a flow, maximizing the marginal product of government expenditure net of its resource cost maximizes the growth rate of capital. But it also maximizes the social return to public expenditure, thereby maximizing overall intertemporal welfare. By contrast, when government expenditure affects output as a stock, public capital needs to be accumulated to attain the growth maximizing level. This involves foregoing consumption, leading to welfare losses relative to the social optimum. Intertemporal welfare is raised by reducing the growth rate, thereby enabling the agent to enjoy more consumption.[26]

Equation (21.2) ensures that if $\tilde{\nu} \leq 0$, so that the economy does not have a shortage of public capital, then if the central planner increases the share of output devoted to public capital, the long-run ratio of public to private capital is increased. However, if $\tilde{\nu} > 0$, one cannot rule out the possibility that the short-run growth in private capital generated by the increase in g in that case will be sufficiently great so as to reduce the steady-state ratio of public to private capital.[27]

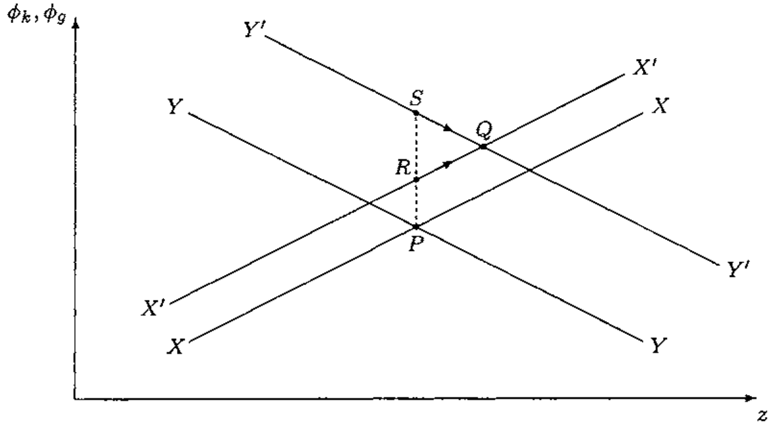
3.5 Transitional Dynamics

Figure 2 illustrates alternative transitional paths for growth rates of the two types of capital following an unanticipated permanent increase in g. We assume the more plausible case where $d\tilde{z}/dg > 0$. In this figure the points P and Q represent initial and final steady-state equilibria.

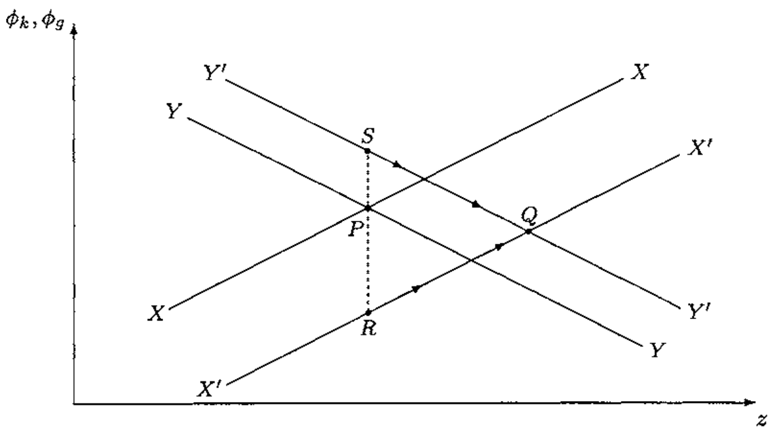
The immediate effect of a larger share of output being devoted to public investment is to raise the initial growth rate of public capital, $\phi_g(0)$, doing so by an amount $(f/z)dg$; see (9.2). The implied long-run increase in the ratio of public to private capital means that during the transition it is always increasing; i.e. $\dot{z} > 0$. As a result, the average productivity of public capital f/z declines over time, so that with g remaining unchanged after the initial increase, the growth rate of public capital, ϕ_g, declines over time. The time path of ϕ_g is represented by a jump from the initial equilibrium P to S, followed by a continuous

26. In Turnovsky (1996c) where we introduce productive government expenditure as a flow, we assume that it not only improves the productivity of existing capital, but also that it reduces the cost of adjustment associated with investment. This latter aspect also leads to the result that the growth-maximizing rate of government expenditure exceeds the welfare-maximizing rate.

27. Setting $\tilde{\nu} = 0$ in (20.1)–(20.4) determines the equilibrium long-run values for $\hat{q}_1$, $\hat{q}_2$, and $\hat{z}$, corresponding to the endogenously determined optimal share of government expenditure $\hat{g}$. It is straightforward to analyze these equations to determine the impact of various shocks on the optimal long-run share of government expenditure.



A. Increase in Long-run Growth Rate



B. Decrease in Long-run Growth Rate

Fig. 2. Increase in government expenditure: Centrally planned economy

decline along the path SQ. The new steady state, Q, may lie above the original steady state P as in figure 2A, or below as in figure 2B, depending upon whether $d\tilde{\phi}/dg \gtrless 0$.

The initial response of the growth rate of private capital, $\phi_k(0)$, is ambiguous, and reflects two, possibly offsetting, effects. From (18.1), we have:

$$d\phi_k(0) = d\tilde{\phi} - \frac{z\zeta/h_1}{[(r+\delta_g) - gf'] - \lambda} d\tilde{z}.$$

The partial effect of the long-run increase in the ratio of public to private capital, $\tilde{z}$, is to reduce the initial growth rate of private capital. This is because, as z increases during the subsequent transition, the relative scarcity of private capital, and therefore its shadow value, increases so that private investment is stimulated. In order to accommodate this while holding $\tilde{\phi}$ constant, in the short run the growth of private investment must decline. Offsetting this is the fact that to the extent that the long-run growth rate of capital may be expected to increase ($d\tilde{\phi} > 0$), that will induce an immediate increase in the growth rate as well. Overall, whether $\phi_k(0)$ rises or falls depends upon which effect dominates.

Figure 2A illustrates the case where $\phi_k(0)$ initially increases. This causes the stable locus XX to shift up to $X'X'$; $\phi_k(0)$ initially jumps up from P to R and then continues to increase along RQ to the new equilibrium. Figure 2B illustrates the case where $\phi_k(0)$ initially drops from P to R and then increases gradually to the new equilibrium at Q. As illustrated, Q lies below P in figure 2B, although that need not necessarily be the case. In either case the transitional dynamics following the initial jump causes the two growth rates to approach their common equilibrium from opposite directions. This is because the declining ratio of public to private capital, z, during the transition is associated with the decreasing productivity of public capital and the increasing productivity of private capital.

4. Decentralized Economy

We now turn to the representative agent operating in a decentralized economy. The objective of the agent is to maximize his constant elasticity utility function (2), subject to his accumulation of private capital (3.1) and his own budget constraint, represented by

$$\dot{b} = r(1-\tau_b)b + (1-\tau_k)f\left(\frac{K_g}{K}\left(\frac{K}{k}\right)^{\sigma}\right)k,$$

$$-C(1+\omega) - I\left(1+\frac{h_1}{2}\frac{I}{K}\right) - T, \tag{5'}$$

where τ_k is the rate of taxation on capital income; τ_b is the rate of taxation on foreign bond income; ω is the rate of taxation on consumption; T is the time-varying rate of lump-sum taxation (or rebate).

Two points concerning this specification merit comment. First, throughout this section we assume that the distortionary tax rates τ_b, τ_k, and ω are constant through time, being subject to at most once-and for-all policy changes at discrete times. As we will show in section 5 below, to replicate the first best optimum, τ_k will need to be time-varying. Second, in performing this optimization, the agent is assumed to treat the stock of public capital, K_g, and the aggregate stock of private capital, K, as given and independent of his own decisions. With the population size being normalized at unity, the condition $k = K$ holds as an equilibrium relationship.

In the absence of government bonds, the government must maintain a continuously balanced budget which, for the above specification of taxation and with G specified in accordance with (6), is:

$$T+\tau_k f(K_g/K)K+r\tau_b b+\omega C = gf(K_g/K)K\left[1+\frac{h_2}{2}\frac{gf(K_g/K)}{K_g/K}\right]. \tag{22}$$

Note that combining (22) with (5′) yields the national resource constraint (5).

4.1 Equilibrium Growth

The representative agent's optimality conditions with respect to private consumption and private investment are:

$$(C^*)^{\gamma-1} = \eta^*(1+\omega), \tag{8.1'}$$

$$\left[1+h_1\left(\frac{I}{K}\right)^*\right] = q^*, \tag{8.2'}$$

where star denotes equilibrium in the decentralized economy and q^* is the market value of private capital. Thus, analogous to (9.1), we have

$$\left(\frac{I}{K}\right)^* = \frac{q^*-1}{h_1}, \quad \left(\frac{\dot{K}}{K}\right)^* = \frac{q^*-1}{h_1} - \delta_k \equiv \phi_k^*. \tag{9.1'}$$

The arbitrage condition with respect to traded bonds is now

$$\rho - \frac{\dot{\eta}^*}{\eta^*} = r(1-\tau_b), \tag{10.1'}$$

so that the equilibrium rate of growth of consumption becomes

$$\left(\frac{\dot{C}}{C}\right)^{*} = \frac{r(1-\tau_b)-\rho}{1-\gamma} \equiv \psi^{*}, \tag{11'}$$

implying that the level of consumption at time t is

$$C^{*}(t) = C^{*}(0)e^{\psi^{*}t}. \tag{12'}$$

The optimality condition with respect to private capital is now modified to:

$$(1-\tau_k)\frac{f(z^{*})-\sigma f'(z^{*})z^{*}}{q^{*}} + \frac{\dot{q}^{*}}{q^{*}} + \frac{(q^{*}-1)^2}{2h_1 q^{*}} - \delta_k = r(1-\tau_b). \tag{10.2'}$$

The interpretation of this is analogous to (10.2), though there are two differences to be noted. First, the relevant return is the net private after-tax return, where the marginal physical product of private capital increases with the degree of congestion (decreases with σ). Transversality conditions analogous to (13) also apply.

4.2 Dynamics of Private Capital Accumulation

The dynamics of private capital in the decentralized economy are now represented by

$$\dot{q}^{*} = [r(1-\tau_b)+\delta_k]q^{*} - \tfrac{(q^{*}-1)^2}{2h_1} - (1-\tau_k)[f(z^{*})-\sigma f'(z^{*})z^{*}], \tag{14.2'}$$

$$\frac{\dot{z}^{*}}{z^{*}} = \left[g\tfrac{f(z^{*})}{z^{*}} - \delta_g\right] - \left[\tfrac{q^{*}-1}{h_1} - \delta_k\right] \quad (\equiv \phi_g^{*} - \phi_k^{*}). \tag{14.4'}$$

Again, there are potential problems of nonexistence and nonuniqueness of equilibrium due to the nonlinearity of the system, further discussion of which is given in the Appendix. One important difference from the centralized economy is that the dynamics of q and z proceed independently of the shadow value of public capital. This is because the private agents in the decentralized economy face given tax rates τ_b, τ_k, whereas in the centralized economy the social rate of return, which in part drives (14.2) is a function of the shadow value ν, which in turn is determined by the shadow value of public capital, q_2.

The linearized dynamics about the steady-state equilibrium $(\tilde{q}^{*}, \tilde{z}^{*})$ in the decentralized economy are represented by:

$$\begin{bmatrix} \dot{q}^{*} \\ \dot{z}^{*} \end{bmatrix} = B \begin{bmatrix} q^{*}-\tilde{q}^{*} \\ z^{*}-\tilde{z}^{*} \end{bmatrix}, \tag{15'}$$

with

$$B = \begin{bmatrix} r(1-\tau_b) + \delta_k - \frac{\tilde{q}^*-1}{h_1} & (1-\tau_k)[\sigma f''\tilde{z}^* - (1-\sigma)f'] \\ -\frac{\tilde{z}^*}{h_1} & -\frac{g}{\tilde{z}^*}(f - \tilde{z}^* f') \end{bmatrix}.$$

The transversality condition will be met if and only if

$$r(1-\tau_b) + \delta_k > \frac{\tilde{q}^* - 1}{h_1}, \tag{16.1'}$$

thus ensuring that the stable locus is a saddlepath. This is equivalent to $r(1-\tau_b) > \tilde{\phi}^*$.

The linearized dynamics of the decentralized system are as follows. The ratio of public to private capital evolves as in (17.1), namely

$$z^*(t) - \tilde{z}^* = (z_0^* - \tilde{z}^*)e^{\lambda^* t}, \tag{17.1'}$$

where λ^* is the stable eigenvalue to (15′), while the growth rates of the two types of capital evolve as:

$$\phi_k^* - \tilde{\phi}^* = \frac{-(1-\tau_k)[\sigma f''\tilde{z}^* - (1-\sigma)f']/h_1}{[r(1-\tau_b) - \tilde{\phi}^*] - \lambda^*}(z^* - \tilde{z}^*), \tag{18.1'}$$

$$\phi_g^* - \tilde{\phi}^* = -\frac{g}{\tilde{z}^{*2}}[f - \tilde{z}^* f'](z^* - \tilde{z}^*). \tag{18.2'}$$

These relationships have analogous properties to those illustrated in figure 1.

The dynamics of the current account are obtained following the procedure discussed in section 2.4. The application of the intertemporal national budget constraint (the transversality condition on traded bonds) leads to an initial sustainable value for consumption.

4.3 Steady-State Fiscal Effects

The steady-state shadow value of private capital and the ratio of the two types of capital are determined by setting $\dot{q}^* = \dot{z}^* = 0$ in (14.2′) and (14.4′), from which the corresponding value of the equilibrium growth rate $\tilde{\phi}^*$ can be derived. This forms the basis for the long-run effects of various types of fiscal policies. Here we shall discuss the effects of changes in the tax rates and in the share of government expenditure, on the assumption that the government budget constraint is met through appropriate adjustments in lump-sum taxes. Note that this aspect of the equilibrium is independent of the consumption tax, ω, which therefore operates as a lump-sum tax; see also Rebelo (1991). Omitting details, the following results can be established:

$$\begin{aligned}
\frac{\partial \tilde{\phi}^*}{\partial \tau_b} &= \frac{r\tilde{q}^* g[f - \tilde{z}^* f']/h_1 \tilde{z}^{*2}}{\tau^*} > 0, \\
\frac{\partial \tilde{z}^*}{\partial \tau_b} &= \frac{-r\tilde{q}^*/h_1}{\tau^*} < 0, \qquad \frac{\partial \psi^*}{\partial \tau_b} = \frac{-r}{1-\gamma} < 0,
\end{aligned} \tag{23.1}$$

$$\begin{aligned}
\frac{\partial \tilde{\phi}^*}{\partial \tau_k} &= \frac{-g[f - \tilde{z}^* f'][f - \sigma \tilde{z}^* f']/h_1 \tilde{z}^{*2}}{\tau^*} < 0, \\
\frac{\partial \tilde{z}^*}{\partial \tau_k} &= \frac{(f - \sigma \tilde{z}^* f')/h_1}{\tau^*} > 0, \qquad \frac{\partial \psi^*}{\partial \tau_k} = 0
\end{aligned} \tag{23.2}$$

$$\begin{aligned}
\frac{\partial \tilde{\phi}^*}{\partial g} &= \frac{-(1-\tau_k)(f/\tilde{z}^*) f f''/h_1}{\tau^*} > 0, \\
\frac{\partial \tilde{z}^*}{\partial g} &= \frac{[r(1-\tau_b) - \tilde{\phi}^*](f/\tilde{z}^*)}{\tau^*} > 0, \qquad \frac{\partial \psi^*}{\partial g} = 0,
\end{aligned} \tag{23.3}$$

where

$$\tau^* \equiv \frac{g}{\tilde{z}^{*2}}(f - \tilde{z}^* f')[r(1-\tau_b) - \tilde{\phi}^*] + \frac{1}{h_1}(1-\tau_k)[(1-\sigma)f' - \sigma f'' \tilde{z}^*] > 0.$$

Intuitively, an increase in the tax on interest income lowers the net rate of return on traded bonds, thereby inducing investors to increase the proportion of private capital in their portfolios, raising the price of capital and inducing long-run growth in private capital. This growth in private capital reduces the equilibrium ratio of public to private capital. In addition, this tax induces agents to switch from savings to consumption, increasing the amount of initial consumption, but slowing down its growth rate.

An increase in the tax on private capital has the opposite portfolio effect, lowering the growth of private capital and public capital and increasing the ratio of public to private capital. It leaves the growth rate of consumption unaffected.

In contrast to the centralized economy, an increase in the share of output claimed by the government, financed by a lump-sum tax, raises the equilibrium growth rate of capital unambiguously. This is because lump-sum taxation avoids the excess burden of taxation associated with distortionary taxes. At the same time, the transversality conditions (13) prevent the growth rate from being increased indefinitely through an ever-increasing share of government expenditure.

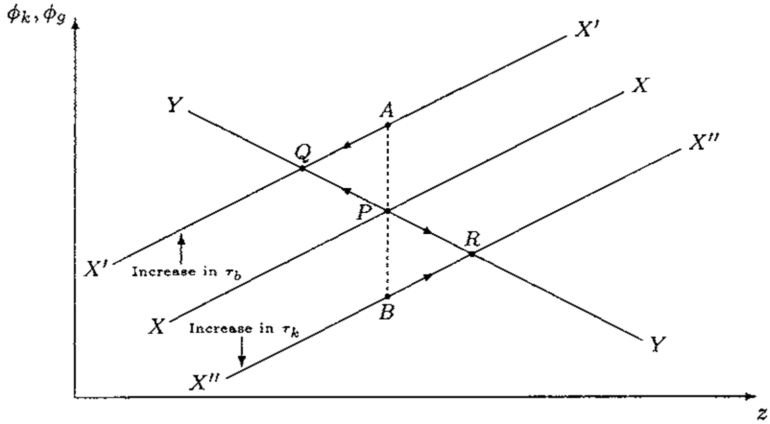
4.4 Transitional Dynamics

Figure 3 illustrates the transitional dynamics in capital following the three types of fiscal disturbances. Figure 3A illustrates the effects of higher income tax rates, while figure 3B traces out the dynamic adjustment in response to a higher proportion of government expenditure. In each case, the economy starts out initially in steady-state equilibrium at the point P.

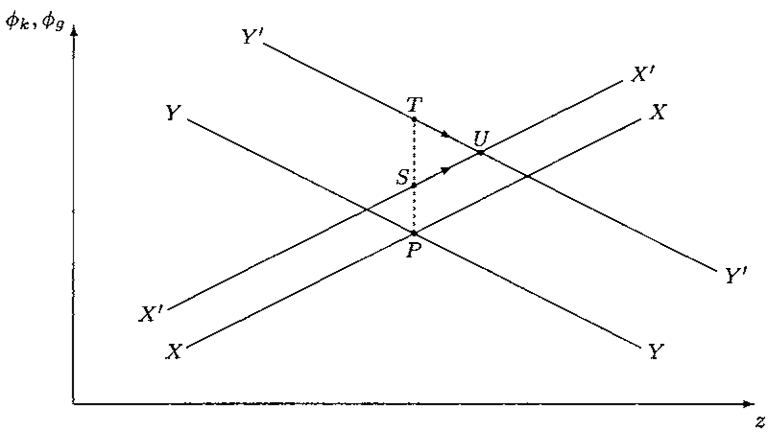
The immediate effect of an increase in the tax rate on interest income, τ_b, is to induce agents to begin switching their portfolios from bonds to capital. The rate of growth of private capital increases, reducing the ratio of public to private capital in the economy. As z declines, (i.e. the relative abundance of private capital increases), its shadow value declines, causing the growth rate of private capital to decline. The transitional adjustment in the growth rate of private capital is illustrated by the initial jump from P to A, on the new stable arm $X'X'$, followed by the continuous decline AQ, to the new steady state at Q. With the growth of public capital being tied through aggregate output to the capital stocks in accordance with (6), the growth rate of public capital does not respond instantaneously to the higher tax rate τ_b. Instead, as z declines, the average productivity of public capital f/z rises, causing the growth rate of public capital to rise gradually over time. The stable arm YY remains fixed and the growth rate of public capital occurs gradually along the path PQ.

The transitional response to a higher tax on capital, τ_k, is the mirror image of that we have just been discussing. The higher tax on capital generates an initial decline in the rate of growth of private investment, followed by a gradual, but only partial, increase. This is represented by the initial jump from P to B, to the new stable path $X''X''$, followed by the continuous increase along BR, to the new equilibrium at R. The growth of public capital does not respond immediately, but declines gradually, as its average productivity f/z falls. This is represented by the continuous movement along PR in figure 3A.

The transitional adjustment of the two types of capital to an increase in government expenditure is illustrated in figure 3B. The dynamic adjustment in the decentralized economy is qualitatively the same as that in figure 2A in the centralized economy. In this case, the long-run increase in the equilibrium growth rate is sufficiently large to generate a corresponding partial increase in the short-run growth rate of private capital, followed by a further continuous increase along SU.



A. Increase in Income Tax Rates



B. Increase in Government Expenditure

Fig. 3. Transitional dynamics of capital: Decentralized economy

5. Optimal Tax Policy

We turn now to the determination of the tax structure that will enable the decentralized economy to replicate the first best outcome of the centrally planned economy. There are two general requirements to be met. The first is that the decentralized economy must ultimately attain the steady state of the centralized economy. Second, having replicated the steady state, the transitional dynamic adjustment paths in the two economies must also coincide.

To replicate the optimal adjustment path for consumption is straightforward. Comparing (11) and (11′), these two growth paths will coincide if and only if

$$\hat{\tau}_b = 0. \tag{24.1}$$

That is, the tax rate on foreign bond income should be zero.

To replicate the growth rates of the capital stocks in the two economies is more involved. First, the rate of adjustment of the relative stocks; i.e. the time path for z^* given in (17.1′) must replicate that of z given in (17.1). This will be so if and only if the stable eigenvalue λ^* for (15′) equals the corresponding eigenvalue λ for (15) and as we shall see, this requires the optimal capital income tax rate, τ_k, to be time-varying. Having matched the relative capital stocks, we also need to replicate their corresponding shadow values and growth rates. As indicated in footnote 29 below, once one has set $\lambda^* = \lambda$, this in fact is assured.

To see the time-varying nature of τ_k, first consider the case where the capital income tax rate remains constant through time at the rate $\tau_k = \hat{\tau}_k$. Comparing the steady-state relationships (20.1)–(20.4) with the corresponding conditions in the decentralized economy, we see that the steady-state equilibrium values $(\tilde{z}^*, \tilde{q}^*)$ will replicate the first-best optimum if and only if $\hat{\tau}_b = 0$ and $\hat{\tau}_k$ satisfies

$$(1 - \hat{\tau}_k)[f - \sigma\tilde{z}f'] = (1 + \tilde{\nu}g)[f - \tilde{z}f'].$$

Simplifying this relationship, the optimal steady-state capital income tax can be expressed as

$$\hat{\tau}_k = -\tilde{\nu}g + \frac{(1-\sigma)f'\tilde{z}(1+\tilde{\nu}g)}{f - \sigma\tilde{z}f'}. \tag{24.2}$$

Setting the two income tax rates in accordance with (24.1)–(24.2) ensures that the steady-state equilibrium of the centrally planned economy will be replicated. We shall discuss the intuition underlying this steady-state tax policy presently, but before doing so we shall show how if τ_k

is maintained at $\hat{\tau}_k$ during the transition, the adjustment path followed by the decentralized equilibrium will fail to mimic that of the first best optimum. To see this we consider the respective eigenvalues and show how in this circumstance $\lambda^* \neq \lambda$.

For notational convenience we denote the elements of the matrix of coefficients in the linearized centralized economy by (a_{ij}). These elements can be immediately identified by referring to (15). The equilibrium (stable) eigenvalue in the centralized economy is thus the unique negative solution to the cubic equation:

$$F(\lambda) \equiv (a_{22} - \lambda)[(a_{11} - \lambda)(a_{33} - \lambda) - a_{13}a_{31}] + a_{12}a_{23}a_{31} = 0. \quad (25)$$

Using this notation and if the tax rates $\hat{\tau}_b$, $\hat{\tau}_k$ in the decentralized economy are set in accordance with (24.1)–(24.2), thereby replicating the steady-state, then from (15′) the corresponding eigenvalue, λ^*, in the decentralized economy is determined where

$$G(\lambda^*) \equiv (a_{11} - \lambda^*)(a_{33} - \lambda^*) + \frac{\tilde{z}}{h_1}(1 + \tilde{\nu} g)\left(\frac{f - \tilde{z}f'}{f - \sigma\tilde{z}f'}\right)[\sigma\tilde{z}f'' - (1-\sigma)f'] = 0. \quad (25')$$

Combining (25) and (25′) we can show that

$$F(\lambda^*) = a_{31}\left[(a_{11} - \lambda^*)\left(\left[\frac{(1+\tilde{\nu}g)f''f'\tilde{z}^2(\sigma-1)}{f-\sigma\tilde{z}f'}\right] + \frac{g^2h_2}{z^2}[f - \tilde{z}f']^2\right) + \frac{(1+\tilde{\nu}g)gf''(f-\tilde{z}f')^2}{f-\sigma\tilde{z}f'}\right].$$

It then follows from the fact that $F(\cdot)$ is cubic in λ and that λ, λ^* are stable eigenvalues that:

$$F(\lambda^*) > 0 \Rightarrow \lambda^* < \lambda < 0,$$
$$F(\lambda^*) < 0 \Rightarrow \lambda < \lambda^* < 0.$$

Thus, if the tax rates are fixed over time at $\tau_b = 0$; $\tau_k = \hat{\tau}_k$ as in (24.1)–(24.2), then the ratio of public to private capital in the decentralized economy, z^*, determined by (17.1′) will in general converge at a nonoptimal rate, relative to the first-best rate of adjustment, as described by (17.1). Whether the adjustment in the decentralized economy is too fast or too slow depends among other things upon: (i) degree of congestion associated with public capital $(1-\sigma)$ and (ii) the adjustment costs (h_2).

In the case that these are both absent $(\sigma = 1, h_2 = 0)$, $F(\lambda^*) > 0$ so that $\lambda^* < \lambda < 0$. The intuition for this result is straightforward and is a consequence of the fact that the private agent treats τ_k as fixed and does not respond to changes in the shadow value of public capital, q_2, as does the central planner. Suppose some change occurs causing z to increase from z_0 to $\tilde{z}$. During the transition as z is increased, the shadow value of public capital declines. This, however, is not reflected by a fixed τ_k, so that during the transition $\hat{\tau}_k$ overstates the proper social value of public capital. Accordingly, private capital is taxed too much and there is an overinvestment in public capital relative to private capital along the transitional path in the decentralized economy . But while the relationship $z^* > z$ holds along the transitional path, asymptotically $z^* \to \tilde{z}$.

If the public capital is subject to substantial adjustment costs, or to congestion, this tends to raise the shadow value of installed public capital. It is now possible for the tax rate on private capital to be too low relative to the social optimum, leading to a relative underinvestment of public capital in the decentralized economy during the transition.

We now propose modifying the tax rate on capital income to

$$\tau_k(t) = \hat{\tau}_k + \theta[z^*(t) - \tilde{z}^*],$$

where $\hat{\tau}_k$ is given by (24.2) and θ is a constant, to be determined. The income tax rate as specified by (26) is time-varying, tracking the evolution of the economy as the relative stocks of capital change over time. Intuitively, the time-varying tax rate $\tau_k(t)$ in effect permits the representative agent to track the endogenous shadow value of public capital. Since θ is relevant only along the transitional path (when $z^* \neq \tilde{z}^*$) it has no impact on the steady-state equilibrium. Consequently, setting $\hat{\tau}_k$ in accordance with (24.2) will still replicate the steady-state capital and shadow value, $\tilde{z}$, $\tilde{q}_1$, of the first best optimum.[28]

However, θ will affect the eigenvalue λ^* in the decentralized economy and therefore the speed of adjustment along the transitional path. In particular, if $\tau_k(t)$ is generated by (26), the critical modification to be made is to the linearization of (10.2′), [appearing as the first row in (15′)] which now becomes

$$\begin{aligned}\dot{q}^* = &\left[r(1-\tau_b) + \delta_k - \tfrac{\tilde{q}^*-1}{h_1}\right](q^* - \tilde{q}^*)\\ &+ [(1-\hat{\tau}_k)[\sigma f'' z^* - (1-\sigma)f'] + \theta(f - \sigma z^* f')](z^* - \tilde{z}^*).\end{aligned}$$

If $\hat{\tau}_k$ is set in accordance with (24.2), the eigenvalue, λ^*, in the decentralized economy is now determined by

$$G(\lambda^*, \theta) \equiv (a_{11} - \lambda^*)(a_{33} - \lambda^*) + \tfrac{z}{h_1}\{(1+\nu g)\left(\tfrac{f - zf'}{f - \sigma z f'}\right)$$

28. Since the time-varying tax rate is specified as a function of the ratio of the aggregate stock of public to private capital, we assume that the representative agent takes this tax rate as given when making his own individual decisions.

$$\cdot [\sigma z f'' - (1-\sigma) f'] + \theta [f - \sigma z^* f']\} = 0. \tag{25''}$$

It then follows that the speed of adjustment in the linearized decentralized economy will replicate that in the centralized economy (i.e. $\lambda^* = \lambda$) if and only if θ is set such that λ^*, the solution to (25″) also satisfies the condition $F(\lambda^*) = 0$; i.e. if and only if $G(\lambda^*, \theta) = F(\lambda^*) = 0$. By appropriate choice of θ this can always be achieved.[29]

Thus the time-varying capital income tax rate (26) where $\hat{\tau}_k$ is determined by (24.2) and θ is determined by (25″) will replicate (to a linear approximation) the first best optimum in the sense that both its transitional path and the steady-state will be attained. Having set the distortionary income taxes in accordance with (24.1)–(24.2), the government budget constraint will be met if and only if lump-sum taxes and/or the consumption tax adjust to satisfy

$$(\hat{T}/K) + \hat{\omega}(C/K) = gf(z)[1 + (h_2/2)gf(z)/z] - \tau_k(t)f(z). \tag{24.3}$$

Note further that with the availability of a full set of tax instruments the problem of time inconsistency of optimal policy does not arise. With the target value for the income tax rate at each instant of time being determined by the time path followed by the first best optimum, the government will always want to choose the income tax rate to attain that given and unchanging target path.

We return to the optimal steady-state capital income tax rate, $\hat{\tau}_k$, given in (24.2). The intuition behind this optimum can be understood by comparing the *social* and *private* returns to private capital accumulation in the presence of public capital. Recalling (10.2), the steady-state social return to accumulating a marginal unit of private capital is:

$$r_s \equiv (1 + \tilde{\nu} g)\frac{[f - \tilde{z} f']}{\tilde{q}_1} + \frac{(\tilde{q}_1 - 1)^2}{2h_1 \tilde{q}_1} - \delta_k.$$

This takes account of the fact since the government maintains a fixed expenditure ratio, gY, the accumulation of private capital indirectly causes the government to increase its rate of investment.

By contrast, the individual in the decentralized economy computes the marginal physical product of private capital on the assumption that the value of the public capital, K_g, remains unaffected by his individual

29. Substituting (25′) into (25) one can show that having set $\lambda^* = \lambda$, the stable solution for the shadow value, q^*, in the decentralized economy with time-varying taxes, replicates the corresponding path in the centralized economy, namely (17.2). The time paths for the corresponding growth rates, ϕ_k and ϕ_k^*, also coincide.

decision. Thus the steady-state after-tax private rate of return on private capital is:

$$r_p \equiv (1-\tau_k)\frac{[f-\sigma\tilde{z}f']}{\tilde{q}^*}+\frac{(\tilde{q}^*-1)^2}{2h_1\tilde{q}^*}-\delta_k,$$

which takes account of the degree of congestion associated with the public capital. The optimal tax rate $\hat{\tau}_k$ is set so as to equate r_p to r_s. The income tax rate thus corrects for two potential sources of externality: (i) the size of the government relative to its social optimum, and (ii) the degree of congestion.

Suppose that there is no congestion, so that $\sigma = 1$, and that $\tilde{\nu} > 0$, i.e. $\tilde{z} < \hat{z}$ so that the relative stock of government capital is less than optimal. In this case, the optimal tax on private capital income is $\hat{\tau}_k < 0$; see (24.2). Since private investment increases output and therefore has the desirable effect of increasing the size of public capital, it generates a positive externality and therefore should be encouraged through a subsidy. On the other hand, if $\tilde{\nu} < 0$ and the government is too large relative to the optimum, capital income should be taxed positively. This is because the induced expansion of the government through private investment now generates a negative externality and should be discouraged through taxation. Finally, if $\tilde{\nu} = 0$, so that the size of the government sector is optimal, the induced change in government expenditure is just worth its cost. There is no externality and so private capital income should be untaxed. The first best optimum can be reached either through lump-sum taxation alone, or equivalently through a consumption tax. At the other extreme, suppose that $\sigma = 0$ so that congestion is proportional. If the stock of public capital is at its social optimum, $\tilde{\nu} = 0$, the income from private capital should now be taxed at the rate $\hat{\tau}_k = \tilde{z}^* f'/f$, the share of public capital in the overall social optimum.

The idea that the presence of congestion favors an income tax over lump-sum taxation or a consumption tax has been shown previously by Barro and Sala-i-Martin (1992a) and Turnovsky (1996a). In these models, in which government expenditure appears as a flow, there are no adjustment costs and if congestion is proportional ($\sigma = 0$), the optimal tax rate turns out to be $\hat{\tau}_k = \hat{g}$ so that the expenditure is fully financed by the capital income tax. In the present case, $\hat{\tau}_k \gtrless \hat{g}$, reflecting the fact that while congestion in public capital enhances the return to private capital, thus providing an incentive for private investment, this needs to be weighed against the adjustment costs associated with the latter.

This result that the optimal tax rate does depend upon the degree of congestion contrasts with that of Glomm and Ravikumar (1994), who reach the opposite conclusion. The difference is due to the formulation of congestion and the fact that we are imposing constant returns to scale in the two forms of capital in the absence of labor. If we adopt the Glomm-Ravikumar specification of congestion, the only assumption consistent with ongoing growth is for $\sigma = 0$, in which case our expression (24.2) with $\tilde{\nu} = 0$ also reduces to the Barro expression $\hat{\tau}_k = f'\tilde{z}/f$.[30]

In general $\theta \gtrless 0$, depending upon the degree of congestion and the adjustment costs. In the absence of adjustment costs, $\theta < 0$ in which case the transitional component of the tax rate, $\tau_k(t)$, is a subsidy as long as $\tilde{z}^* > z^*(t)$, favoring the accumulation of private capital. In the absence of such a subsidy, the ratio of public capital in the decentralized economy will accumulate too fast relative to the social optimum and the effect of $\theta < 0$ is to slow down the speed of adjustment. Notice that as the ratio z approaches its steady-state, the magnitude of the subsidy along the path declines. If $\tilde{z}^* < z^*(t)$ it is a tax slowing down the contraction of z (i.e. speeding up the relative contraction of private capital). If $\theta > 0$ the argument is reversed.

The other aspect of the optimal tax structure – the differential taxation of capital and interest income when g is not at its optimum – is due to the form of the government expenditure rule (6), where gross public investment is assumed to be a fixed proportion of output. It is through this relationship that the accumulation of private capital generates the externality that needs to be corrected by a tax on capital. Since government expenditure is unrelated to interest income, the accumulation of bonds by the agent generates no such externality.

While the expenditure rule (6) is plausible, it is arbitrary, and we therefore briefly consider the implications of modifying (6) to:

$$G = g\left[f\left(\frac{K_g}{K}\right)K + rb\right]. \tag{6'}$$

so that government expenditure is proportional to GNP. In this case the accumulation of bonds generates an externality completely analogous to that generated by private capital. To replicate the first-best optimum will therefore require the taxation of both forms of income and with G being proportional to the sum of the income sources, both sources of income will have to be taxed equally in order to replicate the first-best equilibrium.[31]

30. Glomm and Ravikumar (1994) specify congestion (using our notation) in the form $K_g^s = K_g/K^\rho$, where $\rho \geq 0$, rather than in the form (1.2).

31. If $G = gfK + g'rb$, then the two forms of income will be taxed at differential rates. The specifications in (6) and (6′) correspond to polar cases. See Turnovsky (1996b). When government expenditure is optimally determined, the specifics of the underlying rule cease to matter.

6. Conclusions

Recently, economists have become interested in the role of public expenditure in determining the productive performance of the economy. Virtually all of the analytical work addressing this issue has introduced government expenditure as a flow in the production function. It is therefore subject to the criticism that insofar as it is intended to represent the infrastructure of the economy, it is an inadequate measure of what is really relevant, namely the accumulated stock of publicly provided capital. This chapter has introduced both public and private capital into an endogenous growth model of a small open economy. Apart from its intrinsic importance, the small open economy has the advantage of enabling us to focus on the dynamic interaction in the adjustments of the two types of capital in the most transparent way.

We conclude by drawing the parallels and highlighting the differences between considering productive government expenditure in the form of capital, with the more standard practice of introducing it as a flow. The first difference is that the introduction of public together with private capital generates transitional dynamics in the growth of both types of capital. This is in contrast to the case where government expenditure appears as a flow, when the private capital stock is always on its balanced growth path; see e.g. Barro (1990), Turnovsky (1996a). In this respect, the dynamics are analogous to those characterizing the two sector endogenous growth models that incorporate both physical and human capital; see e.g. Lucas (1988), Mulligan and Sala-i-Martin (1993). Second, not only do the two types of capital evolve at different time-varying rates during the transition, but they also approach their common equilibrium growth rate from opposite directions. In response to an increase in the size of the government, say, the growth of public capital initially overshoots, before gradually declining to the new equilibrium growth rate. The growth rate of private capital always undershoots on impact – and indeed may initially respond perversely – before gradually increasing to its new equilibrium. This pattern of adjustment is reversed in response to tax changes. Now the growth rate of private capital initially overshoots its long-run response – positively in the case of a tax on interest, negatively in the case of a tax on capital – while the growth of public capital adjusts gradually to the new equilibrium.

Third, as in the case where productive government expenditure impacts as a flow, there is a growth-maximizing size of productive government expenditure. However, in contrast to that case, maximizing the equilibrium growth rate does not coincide with welfare maximization. The process of accumulating the public capital necessary to maximize the equilibrium growth rate of capital may involve consumption losses, which more than outweigh the benefits to future production. The economy may be better off with a slightly lower growth rate and higher consumption.

Finally, as in the more conventional formulation, the introduction of government capital introduces an externality in production. As in the simpler model this can be corrected by a combination of income taxes and/or lump-sum taxes, enabling the decentralized economy to replicate the first-best equilibrium of the centrally planned economy. But in contrast to the simple model, the income tax necessary to achieve this varies along the transitional path. The steady-state component has a simple structure aimed at correcting for potential externalities due to: (i) the deviation in government expenditure from its social optimum, and (ii) the effects of congestion associated with public capital. The transitional component is aimed at inducing the representative agent to take proper account of the fact that the shadow value of public capital varies inversely with the changing ratio of public to private capital along the adjustment path.[32]

Appendix

This Appendix discusses the potential problems of existence of equilibrium in the two economies.

A.1 Centrally Planned Economy

Consider the steady-state to the dynamic system (14.1)–(14.4), described by (20.1)–(20.4). Potential problems of nonexistence and nonuniqueness of equilibrium arise from the fact that the returns to capital are quadratic in their respective shadow values $\tilde{q}_1$, $\tilde{q}_2$. A feasible equilibrium is one in which $\tilde{q}_1 > 0$, $\tilde{q}_2 > 0$, and $\tilde{z} > 0$. The shadow value $\tilde{\nu} \gtrless 0$, depending upon the size of the government relative to its steady-state optimum. Existence will depend in part upon the specific form of the production function and for simplicity we shall assume $f(z) \equiv \alpha z^{\theta}$, $0 \leq \theta \leq 1$, so that θ is the elasticity of public capital in the production function. Also for simplicity we shall assume a common depreciation rate $\delta_k = \delta_g = \delta$.

32. We may observe that with the consumption-tax essentially operating as a lump-sum tax, the issue of time inconsistency does not arise. Given an unchanging time path characterizing the first-best optimum, the policy maker will have no incentive to deviate from it.

With this notation, equations (20.1) and (20.4) yield

$$\tilde{q}_1 = 1 + h_1 g\alpha\tilde{z}^{\theta-1}, \qquad \tilde{q}_2 = 1 + \tilde{\nu} + h_2 g\alpha\tilde{z}^{\theta-1},$$

so that (20.2)–(20.3) become:

$$(h_1/2)g^2\alpha\tilde{z}^{2(\theta-1)} + (1+\tilde{\nu}g)(1-\theta)\alpha\tilde{z}^{\theta} = (r+\delta)[1 + h_1 g\alpha\tilde{z}^{\theta-1}], \quad \text{(A.1.1)}$$
$$(h_2/2)g^2\alpha\tilde{z}^{2(\theta-1)} + (1+\tilde{\nu}g)\theta\alpha\tilde{z}^{\theta-1} = (r+\delta)[1 + \tilde{\nu} + h_2 g\alpha\tilde{z}^{\theta-1}]. \quad \text{(A.1.2)}$$

These two equations jointly determine equilibrium solutions for $\tilde{z}$ and $\tilde{\nu}$, from which solutions for $\tilde{q}_1$ and $\tilde{q}_2$ follow. Existence of a feasible equilibrium requires that $\tilde{z} > 0$ and that the implied solutions for $\tilde{q}_i > 0$ and be consistent with the transversality conditions. Whether or not these conditions are met depends upon: (i) the adjustment costs h_i; (ii) the importance of public capital θ; (iii) the size of government relative to its optimum.

To focus on the costs of adjustment it is convenient to abstract from public capital by assuming $\theta = g = 0$. In this case, the solution for $\tilde{q}_1$ is obtained directly from (20.2), written in the form

$$\tilde{q}_1^2 - 2[1 + h_1(r+\delta)]\tilde{q}_1 + [1 + 2h_1\alpha] = 0. \quad \text{(A.2)}$$

This equation has real roots and therefore an equilibrium solution for $\tilde{q}_1$ exists if and only if

$$\alpha < (r+\delta)[1 + h_1(r+\delta)/2]. \quad \text{(A.3)}$$

Assuming (A.3) holds, the solutions for $\tilde{q}_1$ are

$$\tilde{q}_1 = 1 + h_1(r+\delta) \pm \sqrt{[1 + h_1(r+\delta)]^2 - [1 + 2h_1\alpha]}. \quad \text{(A.4)}$$

However, the positive root can be ruled out, since it violates the transversality condition (16.1).

To consider the importance of government, we abstract from adjustment costs, setting $h_1 = h_2 = 0$. In this case $\tilde{q}_1 = 1$, $\tilde{q}_2 = 1 + \tilde{\nu}$ and (20.2)–(20.3) reduce to

$$(1+\tilde{\nu}g)(1-\theta)\tilde{z}^{\theta} = r + \delta, \quad \text{(A.5.1)}$$
$$(1+\tilde{\nu}g)\theta\tilde{z}^{\theta-1} = (r+\delta)(1+\tilde{\nu}). \quad \text{(A.5.2)}$$

Eliminating ν from these two equations leads to the following equation in $\tilde{z}$:

$$[(1-\theta)(1-g)\tilde{z} + g\theta]\tilde{z}^{\theta-1} = r + \delta. \quad \text{(A.6)}$$

It is straightforward to show that this equation will have a positive real solution for $\tilde{z}$ if and only if $(r+\delta) > g^{\theta}(1-g)^{1-\theta}$. In that case there are in fact two solutions and the larger can be rejected as being inconsistent with the transversality conditions.

As a third example, suppose $\tilde{\nu} = 0$, so that the size of the government is at its optimum and that $h_1 = h_2$. In this case (A.1.1)–(A.1.2) imply $\tilde{z} = \theta/(1-\theta)$, again ensuring a well defined equilibrium. Nonexistence of equilibrium is thus associated with having a nonoptimal size of government.

Conditions for the existence of equilibrium in more general cases will require the use of numerical methods.

A.2 Decentralized Economy

In the decentralized economy the relevant conditions pertinent to the existence of an equilibrium are steady-state conditions to (14.2′) and (14.4′). These can be combined to yield

$$\begin{aligned}(h_1/2)g^2\alpha\tilde{z}^{2(\theta-1)} &+ (1-\tau_k)(1-\theta\sigma)\alpha\tilde{z}^{\theta} \\ &= (r(1-\tau_b)+\delta)(1+h_1 g\alpha\tilde{z}^{\theta-1}),\end{aligned} \tag{A.1.1′}$$

and there will be a well defined equilibrium if and only the solution to this equation $\tilde{z}^* > 0$. Whether this is so depends upon h_1, g, and θ as before, as well as now the tax rates, τ_k, τ_b and the degree of congestion σ. By considering this equation one can show that the likelihood of a feasible equilibrium increases with the tax on capital τ_k, but decreases with the degree of congestion (i.e. a declining σ) and the tax on interest income, τ_b. As was shown in the case of the centralized economy it is possible that there are two (or more) solutions $\tilde{z}^* > 0$. The transversality conditions can then be applied to eliminate one or more of these.

Acknowledgement. I am grateful to Theo Eicher and to participants in seminar presentations at the University of Washington for their comments. The comments of the discussant, Alasdair Smith, are also gratefully acknowledged.

References

Arrow, Kenneth J., and Mordecai Kurz (1970). *Public Investment, the Rate of Return, and Optimal Fiscal Policy.* Baltimore: Johns Hopkins University Press.

Aschauer, David A. (1988). "The Equilibrium Approach to Fiscal Policy." *Journal of Money, Credit, and Banking* 20, 41–62.

— (1989a). "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23, 177–200.

— (1989b). "Does Public Capital Crowd Out Private Capital?" *Journal of Monetary Economics* 24, 171–188.

— and Jeremy Greenwood (1985). "Macroeconomic Effects of Fiscal Policy." *Carnegie-Rochester Conference Series on Public Policy* 23, 91–138.

Barro, Robert J. (1989). "The Neoclassical Approach to Fiscal Policy." In Robert J. Barro (ed.), *Modern Business Cycle Theory.* Cambridge: Harvard University Press.

— (1990). "Government Spending in a Simple Model of Endogenous Growth." *Journal of Political Economy* 98, S103–S125.

— and Xavier Sala-i-Martin (1992a). "Public Finance in Models of Economic Growth." *Review of Economic Studies* 59, 645–661.

— and Xavier Sala-i-Martin (1992b). "Convergence." *Journal of Political Economy* 100, 223–251.

— and Xavier Sala-i-Martin (1995). *Economic Growth.* New York: McGraw-Hill.

Baxter, Marianne, and Robert G. King (1993). "Fiscal Policy in General Equilibrium." *American Economic Review* 83, 315–334.

Clarida, Richard, and Ronald Findlay (1992). "Optimal Endogenous Growth, Public Capital, and the Dynamic Gains from Trade." Unpublished Columbia University.

Cornes, Richard, and Todd Sandler (1986). *The Theory of Externalities, Public Goods, and Club Goods.* Cambridge University Press.

Easterly, William, and Sergio Rebelo (1993). "Fiscal Policy and Economic Growth: An Empirical Investigation." *Journal of Monetary Economics* 32, 417–458.

Ebrill, Liam P., and Steven M. Slutsky (1982). "Time, Congestion, and Public Goods." *Journal of Public Economics* 17, 307–334.

Edwards, John H.Y. (1990). "Congestion Function Specification and the 'Publicness' of Local Public Goods," *Journal of Urban Economics* 27, 80–96.

Futagami, Koichi, Yuichi Morita, and Akihisa Shibata (1993). "Dynamic Analysis of an Endogenous Growth Model with Public Capital." *Scandinavian Journal of Economics* 95, 607–625.

Glomm, Gerhard, and B. Ravikumar (1994). "Public Investment in Infrastructure in a Simple Growth Model" *Journal of Economic Dynamics and Control* 18, 1173–1187.

Gramlich, Edward M. (1994). "Infrastructure Investment: A Review Essay." *Journal of Economic Literature* 32, 1176–1196.

Hayashi, Fumio (1982). "Tobin's Marginal and Average q: A Neoclassical

Interpretation." *Econometrica* 50, 213–224.

Jones, Charles I. (1995). "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics* 110, 495–527.

Jones, Larry E., and Rodolfo E. Manuelli (1990). "A Convex Model of Equilibrium Growth: Theory and Policy Implications." *Journal of Political Economy* 98, 1008–1038.

Jones, Larry E., Rodolfo E. Manuelli, and Peter E. Rossi (1993). "Optimal Taxation in Models of Endogenous Growth." *Journal of Political Economy* 101, 485–517.

King, Robert G., and Sergio Rebelo (1990). "Public Policy and Economic Growth: Developing Neoclassical Implications." *Journal of Political Economy* 98, S126–S150.

Lee, Yeonho (1995). "The Effects of Fiscal Policy in a Two-country World Economy: An Intertemporal Analysis." *Journal of Money, Credit, and Banking* 27, 742–761.

Lucas, Robert E. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22, 3–42.

Mulligan, Casey B., and Xavier Sala-i-Martin (1993). "Transitional Dynamics in Two-Sector Models of Endogenous Growth." *Quarterly Journal of Economics* 108, 739–773.

Oakland, William H. (1972). "Congestion, Public Goods, and Welfare." *Journal of Public Economics* 1, 339–357.

Ortigueira, Salvador, and Manuel Santos (1997). "On the Speed of Convergence in Endogenous Growth Models." *American Economic Review* 87, 383–399.

Pecorino, Paul (1993). "Tax Structure and Growth in a Model with Human Capital." *Journal of Public Economics* 52, 251–271.

Rebelo, Sergio (1991)."Long-run Policy Analysis and Long-run Growth." *Journal of Political Economy* 99, 500–521.

Rodrik, Dani (1996). "Why Do More Open Economies Have Bigger Governments?" NBER Working Paper No. 5537.

Romer, Paul M. (1986). "Increasing Returns and Long-run Growth." *Journal of Political Economy* 94, 1002–1037.

Saint-Paul, Gilles (1992). "Fiscal Policy in an Endogenous Growth Model." *Quarterly Journal of Economics* 107, 1243–1259.

Thompson, Earl A. (1976). "Taxation and National Defense." *Journal of Political Economy* 82, 755–782.

Turnovsky, Stephen J. (1996a). "Optimal Tax, Debt, and Expenditure Policies in a Growing Economy." *Journal of Public Economics* 60, 21–44.

— (1996b). "Fiscal Policy, Growth, and Macroeconomic Performance in a Small Open Economy." *Journal of International Economics* 40,

41–66.
— (1996c). "Fiscal Policy, Adjustment Costs, and Endogenous Growth." *Oxford Economic Papers* 48, 361–381.
— (1997). "Fiscal Policy in a Growing Economy with Public Capital." *Macroeconomic Dynamics* 1.
— and Walter H. Fisher (1995). "The Composition of Government Expenditure and its Consequences for Macroeconomic Performance." *Journal of Economic Dynamics and Control* 19, 747–786.

CHAPTER 7

Trade and Growth with Endogenous Human and Physical Capital Accumulation

Eric W. Bond and Kathleen Trask

1. Introduction

The analysis of capital accumulation has been a central focus of growth theory, starting with the one sector growth model of Solow (1956). More recently, growth models have been enriched to consider the accumulation of both physical and human capital. This interest in human capital accumulation was spurred by two influential papers by Lucas (1988, 1993), who argued that human capital accumulation is the "engine of development" and a critical factor in explaining the East Asian "miracles" of economic growth. Empirical evidence in support of the role of human capital in the growth process has been found both in cross-sectional studies of growth rates [e.g. Romer (1990), Barro (1991)] and in studies of the East Asian miracle countries [Young (1995), Tallman and Wang (1994)].[1] This has led to the development of two sector models of endogenous economic growth with physical and human capital accumulation [e.g. Rebelo (1991), Caballé and Santos (1993), Bond, Wang, and Yip (1996)] which examine how interactions between the stocks of physical and human capital affect the growth process when physical and human capital are not perfect substitutes in production.[2] These models provide conditions under which there is a balanced growth path (BGP) in which physical and human capital grow at the same rate, and show the existence of a saddle path adjustment process to the BGP in which physical capital grows more rapidly than human capital when its relative stock is below the BGP value. The relationship between the growth rates of the capital stocks and consumption during the transition process depends

1. Tallman and Wang (1994) calculate that for the case of Taiwan, human capital growth accounted for 45% of output growth.

2. Rebelo (1991) analyzes a two sector model with Cobb-Douglas production technologies in each sector when there are no externalities from capital accumulation. Mulligan and Sala-i-Martin (1993) consider a two sector model in which there are externalities from the stocks of human capital. Caballé and Santos (1993) and Bond, Wang, and Yip (1996) examine the existence of balanced growth equilibria and characterize transitional dynamics for the case without externalities under more general assumptions regarding functional forms.

on the factor intensities of the respective production sectors. While these two sector endogenous growth models provide useful insights about the transition process, they have generally been closed economy models that ignore how factor accumulation decisions are affected by the presence of international trade.

The purpose of this chapter is to present an endogenous growth model of physical and human capital accumulation in a small open economy. This chapter will focus on two aspects of the interaction between international trade and the capital accumulation process. The first concerns how international trade can affect the economy's adjustment to imbalances in factor stocks in the absence of international capital mobility. In an open economy, a shortage of one factor can be dealt with by importing goods which use the scarce factor intensively. Second, international trade may alter the relative returns to investment in physical and human capital and thus affect the long run factor stocks in the economy. Since much of the motivation for the analysis of human capital accumulation has come from the experience of small open economies, it is important to understand the role of trade in the factor accumulation process.

We utilize a dynamic general equilibrium model in which there are two traded goods, a consumption good and an investment good, and a non-traded good, education, with additions to the stock of physical (human) capital produced by output of the investment (education) sector. Output is assumed to be produced under conditions of constant returns to scale and perfect competition, so that the model exhibits endogenous growth because it has constant returns to scale in the reproducible factors. Since it is natural to think of human capital as being a non-traded good, we have generalized the two sector growth models by adding an additional sector to allow for international trade.[3] The small open economy will thus face given prices for the consumption and investment goods, but the price of education is endogenously determined.

This model is related to the 2×2 Heckscher-Ohlin model of international trade theory, the cornerstone of factor proportions trade theory, because it has two primary factors and two traded goods. The model differs, however, from the existing dynamic literature on Heckscher-Ohlin surveyed by Smith (1984) because the supplies of both of the primary factors of production are endogenously determined in the long run. Dynamic versions of the Heckscher-Ohlin model have generally focused on

3. Stokey and Rebelo (1995) use a closed economy model with a similar production structure to analyze the effects of factor taxes on the long run rate of growth. They perform simulation analysis to calculate the effects of changes in the policy parameters.

the case in which the physical capital stock is determined endogenously by investment decisions, but the growth of the labor force is exogenously given.[4] We show that when both factors are reproducible, the model has a distinctly Ricardian flavor in the long run because long run comparative advantage must be based on technological differences.

We first establish the existence of a balanced growth path (BGP) for the small open economy given a fixed world price of traded goods, and show the relationship between the world price and the pattern of production on the BGP. We show that there is a unique world price at which the small country is incompletely specialized. The world price associated with incomplete specialization is also the price on the BGP for the small open economy under autarky. If the price of the investment good is greater (less) than this critical value, the country will specialize in production of the investment (consumption) good and the non-traded good. The reason for the knife edge nature of the incomplete specialization equilibrium is the intertemporal arbitrage condition, which requires that the returns to physical and human capital be equalized at the margin. We then show that there is a unique capital/labor ratio on the BGP for the equilibria with specialization in one of the traded goods and we establish the saddle path stability of these equilibria. However, the equilibrium with incomplete specialization is shown to be consistent with a continuum of capital/labor ratios in which there is balanced growth.

We also analyze how the BGP is affected by changes in the world price. We show that when the country is specialized in the investment good, increases in the world price of the investment good have no effect on the growth rate or on the domestic capital/labor ratio but do lead to a proportional increase in consumption per unit of labor. When the country is specialized in the consumption good, increases in the relative price of the investment good will decrease the rate of growth and raise the domestic rental on capital relative to the wage rate. This will reduce the

4. An exception is Grossman and Helpman (1991, Chapter 5) who examine accumulation of two "factors", technology and (physical or human) capital, in a small open economy. Their model differs from ours in that knowledge is the engine of growth due to the presence of scale economies. In this chapter, we limit our analysis to the case in which education is a private good whose returns are fully internalized by the owner. Growth in human capital thus represents accumulation of a productive factor, rather than pure technical change. The importance of factor accumulation is stressed by Young (1995), who finds that factor accumulation (physical and human) is responsible for most of the rapid growth in several East Asian economies. Also, Jensen and Wang (1997) examine a case in which accumulation of both capital and labor are endogenously determined, with the rate of labor force growth proportional to the level of per capita consumption. Under their specification, population growth is not a result of an explicit optimizing decision on the part of households, which may account for their finding that incomplete specialization occurs for a range of prices.

sectoral capital/labor ratios on the BGP and will reduce the aggregate capital/labor ratio if the education sector is capital intensive relative to the consumption good sector. If the education sector is labor intensive relative to the consumption good sector, however, the capital/labor ratio may either rise or fall with increases in the price of the investment good.

The results on the relationship between the terms of trade and the real return to the investment good can be used to show that the growth rate on the BGP under free trade can never be lower than the growth rate on the autarky BGP. The growth rate with trade is strictly higher than the autarky growth rate if the country specializes in the consumption good, and is equal to the autarky growth rate if the country specializes in the investment good. This establishes a sense in which the opening of trade is favorable to economic growth, even though there are no scale economies associated with obtaining access to the world market.[5]

We also examine the effect of technical progress on the growth rate and comparative advantage. We show that technical progress in any sector where production is taking place will result in an increase in the rate of growth on the BGP. Technical progress in either of the traded goods expands the set of prices for which the country exports that good. Technical progress in the education sector will make it more likely that the country exports the labor intensive good. Thus, BGP comparative advantage in this model is determined by the technological factors, as in the static Ricardian model. The trade pattern on the BGP is independent of the initial factor endowment ratio of the country, and also of the country's discount rate.

In section 2 we present the basic growth model and establish existence, uniqueness, and saddle path stability results for the BGP. An analysis of the effects of changes in world prices and technologies on the BGP is presented in section 3, and section 4 offers some concluding remarks on the results for the small country case and their implications for the world equilibrium.

2. Balanced Growth Paths for the Small Open Economy

In this section we present a model of endogenous growth in which there are two reproducible factors of production, physical and human capital,

5. A favorable effect of trade on the rate of growth is obtained in models such as those of Grossman and Helpman (1991, Chapter 9) or Rivera-Batiz and Romer (1991), where economic integration results in the access to knowledge spillovers from the rest of the world. On the other hand, an unfavorable effect of trade on the rate of growth may occur if sectors differ in the extent of knowledge spillovers to the rest of the economy. The unfavorable effect results when the opening of trade results in specialization in goods where knowledge spillovers are small [e.g. Lucas (1988)].

which are used to produce a consumption good, an investment good, and education. Endogenous growth results from the assumption of constant returns to scale in the reproducible factors of production. Consumption and investment goods are assumed to be traded, but education is non-traded. We establish the existence and uniqueness of a balanced growth path (BGP) in which consumption, human capital, and physical capital all grow at the same rate and the relative price of the non-traded good is constant. The production pattern on the BGP will be one of three types, depending on the value of the world price. There are two types of BGP in which the economy specializes in production of one traded good and imports the other, and one type of incomplete specialization equilibrium in which the country produces all three goods. We also show that the BGP equilibria with specialization exhibit saddle path stability.

2.1 The Model

We denote the capital goods sector by X, the education sector by Y, and the consumption goods sector by Z. The stock of physical (human) capital is denoted by K (H) and both factors are assumed to be perfectly mobile across sectors. All sectors are assumed to have production functions exhibiting constant returns to scale with perfect competition in goods and factor markets, so that the production technologies for the respective sectors can be expressed as

$$\begin{aligned} X &= F(s_X K, u_X H) = u_X H f(k_X), \\ Y &= G(s_Y K, u_Y H) = u_Y H g(k_Y), \\ Z &= J(s_Z K, u_Z H) = u_Z H j(k_Z), \end{aligned} \tag{1}$$

where s_i (u_i) is the share of physical (human) capital allocated to sector i and $k_i \equiv (s_i K)/(u_i H)$ is the capital/labor ratio sector $i \in \{X, Y, Z\}$. The output per unit labor functions, f, g, and j are assumed to be strictly increasing and strictly concave. The consumption good is chosen as the numeraire and the economy faces a constant relative price, p_X, for the investment good that is determined on world markets.

It is assumed that there is no international lending and borrowing, which requires that the value of purchases of traded goods is equal to the value of production of traded goods at each point in time. Demand for traded goods is the sum of consumption, C, and gross investment in physical capital, $p_X(\dot{K} + \delta K)$, where δ is the rate of depreciation on physical capital. Using this trade balance condition, the evolution of the capital stock can be written as

$$\dot{K} = u_X H f(k_X) - \delta K + (1/p_X)[u_Z H j(k_Z) - C]. \tag{2.1}$$

The education good is non-traded, so domestic production and consumption in the Y sector are equal and

$$\dot{H} = u_Y H g(k_Y) - \eta H, \tag{2.2}$$

where η is the rate of depreciation of human capital.

A representative agent's optimization problem is

$$\begin{aligned} &\max_{C, s_i, u_i} \int_0^\infty \frac{C(t)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt \qquad \text{(P)} \\ &\text{s.t.} \quad (2.1)\text{–}(2.2), \sum_{i=X,Y,Z} u_i \le 1, \sum_{i=X,Y,Z} s_i \le 1, \\ &\qquad u_i \ge 0, \;\; s_i \ge 0, \;\; i = X, Y, Z, \\ &\qquad H(0) = H_0 > 0, \quad K(0) = K_0 > 0. \end{aligned}$$

The current value Hamiltonian for this problem can be written as

$$\begin{aligned} \max_{C, u_i, s_i} \frac{C^{1-\sigma}}{1-\sigma} &+ \mu \left[u_X H f(k_X) - \delta K + \frac{1}{p_X}(u_Z H j(k_Z) - C) \right] \\ &+ \lambda \left[u_Y H g(k_Y) - \eta H \right] \\ &+ \left[\psi_K + \sum_{i \in \{X,Y,Z\}} (\beta_i - \psi_K) s_i \right] K \\ &+ \left[\psi_H + \sum_{i \in \{X,Y,Z\}} (\alpha_i - \psi_H) u_i \right] H, \end{aligned}$$

where λ and μ are the costate variables associated with the state variables H and K respectively. ψ_K (ψ_H) is the Lagrange multiplier associated with the full employment condition for physical (human) capital and β_i (α_i) is the multiplier for the requirement that the shares of physical (human) capital devoted to sector i be non-negative. As will be demonstrated below, it is possible for the economy to shut down production of one of the traded goods on the BGP, so the non-negativity conditions may bind.

The costate variables have the interpretation of being the value of an increment of the respective capital goods, μ/p_X is the value of an increment of good Z to current income, and ψ_H (ψ_K) is the flow value of an increment of physical (human) capital at time t. Since these marginal values are all measured in utility units, it is convenient to normalize these values by the utility value of an increment of good Z to obtain shadow prices measured in terms of good Z. Therefore, we can define

$r = p_X\psi_K/\mu$ ($w = p_X\psi_H/\mu$) to be the rental value of physical (human) capital in units of Z and $p_Y = p_X\lambda/\mu$ to be the relative price of education output in units of Z. Utilizing these definitions of domestic relative prices, we can express the necessary conditions associated with a solution to (P) as

$$C^{-\sigma} - \frac{\mu}{p_X} = 0, \tag{3.1}$$

$$r = p_X f'(k_X) + \frac{p_X\beta_X}{\mu} = p_Y g'(k_Y) + \frac{p_X\beta_Y}{\mu} = j'(k_Z) + \frac{p_X\beta_Z}{\mu}, \tag{3.2}$$

$$\begin{aligned} w &= p_X[f(k_X) - k_X f'(k_X)] + p_X\alpha_X/\mu \\ &= p_Y\ [g(k_Y) - k_Y g'(k_Y)] + p_X\alpha_Y/\mu \\ &= [j(k_Z) - k_Z j'(k_Z)] + p_X\alpha_Z/\mu, \end{aligned} \tag{3.3}$$

$$\frac{\dot{\mu}}{\mu} = \rho + \delta - \frac{r}{p_X}, \tag{3.4}$$

$$\frac{\dot{\lambda}}{\lambda} = \frac{\dot{p}_Y}{p_Y} + \frac{\dot{\mu}}{\mu} = \rho + \eta - \frac{w}{p_Y}, \tag{3.5}$$

$$\lim_{t\to\infty} e^{-\rho t}\mu(t)K(t) = 0, \tag{3.6}$$

$$\lim_{t\to\infty} e^{-\rho t}\lambda(t)H(t) = 0. \tag{3.7}$$

(3.1) equates the marginal utility of consumption to the value of an increment of good Z. (3.2)–(3.3) imply that the marginal revenue product of physical and human capital are equalized across all operational sectors. (3.4)–(3.5) describe the evolution of the costate variables and (3.6)–(3.7) are the transversality conditions.

2.2 Existence of Balanced Growth Path

We now illustrate how (2.1)–(3.7) can be used to examine the existence of potential balanced growth paths for the economy. A BGP for the economy requires that the level of consumption and the stocks of human and physical capital grow at the same (non-degenerate) rate $\nu_C = \nu_K = \nu_H > 0$, and that relative prices of goods be constant. The relative price of the investment good is constant by assumption. Constancy of the relative prices of human to physical capital (p_Y/p_X) requires that the costate variables must also grow at a common rate ($\nu_\mu = \nu_\lambda$). In this section we show that the model has a block recursive structure. The solutions for w, r, and p_Y can be obtained from the zero profit conditions for the respective production sectors and from an intertemporal arbitrage condition that must hold on the balanced growth path. We identify two types of balanced growth equilibria: an equilibrium with incomplete specialization in which all three goods are produced and two equilibria with production specialization in the non-traded good and one of the traded goods. We begin by showing that domestic prices (p_Y, w, r) and the pattern of specialization consistent with balanced growth are uniquely determined by the world relative price, p_X. We then use these prices to solve for the capital/labor ratio, growth rate, and consumption/wealth ratio on the BGP.

Utilizing (3.4)–(3.5), the requirement that $\nu_\mu = \nu_\lambda$ on the balanced growth path yields the intertemporal arbitrage (IA) condition

$$\frac{r}{p_X} - \frac{w}{p_Y} + \eta - \delta = 0. \tag{4}$$

This condition requires that the net return on investment in physical capital (rental rate less depreciation rate) equal the net return on investment in human capital. The factor market equilibrium conditions (3.2)–(3.3) require that factor prices be equalized across all operating sectors. Following Bond, Wang, and Yip (1996), we adopt a dual approach to the solution for goods and factor prices by utilizing the requirement that unit costs be no less than price in each sector, strict equality holding in sectors that are producing. Letting ϕ_i denote the cost function for sector i, these conditions can be expressed as

$$p_Y = \phi_Y(w, r), \tag{5.1}$$

$$p_X \leq \phi_X(w, r), \tag{5.2}$$

$$1 \leq \phi_Z(w, r). \tag{5.3}$$

Given that the education good is non-traded, an immediate implication of the assumption of balanced growth is that the economy may not shut down the production of the Y sector. In terms of (3.2)–(3.3) this implies that the non-negativity conditions associated with s_Y and u_Y never bind (i.e. $\beta_Y = \alpha_Y = 0$). Although we must have positive production of the education good on any possible balanced growth path, it is possible for the economy to shut down production of one of the traded goods sectors.

(4)–(5.3) are four equations to determine the prices w, r, and p_Y consistent with balanced growth given the value of p_X. The production structure in (5.1)–(5.3) has the characteristic that there are as many traded goods as there are factors of production. In static trade models this production structure would typically generate a range of factor endowments consistent with production of all three goods for an exogenously given p_X [e.g. Komiya (1967)]. This result for the static case follows from the fact that equations (5.2)–(5.3) can be solved for (w, r) given p_X. These domestic factor costs then determine the price of the non-traded good from its zero profit condition (5.1). However, in the dynamic model the intertemporal arbitrage condition links the price of the non-traded capital good to that of the traded capital good. This additional restriction reduces the likelihood that a given p_X is consistent with production of all goods on the BGP.

We now show that there can exist at most one world price at which all of the equations in (4)–(5.3) are satisfied with strict equality. The following condition will be imposed on the technologies, which assures the existence of a solution.

Condition FP (Factor Price). *Let* $\Omega_{XYZ} = \{(w, r, p_X, p_Y) \mid p_i = \phi_i(w, r) \text{ for } i \in \{X, Y, Z\}\}$. *Then*

$$\sup_{\Omega_{XYZ}} \left(\frac{r}{p_X} - \frac{w}{p_Y} \right) > \delta - \eta > \inf_{\Omega_{XYZ}} \left(\frac{r}{p_X} - \frac{w}{p_Y} \right).$$

If this condition fails, the technology for producing one of the factors is so inefficient that the net return to that factor (at constant prices) is always dominated by that of the other factor.[6]

Condition FP can be used to establish:

Proposition 1. *If Condition FP holds, there exists a unique world price p_X^* and corresponding domestic prices (w^*, r^*, p_Y^*) and sectoral factor intensities k_i^* $(i \in \{X, Y, Z\})$ at which the zero profit conditions (5.1)–(5.3) are satisfied with strict equality for all three sectors and the intertemporal arbitrage condition (4) is satisfied.*

Proof. The zero profit conditions yield three equations to solve for the prices (p_X, p_Y, w, r). These equations can be inverted to solve for the w, p_X, and p_Y as functions of r. Totally differentiating the system (5.1)–(5.3) yields

$$\hat{w}|_{XYZ} = -\frac{1 - \theta_{HZ}}{\theta_{HZ}} \hat{r}, \quad \hat{p}_i|_{XYZ} = \frac{\theta_{HZ} - \theta_{Hi}}{\theta_{HZ}} \hat{r}, \quad i = X, Y, \tag{6}$$

where a hat over a variable denotes a rate of change, θ_{Hi} is the share of labor costs in unit costs of good i, and XYZ denotes that the comparative statics exercises are performed with all three sectors producing. Utilizing (6), it follows that $r/p_X(r)$ is a continuous and increasing function of r

6. This condition is analogous to the one required to prove the existence of a BGP in the two sector, closed economy endogenous growth model in Bond, Wang, and Yip (1996).

and $w(r)/p_Y(r)$ is a continuous and decreasing function of r. Therefore, there can be at most one value of r, denoted r^*, at which (4) holds. Condition FP ensures the existence of such a value, and $p_X^* = p_X(r^*)$ is the unique world price consistent with balanced growth and incomplete specialization. □

Note that Proposition 1 holds even in the presence of factor intensity reversals. Factor intensity reversals raise the possibility that the relations $p_i(r)$ are not monotonic. However, this possibility does not alter the fact that $[r/p_X(r)] - [w(r)/p_Y(r)]$ is an increasing function of r, which yields the uniqueness result.

We next examine whether there exist domestic prices consistent with a BGP in which one of the traded goods sectors is shut down when $p_X \neq p_X^*$.

Proposition 2. *(i) If $p_X > p_X^*$ there exist unique prices (r, w, p_Y) satisfying (4)–(5.3) in which only goods X and Y are produced. These prices have the property $r(p_X)/p_X = r^*/p_X^*$ and $w(p_X)/p_X = w^*/p_X^*$, with sectoral factor intensities constant at $k_i = k_i^*$ ($i \in \{X, Y\}$) in all these equilibria.*

(ii) Let $\Omega_{XZ} = \{(w, r, p_Y) \mid p_i = \phi_i(w, r)$ for $i \in \{Y, Z\}\}$ and $p_X^{\min} \equiv \min_{\Omega_{YZ}} r \left(\dfrac{w}{p_Y} + \delta - \eta\right)^{-1} \geq 0$. If $p_X \in [p_X^{\min}, p_X^*]$, there exist unique prices $r(p_X)$, $w(p_X)$, and $p_Y(p_X)$ consistent with (4)–(5.3) in which only goods Y and Z are produced. $r(p_X)/p_X$ is decreasing in p_X for these equilibria, and sectoral factor intensities $k_i(p_X)$ are non-increasing in p_X ($i \in \{Y, Z\}$).

Proof. We begin by solving for the values of (w, r, p_Y, p_X) consistent with balanced growth when the economy produces goods X and Y only. If factor proportions differ across sectors X and Y, (5.1)–(5.2) can be inverted to obtain expressions $p_Y = p_Y(p_X, r)$ and $w = w(p_X, r)$ which have the properties

$$\begin{aligned} \hat{w}\,|_{XY} &= \frac{1}{\theta_{HX}}\,\hat{p}_X - \frac{1-\theta_{HX}}{\theta_{HX}}\,\hat{r}, \\ \hat{p}_Y\,|_{XY} &= \frac{\theta_{HY}}{\theta_{HX}}\,\hat{p}_X + \frac{\theta_{HX}-\theta_{HY}}{\theta_{HX}}\,\hat{r}. \end{aligned} \tag{7}$$

(7) can be used to show that w/p_Y is a decreasing function of r/p_X when goods X and Y are produced. Thus, there can be at most one value of r/p_X at which (4) is satisfied. Since the prices (p_X^*, p_Y^*, w^*, r^*) from Proposition 1 satisfy (4)–(5.2) with strict equality, it follows that $(ap_x^*, ap_Y^*, aw^*, ar^*)$ satisfies (4)–(5.2) with strict equality for any $a > 0$.

The values of w and r/p_X for this case are illustrated by the line DBE in $(r/p_X, p_X)$ space and the ray DBE in (w, p_X) space in figure 1. In order for the factor prices on these loci to be equilibria in which X and Y are the only goods produced it must also be the case that the Z sector is not earning positive profits (i.e. (5.3) is satisfied). Since w and r are increasing in p_X, the unit cost of good Z is increasing in p_X along the DBE locus. Since the production of Z earns zero profits at p_X^*, it follows that $\phi_Z(w,r) > (<)\ 1$ for $p_X > (<)\ p_X^*$. Therefore, the Z sector is unprofitable when $p_X > p_X^*$, and the factor prices given by the segment BE in figure 1 are consistent with (4)–(5.3) being satisfied with production of X and Y. This establishes (i) of the proposition.

Next, we solve for the values of (w, r, p_Y, p_X) consistent with balanced growth when only goods Y and Z are produced. (5.3) and (5.1) must hold with strict equality in this case, and these conditions can be inverted to yield $w = w(r)$ and $p_Y = p_Y(r)$. Differentiation of these conditions implies

$$\hat{w}|_{YZ} = -\frac{1-\theta_{HZ}}{\theta_{HZ}}\,\hat{r}, \quad \hat{p}_Y|_{YZ} = \frac{\theta_{HZ}-\theta_{HY}}{\theta_{HZ}}\,\hat{r}. \tag{8}$$

For a fixed p_X, this implies that $(r/p_X) - (w(r)/p_Y(r))$ is an increasing function of r, so that (4) has at most one solution for a given p_X. Clearly a solution will exist for $p_X = p_X^*$, since the prices (p_Y^*, w^*, r^*) from Proposition 1 must satisfy (4)–(5.1) and (5.3). For $p_X < p_X^*$, a solution will exist as long as $p_X \geq p_X^{\min}$. Differentiating (4) and substituting from (8) yields the effect of changes in p_X on equilibrium factor prices with Y and Z produced

$$\begin{aligned} \frac{\hat{r}-\hat{p}_X}{\hat{p}_X}\bigg|_{YZ,IA} &= -\frac{wp_X(1-\theta_{HY})}{rp_Y\theta_{HZ}+wp_X(1-\theta_{HY})} < 0, \\ \frac{\hat{w}}{\hat{p}_X}\bigg|_{YZ,IA} &= -\frac{rp_Y(1-\theta_{HZ})}{rp_Y\theta_{HZ}+wp_X(1-\theta_{HY})} < 0, \end{aligned} \tag{9}$$

where IA denotes the fact that the intertemporal arbitrage condition (4) is also assumed to hold. The locus of factor prices consistent with production of Y and Z is illustrated by the loci ABC in figure 1. Note also that since r/w is an increasing function of p_X from (9), the cost-minimizing factor proportions in sector i, $k_i(p_X)$, are non-increasing in p_X for $i \in \{Y, Z\}$.

In order for the factor prices on the ABC loci in figure 1 to be equilibria with production of Y and Z, we must also establish that the X sector does not earn positive profits. The effect of an increase in the price of good X on the cost of production of good X is $\hat{\phi}_X/\hat{p}_X =$

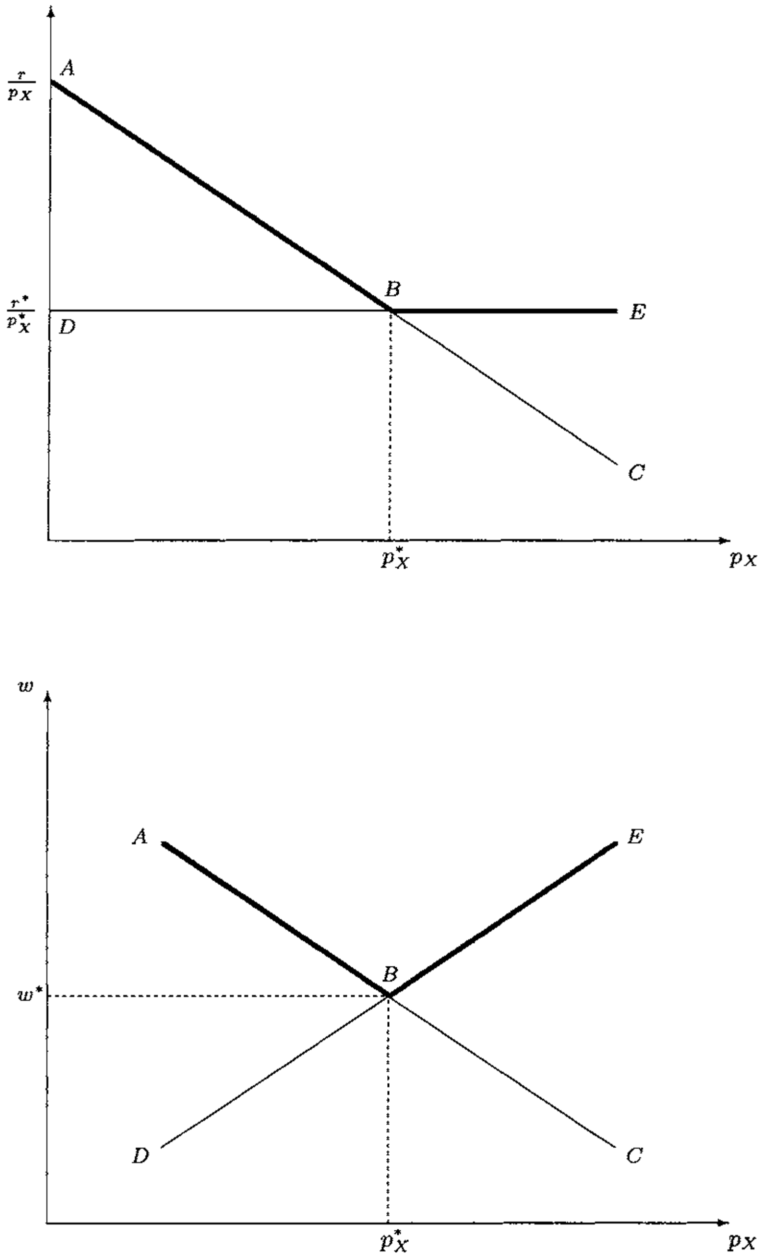


Fig. 1. Factor prices satisfying intertemporal arbitrage and zero profits: Efficient specialization in Y and Z on segment AB and specialization in X and Y on segment BE

$\theta_{HX}(\hat{w}/\hat{p}_X)+(1-\theta_{HX})(\hat{r}/\hat{p}_X)$. Substituting from (9) into this expression yields $\hat{\phi}_X - \hat{p}_X < 0$, so that the cost of producing X is rising more slowly than the price of good X along the ABC locus. Since $\phi_X(w^*, r^*) = p_X^*$, it follows that $\phi_X > (<)\, p_X$ for $p_X < p_X^*$. Therefore, the X sector is unprofitable for $p_X < p_X^*$ and the factor prices on the segments AB in figure 1 satisfy (4)–(5.3) with production of goods Y and Z. □

Propositions 1–2 establish the existence of a unique set of constant domestic prices satisfying the necessary conditions (3.2)–(3-5) for the optimization problem (P) given condition FP. In order to complete the proof of the existence of a unique BGP, we must show that there is a common non-degenerate growth rate ν and constant values of $c \equiv C/H > 0$ and $k \equiv K/H > 0$ that satisfy the remaining necessary conditions and the constraints (2.1)–(2.2). The following result provides sufficient conditions on the technology for the existence of a unique balanced growth path (given p_X) with non-degenerate growth:

Proposition 3. *If $r^*/p_X^* - \delta - \rho > 0$ and the maximal growth condition $\rho > (1-\sigma)\nu(p_X^{\min})$ is satisfied, then the small open economy will exhibit non-degenerate growth for $p_X \in [p_X^{\min}, \infty)$ with the growth rate given by*

$$\nu(p_X) = \frac{1}{\sigma}\left(\frac{r(p_X)}{p_X} - \delta - \rho\right). \tag{10}$$

(i) For $p_X < (>)\, p_X^*$ the country produces goods Y and Z (Y and X). The capital/labor ratio on the balanced growth path is unique.

(ii) For $p_X = p_X^*$ the country can produce all three goods. The capital/labor ratio on the BGP may take any value in the interval $[u_Y^* k_Y^* + \min_{i\in\{X,Z\}}(1-u_Y^*)k_i^*,\ u_Y^* k_Y^* + \max_{i\in\{X,Z\}}(1-u_Y^*)k_i^*]$.

Proof. The consumption growth rate on the BGP can be obtained by differentiating (3.1) with respect to t and substituting from (3.4), which yields (10). As shown in figure 1, r/p_X is lowest at p_X^* so a sufficient condition for non-degenerate growth is that $\nu(p_X^*) > 0$. The upper bound on the feasible growth rate is required to guarantee that the transversality conditions are satisfied on the balanced growth path. The growth rate of $e^{-\rho t}\mu(t)K(t)$ will be $\nu(1-\sigma) - \rho$, which must be negative to satisfy (3.6). This will be satisfied at all p_X if it is satisfied at the maximal growth rate, $\nu(p_X^{\min})$. A similar argument shows that (3.7) is satisfied.

It remains to be shown that the growth rate defined in (10) satisfies (2.1)–(2.2). From (2.2), we have

$$u_Y(p_X) = \frac{\nu(p_X) + \eta}{g(k_Y(p_X))}. \tag{11}$$

Feasibility requires $0 < u_Y < 1$. $u_Y > 0$ follows immediately from the non-degenerate growth result. To establish $u_Y < 1$, note that from the competitive profit condition we have $p_Y g = w + rk_Y$. Therefore, a sufficient condition for $u_Y < 1$ (using (11)) is $(w/p_Y) - \eta = \rho + \sigma\nu > \nu$, which is guaranteed by the maximal growth condition. For $p_X \neq p_X^*$, only one of the traded goods will be produced and the solution for k is obtained from the full employment condition to be

$$k(p_X) = u_Y(p_X)k_Y(p_X) + (1 - u_Y(p_X))k_i(p_X), \tag{12.1}$$

where $i = X$ (Z) if $p_X > (<)\, p_X^*$.

For $p_X = p_X^*$, all three sectors are potentially operational so any capital stock consistent with

$$k = u_X k_X(p_X^*) + (1 - u_X - u_Y(p_X^*))k_Z(p_X^*) + u_Y(p_X^*)k_Y(p_X^*), \tag{12.2}$$
$$u_X \in [0,\ 1 - u_Y(p_X^*)],$$

is an equilibrium.

It remains to show that $c \geq 0$. The budget constraint for the small open economy can be written as $c + p_X(\nu + \delta)k + p_Y(\nu + \eta) = w + rk$. Utilizing (4) and (10) and rearranging terms, we can solve for the consumption/wealth ratio on the balanced growth path

$$\frac{c}{p_Y + p_X k} = \rho + (\sigma - 1)\nu. \tag{13}$$

$c > 0$ then follows immediately from the maximal growth condition. $\square$

It is shown in Bond, Trask, and Wang (1997) that in the closed economy case, there will be a unique BGP with prices given by the values (w^*, r^*, p_Y^*, p_X^*) from Proposition 1. The capital/labor ratio on the BGP in the closed economy case, denoted k^*, will be uniquely determined from the full employment conditions using $u_Y^* = (\nu + \eta)/g(k_Y^*)$ and $u_Z = c^*/j(k_Z^*)$. k^* must lie in the interior of the interval identified in Proposition 3(ii). The growth rate is positively related to the real return to the investment good by (10), so it follows from figure 1 that the economy's growth rate is lowest at autarky. Free trade will lead to a BGP growth rate that is no lower than that under autarky, and it will be strictly higher if the country exports the consumption good. Since the growth rate is determined by the productivity of investment goods, a country which has a comparative disadvantage in investment goods will raise its growth rate by being able to import more productive investment goods.

2.3 Transitional Dynamics

We conclude our characterization of the BGP for the small open economy by examining the transitional dynamics in the neighborhood of the BGP. First consider the case in which $p_X = p_X^*$. Proposition 3(ii) shows that in this case the economy will be on the balanced growth path for any $k \in [u_Y^* k_Y^* + \min_{i \in \{X,Z\}} (1-u_Y^*)k_i^*,\ u_Y^* k_Y^* + \max_{i \in \{X,Z\}} (1-u_Y^*)k_i^*]$, so that there will be no transitional dynamics for any initial endowment in this range. This result is substantially different from the closed economy version of the model, where there is a unique k associated with balanced growth. In the closed economy case, an initial value of k below (above) the BGP value will result in a gradual transition to the BGP in which physical capital is accumulated more rapidly (slowly) than human capital. In contrast, there is no need for adjustment of the relative factor stocks in the open economy case because the difference in relative factor supplies can be met by an increase in the production of the good that uses the abundant factor intensively. The open economy is able to compensate for the scarce factor by importing more of the good that uses the scarce factor intensively. One can think of domestic and foreign factor services as being perfect substitutes in this case.

For $p_X \neq p_X^*$, the economy will be in the position of producing only one of the traded goods in addition to the non-traded good. Domestic and foreign factor services are not perfect substitutes in this case because domestic factors are not being used to produce the import-competing good. This limits the ability of the economy to alter the composition of output in response to imbalances in factor stocks. We will show that the transitional dynamics in the case of specialization in one traded good are quite similar to those for the closed economy case. We begin by showing that the dynamics of the system in either of the specialization cases can be expressed in terms of (r, c, k), and then use this system to prove that the economy exhibits saddle path stability for all sectoral factor intensity rankings in the neighborhood of the BGP.

During the transition to the BGP, the relative price of human capital may be changing because accumulation rates of the factors are not necessarily constant. The intertemporal arbitrage condition equating the returns to human capital obtained from (3.4)–(3.5) will be

$$\nu_\lambda - \nu_\mu = \frac{\dot{p}_Y}{p_Y} = \frac{r}{p_X} - \frac{w}{p_Y} + \eta - \delta. \tag{14}$$

When capital gains exist on the transition path, the equalization of returns to physical and human capital requires that the difference in net returns between physical and human capital equal the rate of capital gain on human capital. Therefore, factor prices must satisfy (5.1)–(5.3) and (14) along the transition path.

Figure 2 can be used to illustrate the feasible values for factor prices during the transition process for the case $\theta_{HX} > \theta_{HZ}$. The ABC locus in (w, r) space is the locus of factor prices consistent with zero profits in the Y and Z sectors, while the DBE locus is the locus giving zero profits in the X and Y sectors. Using (7)–(8), it can be seen that the assumption $\theta_{HX} > \theta_{HZ}$ ensures that the DBE locus is flatter than ABC at the intersection point B. Using (7), it is straightforward to show that since $\theta_{HX} > \theta_{HZ}$, the unit cost of the Z sector is $\phi_Z(w, r)$ is increasing in r along the DBE locus consistent with zero profits in $\{Y, X\}$. Therefore, the Z sector earns positive profits for $r < r^B$ on the DBE locus, so the production specialization must be $\{Y, Z\}$ for $r < r^B$. A similar argument can be used to show that the unit cost of the X sector is decreasing in r along the ABC locus, so that the X sector earns positive profits for $r > r^B$ on the segment ABC locus. The production specialization is $\{X, Y\}$ for $r > r^B$. Note that this argument depends only on the relative factor intensity of the two traded goods sectors, and not on their identity. Therefore, in the absence of factor intensity reversals there will be a critical value of r such that the economy will be completely specialized in the capital intensive (labor intensive) traded good for r less (greater) than the critical value.[7]

Assuming that good Y is produced throughout the transition process, domestic factor prices must lie on the ABE locus in figure 2 during the transition process. Let (w^0, r^0) denote the point on this frontier which is also consistent with intertemporal arbitrage at constant p_Y. This point must be unique as a result of Propositions 1–2. For factor prices in the neighborhood of this point, the economy must have the same specialization pattern as on the BGP. The evolution of factor prices will be given by (14), with the domestic prices determined by (7). Using (7) in (14) yields

$$\nu_r|_{XY} = \frac{\theta_{HX}}{\theta_{HX} - \theta_{HY}} \left(\frac{r}{p_X} - \frac{w(r)}{p_Y(r)} + \eta - \delta \right). \tag{15}$$

(15) indicates that the dynamics of local prices are a function of r alone. Since the expression in parentheses must be an increasing function of r (see the discussion following (7)), we have $\partial \nu_r / \partial r < 0$ iff $\theta_{HX} < \theta_{HY}$ in

7. With factor intensity reversals between the traded goods, the ABC and DBE loci may have multiple intersections. There may then be multiple switches of the specialization pattern as r increases in this case. However, we will always have the conclusion that there is specialization in the labor (capital) intensive good for values of r above (below) an intersection point.

the neighborhood of $\nu_r = 0$. The rental adjustment process is stable iff the Y sector is labor-intensive relative to the X sector. Suppose that the return to physical capital exceeds that to human capital, which requires a capital gain on human capital investments (i.e., $\dot{p}_Y > 0$). In order for the rental process to be stable, the increase in p_Y must reduce r. By the Stolper-Samuelson theorem, an increase in p_Y will reduce r iff the Y sector is human capital intensive.

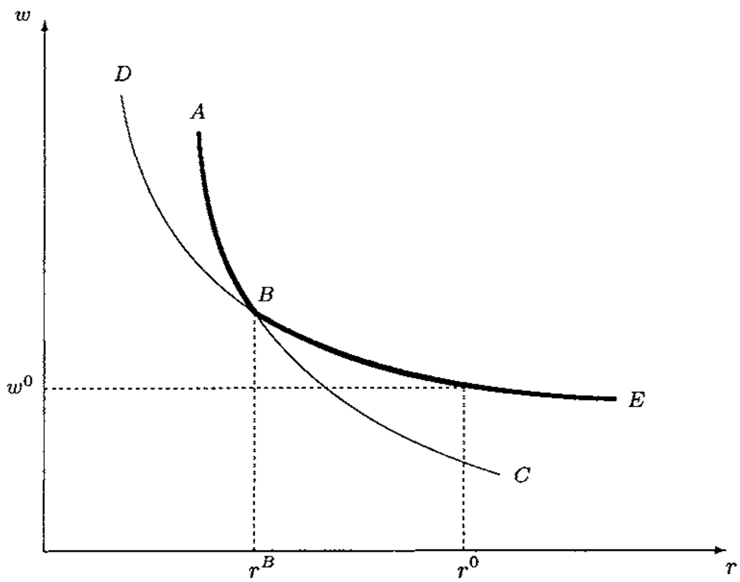


Fig. 2. Price frontiers consistent with production of X and Y (segment BE) and production of Y and Z (segment AB) for given p_X

Note that with specialization in Y and Z on the BGP, this argument is identical with factor prices being determined by (8) in the neighborhood of their BGP values. Since the two cases are so similar, we will present the analysis for the case where the country is specialized in X and Y. The results for specialization in Y and Z follow in a similar manner. With specialization in X and Y, full employment requires $k = (1-u_Y)k_X(r) + u_Y k_Y(r)$, where the sectoral factor intensities $k_i(r)$ are determined by cost minimization given $(w(r), r)$ from (7). The full employment condition can be used to solve for $u_Y = u_Y(r, k)$. Defining $x(r,k) = X/H = u_X(r,k)f(k_X(r))$ and $y(r,k) = Y/H =$

$u_Y(r,k)g(k_Y(r))$, (2.1)–(2.2) can be used to express the evolution of c and k as

$$\frac{\dot{c}}{c} = \frac{\dot{C}}{C} - \frac{\dot{H}}{H} = \nu_C(r) - y(r,c,k) + \eta, \tag{16}$$

$$\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{H}}{H} = \frac{p_X\, x(r,c,k) - c}{p_X k} - y(r,c,k) + \eta - \delta. \tag{17}$$

The system (15)–(17) describes the dynamics of the system along the transition path.

The following result is proven in the Appendix:

Proposition 4. *Suppose that the economy specializes in one traded good, $j \in \{X, Z\}$ and the non-traded good Y on the BGP. This BGP will exhibit saddle path stability for all factor intensity rankings.*

(i) If $k_Y > k_j$, then the relative price of physical to human capital, p_X/p_Y, will be constant along the transition path.

(ii) If $k_j > k_Y$, p_X/p_Y will be a non-increasing function of k along the transition path.

The existence of saddle path stability for all sectoral factor intensity rankings in the open economy case is similar to the result obtained in Bond, Wang, and Yip (1996) for the closed economy version of a two sector endogenous growth model in which there is a unified consumption/investment good sector and an education sector. The main role of the sectoral factor intensity rankings is in the behavior of the relative prices of the capital goods along the transition path. This results from the role of factor intensities on factor price and output adjustments as reflected in the Stolper-Samuelson and Rybczynski theorems of international trade theory. If the education sector is human capital intensive, the price adjustment process in (15) is stable as noted above. However, if the education sector is physical capital intensive, the price adjustment process is unstable and the price must jump to its BGP value.

The jump of the price to its BGP value is consistent with the transitional adjustment in the case where $k_Y > k_j$ because the quantity adjustment process is stable and allows k to adjust to its BGP value at fixed prices. By the Rybczynski theorem, the output of the capital intensive sector will be higher than its BGP value when k exceeds its BGP value. If the education sector is capital intensive, this adjustment in outputs results in a fall in k and the economy converges to the BGP value. In contrast, when $k_j > k_Y$, the output of physical capital is higher when k exceeds its BGP value, causing k to diverge. In this case the relative price of capital goods must be adjusting along the transition path in order to obtain convergence to the BGP value of k.[8] For each factor intensity ranking, instability in one of the adjustment processes is offset by stability in the other process to obtain saddle path adjustment. The saddle path stability in the specialization equilibrium results from the same "polarization" of the price and quantity adjustment processes exhibited by the two sector model. Note that the forward looking behavior of agents, which underlies the intertemporal arbitrage equation is critical to this stability result.[9]

Proposition 4 characterizes transitional dynamics for $p_X \neq p_X^*$. In $p_X = p_X^*$, then the BGP factor prices are given by point B in figure 2. If the initial factor endowment lies outside the range identified in Proposition 3(ii), then full employment is not possible with production of all three goods at the factor prices (r^B, w^B). Therefore, the convergence to the range of factor endowments consistent with balanced growth must occur along one of the branches in which the economy is specialized in one of the traded goods.

3. Balanced Growth Path Effects of Price and Technology Changes

In this section we use the results of Propositions 1–3 to characterize the effect of changes in the world price on the values of ν, k, and c on the BGP. These results indicate how changes in the terms of trade affect the growth rate and welfare levels of a small open economy on the BGP. We also examine how technological change affects the rate of growth and the critical value, p_X^*, at which the country is incompletely specialized. An increase in p_X^* can be interpreted as a shift in comparative advantage toward the consumption good, since it expands the range of prices for which the country exports the consumption good.

---

8. It is straightforward to show that the value function for this problem, $V(K, H, p_X)$, is homogeneous of degree $1-\sigma$ in K and H. Since the costate variables to the representative agent optimization problem (P) are equal to the derivative of the value function with respect to the appropriate state variable, it follows that $p_X/p_Y = V_K/V_H = \phi(k)$. The concavity of V in K and H ensures $\phi' \leq 0$. An implication of this result is that the two capital goods must be perfect substitutes in the case where $k_Y < k_j$, in the sense that the isoquants for the value function must have flat segments.

9. In the case of proportional savings, it is well known that instability can result in two sector models with endogenous capital accumulation and exogenous labor force growth when the investment good sector is capital intensive. See, for example, Inada (1963).

3.1 Terms of Trade Changes

For the case in which $p_X > p_X^*$, the analysis is simplified due to the fact that r/p_X and the sectoral factor intensities are independent of p_X. With a constant r/p_X, the growth rate is independent of p_X from (10) as is the share of labor allocated to human capital production. An increase in p_X has no effect on the production side of the economy, but it does increase the purchasing power of the small open economy because it increases the price of the exportable good. Utilizing (13), it can be seen that the effect of the increase in p_X is a proportional increase in c. Clearly, this increase in the price of investment goods will raise the BGP welfare level for a small open economy that is exporting investment goods.

For $p_X < p_X^*$, r/p_X is a decreasing function of p_X as illustrated in figure 1. When the small open economy is specialized in consumption goods, an increase in the price of capital goods reduces the return to investment (r/p_X) in equilibrium and lowers the growth rate. The increase in p_X will also reduce the sectoral capital/labor ratios, as established in Proposition 2(ii). Differentiating (11), it can be seen that these two effects of an increase in p_X have a conflicting impact on u_Y

$$\frac{du_Y}{dp_X} = \frac{\nu'(p_X)}{g} - (1 - \theta_{HY})u_Y \frac{k_Y'(p_X)}{k_Y}. \tag{18}$$

The first term is negative, because a lower growth rate reduces the amount of labor required to produce human capital. The second term tends to raise u_Y, because the declining capital/labor ratio results in a greater requirement of human capital per unit of Y produced.

The effect of a change in p_X on k is given by $dk/dp_X = [u_Y k_Y' + u_Z k_Z'] + (k_Y - k_Z)u_Y'$. The term in brackets is negative, because the rising r/w causes substitution away from capital in both sectors. This substitution will reduce the relative usage of physical capital on the BGP at a fixed u_Y. The second term reflects the reallocation of labor between sectors at given factor proportions, which will tend to raise k if the reallocation is toward the capital intensive sector (i.e. $k_Y > k_Z$ and $u_Y' > 0$ or $k_Z > k_Y$ and $u_Y' < 0$). Substituting into this expression from (18) yields

$$\frac{dk}{dp_X} = u_Y\left(\theta_{HY} + \frac{k_Z g'}{g}\right)k_Y' + u_Z k_Z' + \frac{(k_Y - k_Z)\nu'}{g}. \tag{19}$$

The first two terms in (19) represent the effect on demand for capital of sectoral substitution effects at a fixed ν. These two terms must be negative. Since $\nu'(p_X) < 0$, a sufficient condition for an increase in the cost of investment goods to reduce k is $k_Y > k_Z$. This yields the intuitive conclusion that an increase in the price of investment goods results in a lower use of physical (relative to human) capital on the BGP. If $k_Y < k_Z$, however, it is possible that the BGP capital/labor ratio is increasing in p_X if the substitution effects generated by the rise in r/w are small enough. For example, in the limiting case of fixed coefficients production processes in the Y and Z sectors, the first two terms in (19) will be zero and an increase in p_X must raise the capital labor ratio when $k_Y < k_Z$. An increase in the price of investment goods lowers the growth rate, which results in a shift of resources from the Y sector to the capital intensive Z sector. This raises the relative usage of physical capital on the BGP. The above discussion on the relationship between k and p_X is summarized in figure 3. When the world price is p_X^*, any k in the range identified in Proposition 3(ii) is consistent with balanced growth. For $p_X > p_X^*$, the capital/labor ratio will equal the value associated with the lower (upper) end of this range when $k_X^* < k_Z^*$ ($k_X^* > k_Z^*$). For $p_X < p_X^*$, the locus $ABCD$ indicates the case in which the substitution effects dominate, while the locus $EBCD$ will arise if $k_Z > k_Y$ and the sectoral reallocation effects dominate the substitution effects.

When the country is specialized in Z, consumption per effective labor unit is the difference between output of Z and the demand for imported investment goods. Differentiating (2.1), using $u_X = 0$, gives

$$\begin{aligned} c'(p_X) = \; & [u_Z r k_Z'(p_X) + j u_Y'(p_X)] \\ & -[k(\nu+\delta) + p_X(\nu+\delta)k'(p_X) + p_X k \nu'(p_X)]. \end{aligned} \tag{20}$$

The term in the first bracket is the impact of a change in p_X on output of Z per effective labor unit. There are two effects: substitution away from capital in production will reduce output of Z at a given labor allocation, while the reallocation of labor may either raise or lower output of Z. The term in the second bracket is the effect of changes in the demand for investment goods. The increase in the price of imported goods will tend to decrease consumption at a given level of investment demand, while a decrease in the growth rate will tend to increase consumption. Finally, an increase (decrease) in the capital/labor ratio will lower (raise) the demand for the consumption good. Overall, the sign of $c'(p_X)$ is indeterminate. Although a reduction in consumption is more likely when $k_Y > k_Z$ [and hence $k'(p_X) < 0$], this condition is not sufficient to guarantee a decline in consumption.

Although the effect of an increase in p_X on c is ambiguous when the country is importing capital goods, the fact that the growth rate is declining suggests that it may be possible to derive results on the change in welfare on the BGP. The welfare level of the representative agent on

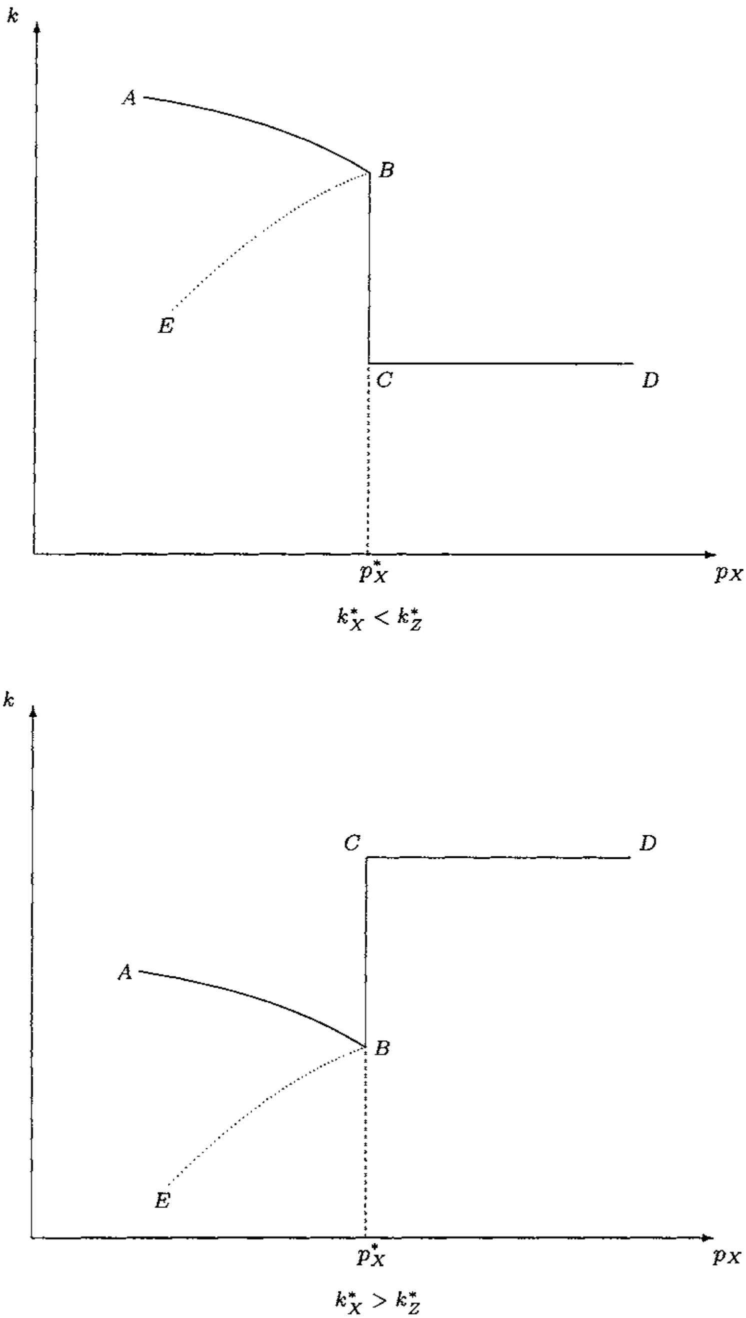


Fig. 3. Relationship between capital/labor ratios and p_X on the BGP

the BGP may be found from (P) as

$$V(p_X) = \frac{c(p_X)^{1-\sigma} H_0^{1-\sigma}}{(\rho + (\sigma - 1)\nu(p_X))(1-\sigma)}. \tag{21}$$

Differentiating (21) with respect to p_X and using (18)–(20) yields

$$\begin{aligned} \frac{V'(p_X)}{V} = & \left(u_Z k_Z'(p_X) + \frac{u_Y j}{p_Y g} k_Y'(p_X) + \frac{k_Y - k_Z}{g} \nu'(p_X) \right) \\ & \cdot \frac{p_X(\rho + (\sigma - 1)\nu)}{c} - \frac{k(\nu + \delta)}{c}. \end{aligned} \tag{22}$$

The first two terms in parentheses are negative, since $k_i'(p_X) < 0$ as established above. The last term (outside parentheses) must also be negative. Therefore, a sufficient condition for an increase in p_X to unambiguously reduce V is $k_Y > k_Z$.

The results of this section can be summarized in the following result:

Proposition 5. *The effect of a change in the terms of trade on the BGP values of ν, c, k, and V depend on the pattern of specialization.*

(i) When $p_X > p_X^*$: $\nu'(p_X) = 0$, $k'(p_X) = 0$, $c'(p_X) > 0$, $V'(p_X) > 0$.

(ii) When $p_X < p_X^*$: $\nu'(p_X) \leq 0$, $c'(p_X) \gtreqless 0$. If $k_Y > k_Z$, then $k'(p_X) < 0$, $V'(p_X) < 0$.

In static trade models, an increase in the price of the importable good is associated with a decline in welfare. Proposition 5 establishes that BGP welfare must be decreasing in the price of the importable when the country imports the consumption good, and when the country imports the investment good and $k_Y > k_Z$. Thus, when $k_Y > k_Z$ welfare is always higher on the BGP with trade than it is at the autarky BGP, which is associated with the price, p_X^*. The possibility that an increase in p_X may reduce BGP welfare in the case of $k_Y < k_Z$ arises from the possibility that k increases. Note however that even if the deterioration in the terms of trade raises the BGP utility level, it may still lower welfare sufficiently during the transition period so that the overall effect on welfare is negative.

It is also useful to compare the impact of changes in the terms of trade on factor incomes to those of the Stolper-Samuelson theorem for static trade models. The Stolper-Samuelson result would suggest that the interest of factor owners are strongly opposed: a change in the terms of trade will make owners of one factor unambiguously better off and owners of the other factor unambiguously worse off. In the present model the result is quite different, because the long run returns to the two factors are tied together by the intertemporal arbitrage condition. Any changes in the real return to physical capital (r/p_X) are matched by equivalent changes in (w/p_Y). However, windfall gains or losses may be experienced by factor owners who have accumulated stocks prior to the price change. For example, consider the effect of an increase in p_X when X and Y are being produced. Since the aggregate capital/labor ratio is independent of p_X in this region, there will be no transitional dynamics in k. By Proposition 4(i), local factor prices will increase proportionally with p_X. Since the aggregate k is independent of p_X in this case the factor prices will jump immediately to the new BGP values. Owners of existing physical and human capital will be made better off by the increase, since their purchasing power in terms of the consumption good has increased. For the case in which Y and Z are being produced, (9) shows that an increase in p_X must result in a less than proportional increase in r and a decrease in w on the new BGP. If $k_Y > k_Z$, the results of Proposition 4(i) imply that domestic prices will jump immediately to the new BGP values. Owners of existing physical capital will be made better off and owners of existing human capital will be made worse off as a result of these changes. If $k_Y < k_Z$, p_Y will be changing along the transition path [Proposition 4 (ii)] as k adjusts. Since k may either rise or fall as a result of the change in p_X, r may be either increasing or decreasing along the transition path.

3.2 Technical Change

We now turn to the effect of changes in technology on growth rates and comparative advantage. We can analyze the effects of technical progress by considering how changes in technology affect the real rental on physical capital, r/p_X, for each of the cases in which the country produces one traded good and the non-traded good.

We model changes in technology by writing the unit cost function in sector i as $\phi_i(w, r, a_i)$, where a_i is a parameter reflecting the level of the technology, $\partial\phi_i/\partial a_i < 0$. For the case in which the country produces good Y and traded good i $(i \in \{X, Z\})$, the prices on the BGP are determined by the intertemporal arbitrage condition (4) and the zero profit conditions

$$\phi_i\ (w, r, a_i\ ) = p_i, \tag{23.1}$$

$$\phi_Y(w, r, a_Y) = p_Y. \tag{23.2}$$

Totally differentiating these conditions and defining $\hat{b}_i = -(\partial\phi_i/\partial a_i)$ (da_i/ϕ_i) to be the rate of cost reduction in industry i from technical progress, we obtain

$$\frac{\hat{w}}{\hat{b}_Y} = -\frac{wp_X(1-\theta_{Hi})}{\Delta} < 0, \qquad \frac{\hat{r}}{\hat{b}_Y} = \frac{\theta_{Hi}wp_X}{\Delta} > 0,$$
$$0 < \frac{\hat{r}}{\hat{b}_i} = \frac{(1-\theta_{HY})wp_X}{\Delta} < \frac{\hat{w}}{\hat{b}_i} = \frac{wp_X(1-\theta_{HY}) + rp_Y}{\Delta}, \qquad (24)$$

where $\Delta \equiv \theta_{Hi}rp_Y + (1-\theta_{HY})wp_X > 0$.

Technical progress in the traded good will raise the returns to both physical and human capital, with the wage rate rising by relatively more. The direction of factor price changes is driven by the intertemporal arbitrage condition, which requires equal changes in r/p_X and w/p_Y, and the fact that p_Y is endogenous. The increase in both factor prices must raise p_Y, so w must rise by relatively more to maintain equal returns from investment in the two factors. Technical progress in the non-traded good will raise the return to physical capital, but reduce the return to human capital. When technological improvement occurs in the non-traded good, factor prices cannot both move in the same direction because the price of the traded good is constant. Since technical progress in Y reduces p_Y, we must have an increase in r and reduction in w to maintain equal increases in r/p_X and w/p_Y.

The rise in r/p_X resulting from technical progress must increase ν on the BGP in all cases by (10). The impact on comparative advantage can be seen by referring to figure 1. Technical progress in one of the traded goods will raise r/p_X on the segment in figure 1 associated with production of that good, which expands the range of p_X for which the country specializes in that traded good. Technical progress in the non-traded good will raise the r/p_X associated with both patterns of specialization, so whether p_X^* rises or falls will be determined by the relative increase in r/p_X in the two specializations. Using (24), it can be seen that p_X^* will fall with technical progress in Y iff $\theta_{HX} > \theta_{HZ}$. Technical progress in the education sector is associated with an increase in the range of prices for which the labor intensive good is exported.

Proposition 6. *Technological improvement in any of the goods will raise the growth rate at a given world price. The critical price at which the country will export the investment good, p_X^*, is an increasing function of the level of technology in the Z sector and decreasing in the level of technology in the X sector. It is increasing in the technology of the Y sector iff the X sector is labor intensive relative to Z.*

4. The Small Open Economy and the World Equilibrium

The above analysis has established that for the case of a small open economy accumulating physical and human capital and facing a constant world price, there is a unique world price consistent with incomplete specialization. For all other prices, the country will choose to specialize in one of the traded goods. One might be tempted to argue based on this result that incomplete specialization is a very unlikely event, since the probability of a randomly drawn world price being exactly equal to p_X^* would be equal to zero. However, this line of reasoning would be incorrect because it fails to take into account how world prices are determined.

For example, suppose that we make the Heckscher-Ohlin assumption of identical technologies across countries. It is shown in Bond, Trask, and Wang (1997) that the world economy with free trade must converge to a BGP with price p_X^*, so factor price equalization will hold across countries on the world BGP. Since countries are indifferent between producing the two traded goods at these factor prices, the factor accumulation pattern at the country level is indeterminate. The assumption of a constant world price p_X^* is thus consistent with a case where the world economy is on the BGP. Bond, Trask, and Wang (1997) also show that there are a continuum of balanced growth paths for the individual countries, as well as paths with unbalanced growth, which are consistent with the optimal accumulation of factors at p_X^* and with balanced growth for the world as a whole. Therefore, the long run trade pattern in these models is indeterminate.

This example highlights the importance of care in examining the relationships between assumptions made at the country level with those made regarding the time path of world prices. It also suggests that the pattern of trade in the long run in this model has a Ricardian flavor. If countries have identical technologies, the long run trade pattern is indeterminate and the long run world price is equal to that which would be the long run autarkic price of the individual countries. The gains from trade in this case occur during the transitional phase to the BGP, where initial endowment differences across countries cause autarkic prices to differ. The results on the effects of technological differences suggest that if one country has a technical advantage in one of the traded goods, that country will have a lower long run autarkic price for that good. There will then exist world prices at which each country specializes in the traded good in which it has comparative advantage, so the long run trade pattern will be determined by technologies.[10]

10. Note however that it is still important to establish under what conditions these

It should also be noted that the discount rate of consumers plays no role in the determination of comparative advantage in this model. This contrasts with results obtained by Findlay (1970) and Deardorff and Hansen (1978), who examine two sector exogenous growth models in which only physical capital is accumulated. In these models, a dynamic Heckscher-Ohlin theorem is obtained if there are differences in savings rates across countries, with the patient (i.e. high savings rate) country having a higher capital/labor ratio in the long run. In the model of this chapter, both of the primary factors are being accumulated in the long run. Changes in the rate of time preference will affect the absolute incentive to accumulate factors, as indicated by the fact that the growth rate in (10) is a decreasing function of the discount parameter. However, the rate of time preference does not alter the relative attractiveness of the two factors of production. This is reflected in the fact that the autarkic price, p_X^*, and the autarkic capital/labor ratio, k^*, are independent of the discount rate.

Appendix (Proof of Proposition 4)

For the case where the country produces only X and Y, the linearized dynamic system of (15)–(17) around the BGP is given by

$$\begin{bmatrix} \dot{r} \\ \dot{c} \\ \dot{k} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} r - r^* \\ c - c^* \\ k - k^* \end{bmatrix}, \tag{A.1}$$

where

$$a_{11} = \frac{1}{k_Y - k_X}\left(\frac{\partial(r/p_X)}{\partial r} - \frac{\partial(w/p_Y)}{\partial r}\right),$$
$$a_{23} = -\frac{cg}{k_Y - k_X}, \qquad a_{32} = -\frac{1}{p_X},$$
$$a_{33} = -\frac{gk + f}{k_Y - k_X} - (\nu^* + \delta).$$

This model has a block recursive structure, since the dynamics of the rental on capital (and hence p_Y/p_X) are independent of c and k. Thus, the system will have a real root $\gamma_1 = a_{11}$, with the remaining two roots satisfying the characteristic equation of a 2×2 subsystem (whose Jacobian matrix evaluated at the BGP is denoted as J_2^*):

$$\gamma^2 - \operatorname{tr}(J_2^*)\gamma + \det(J_2^*) = 0. \tag{A.2}$$

assumptions of technological differences at the country level are consistent with constant world price for the world economy. This remains an area for future work.

The trace and determinant of the 3×3 matrix in (A1) are

$$\det(J_2^*) = -a_{23}a_{32} = -\frac{cg}{p_X(k_Y - k_X)}, \qquad \operatorname{tr}(J_2^*) = a_{33}. \tag{A.3}$$

Using (A1) and (A3), we have the following two facts:

$$\begin{aligned} \operatorname{sign} a_{11} &= \operatorname{sign} k_Y - k_X, \\ \operatorname{sign} \det(J_2^*) &= \operatorname{sign} \operatorname{tr}(J_2^*) = \operatorname{sign} k_X - k_Y. \end{aligned} \tag{A.4}$$

The first is obtained from the definition of a_{11}, and the second follows immediately from (A3).

We can now use (A4) to establish that the system has one negative root and two positive roots, regardless of the factor intensity rankings, which yields saddle path stability of the dynamic system. First, consider the case where $k_Y > k_X$. One (real) root will be $a_{11} > 0$. The other two roots will be the solution to (A2) with $\operatorname{tr}(J_2^*) < 0$ and $\det(J_2^*) < 0$. Since the discriminant, $\operatorname{tr}(J_2^*)^2 - 4\det(J_2^*)$, is positive the solution to (A2) yields one positive real root and one negative real root. There will thus be two positive real roots and one negative real root and the system will exhibit saddle path stability with monotone transitional adjustment. Next consider the case where $k_X > k_Y$. One (real) root will be $a_{11} < 0$. The solution to (A3) with $\operatorname{tr}(J_2^*) > 0$ and $\det(J_2^*) > 0$ yields two roots with positive real parts. Again we obtain saddle path stability, but the transitional dynamics may be oscillating.

One difference between the two cases involves the adjustment of relative prices along the BGP. In the case where $k_Y > k_X$, we have $a_{11} > 0$ (with $a_{12} = a_{13} = 0$) which means that the adjustment process for r is unstable. An increase in r will raise the relative rental rate on investments in physical capital (i.e., r/p_X increases and w/p_Y decreases), which requires a capital gain on human capital to satisfy intertemporal no-arbitrage. However, when the human capital sector is capital intensive, a rise in p_Y/p_X over time requires $\dot{r}/r > 0$. Since the adjustment process for r is unstable, r (and hence p_X and p_Y) must jump to the corresponding BGP and remain constant along the saddle path.

When $k_X > k_Y$, on the other hand, $\dot{r}/r < 0$. An increase in r will reduce the growth rate of r in this case, which is a stabilizing force. The relative price ratio p_Y/p_X must be adjusting along the saddle path, because $\det(J_2^*) > 0$ and $\operatorname{tr}(J_2^*) > 0$ ensure that the adjustment process of c and k would be unstable at a fixed value of p_Y/p_X. □

Acknowledgement. We thank participants in the conference and the international workshop at Penn State for helpful comments on an earlier draft.

References

Barro, Robert J. (1991). "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics* 106, 407–443.

Bond, Eric W., Kathleen Trask, and Ping Wang (1997). "Factor Accumulation and Trade: A Dynamic Heckscher-Ohlin Model with Physical and Human Capital Accumulation," Penn State Working Paper.

Bond, Eric W., Ping Wang, and Chong Yip (1996). "A General Two-Sector Model of Endogenous Growth with Human and Physical Capital: Balanced Growth and Transitional Dynamics," *Journal of Economic Theory* 68, 149–173.

Caballé, Jordi, and Manuel S. Santos (1993). "On Endogenous Growth with Physical and Human Capital," *Journal of Political Economy* 101, 1042–1067.

Deardorff, Alan, and J. Hansen (1978). "Accumulation and a Long Run Heckscher-Ohlin Theorem," *Economic Inquiry* 16, 288–292.

Findlay, Ronald (1970). "Factor Proportions and Comparative Advantage in the Long Run," *Journal of Political Economy* 78, 27–34.

Grossman, Gene, and Elhanan Helpman (1991). *Innovation and Growth in the Global Economy*, Cambridge: MIT Press.

Inada, Kenichi (1963). "On a Two-Sector Model of Economic Growth: Comments and a Generalization," *Review of Economic Studies* 30, 119–127.

Jensen, Bjarne S., and Chunyan Wang (1997) "General Equilibrium Dynamics of Basic Trade Models for Growing Economies." (this volume).

Komiya, Ryutaro (1967). "Non-traded Goods and the Pure Theory of International Trade," *International Economic Review* 8, 132–152.

Lucas, Robert E. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22, 3–42.

— (1993). "Making a Miracle," *Econometrica* 61, 251–272.

Mulligan, Casey, and Xavier Sala-i-Martin (1993). "Transitional Dynamics in Two-Sector Models of Endogenous Growth," *Quarterly Journal of Economics* 108, 739–773.

Rebelo, Sergio (1991)."Long-run Policy Analysis and Long-run Growth," *Journal of Political Economy* 99, 500–521.

Rivera-Batiz, Luis A., and Paul M. Romer (1991). "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics* 106, 531–555.

Romer, Paul M. (1990). "Human Capital and Growth: Theory and Evidence," *Carnegie Rochester Conference Series on Public Policy* 32, 251–286.

Smith, Alasdair (1984). "Capital Theory and Trade." In R.W. Jones and P.B. Kenen (eds.), *Handbook of International Economics*, Amsterdam: North Holland.

Solow, Robert M. (1956). "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70, 65–94.

Stokey, Nancy, and Sergio Rebelo (1995) "Growth Effects of Flat-Rate Taxes," *Journal of Political Economy* 103, 519–550.

Tallman, Ellis W., and Ping Wang (1994). "Human Capital and Endogenous Growth: Evidence from Taiwan," *Journal of Monetary Economics* 34, 101–124.

Young, Alwyn (1995). "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quarterly Journal of Economics* 110, 641–680.

Part IV

Economic Growth, Technological Progress, and International Trade

CHAPTER 8

Efficiency and Productivity in Rich and Poor Countries

Rolf Färe and Shawna Grosskopf

1. Introduction

The motivation for our work is based in part on several recent articles in *The Economist* concerning growth and income.[1] Part of the discussion concerns new and old growth theory, including the debate over convergence. More specifically we are interested in looking at Mancur Olson's hypothesis that the persistence of low income in some poor countries may have to do with inefficiency rather than with their endowments of productive inputs and human capital.[2] In his own words,

> The argument offered here also fits the relationships between levels of per capita income and rates of growth better than does either the old growth theory or the new. As has often been pointed out, the absence of any general tendency for the poor countries with their opportunities for catch-up growth to grow faster than the rich countries argues against the old growth theory... The argument offered here suggests that poor countries on average have poorer economic policies and institutions than rich countries, and, therefore, in spite of their opportunity for rapid catch-up growth, they need not grow faster on average than the rich countries. But any poorer countries that adopt relatively good economic policies and institutions enjoy rapid catch-up growth: since they are far short of their potential, their per capita incomes can increase not only because of their technological and other advances that simultaneously bring growth to the richest countries, but also by narrowing the huge gap between their actual and potential income... (Olson, 1996, p. 20)

1. See, for example "Economic Growth: The Poor and the Rich", *The Economist*, May 25, 1996, pp. 23–25.

2. In fact, this general idea was first brought to our attention by Bob Parks from Washington University. He told us that his colleague, Douglass North argued that poor countries were poor because they have higher transactions costs than rich countries.

The chapter begins with a discussion of the static activity analysis model which we use to compute a measure of relative efficiency levels for the APEC countries in each of the years 1975–1990. Our results suggest a positive correlation between efficiency and per capita income in any given year.

Next we turn to a comparative static model. In our activity analysis framework, this yields another measure of performance which has turned out to play a central role in the renewed interest in growth theory, namely productivity change. Following Färe, Grosskopf, Lindgren, and Roos (1994), our activity analysis framework allows us to decompose productivity change into a "catching-up" and technical change component. This allows us to look at the relationship between imitation and innovation and per capita income: we compute productivity change for the APEC countries over the 1975–1990 period and relate this to per capita income.

Finally we turn to specification of a dynamic activity analysis model, in order to provide a dynamic measure of efficiency. This is based on the idea of a network model, in the spirit of Shephard and Färe (1980). The idea is to allow for intermediate outputs (in our case investment) that link adjacent periods. One of the goals of this exercise is to provide an estimate of the loss in potential output due to dynamic misallocation of resources. This is accomplished by including investment as endogenous in the model. We use our data from the APEC countries to compute dynamic efficiency and relate it to per capita income.

The three models used here capture different aspects of the Olson hypothesis. The static model is used to see whether there is a positive relationship between *levels* of relative efficiency and per capita income. The comparative static model allows us to consider whether productivity *change*, and its components, technical change and efficiency change (catching up or imitation) are correlated with per capita income, i.e., do successfully developing countries succeed through catching up [as suggested by Van and Wan (1997)] or technical advance? Our dynamic model is used to find out whether there is a dynamic relationship between efficiency and per capita income, in particular, does endogenous investment play an important role in the relationship between efficiency and income?

2. The Static Model

In this section, we present the static activity analysis or DEA[3] model, together with some of the properties it satisfies. This model is used to compute measures of technical efficiency which are measures of the relative *level* of efficiency or productivity. These computed measures are then used to shed light on the Olson idea that poor countries are less efficient than rich countries.

Our static activity analysis model relates input vectors $x = (x_1, \ldots, x_N) \in \mathbf{R}_+^N$ to output vectors $y = (y_1, \ldots, y_M) \in \mathbf{R}_+^M$ through a "piecewise linear" transformation. We assume that there are $k = 1, \ldots, K$ activities, which can be individual firms or as in our case individual countries. Each activity is characterized by its input-output vector $(x^k, y^k) = (x_{k1}, \ldots, x_{kN}, y_{k1}, \ldots, y_{kM})$. These $k = 1, \ldots, K$ vectors form the coefficients of the model, and, together with the intensity variables, $z_k \geqq 0$, $k = 1, \ldots, K$ construct the output set as

$$
\begin{aligned}
P(x) = \{(y_1, \ldots, y_M)\colon\ & y_m \leqq \sum_{k=1}^{K} z_k y_{km}, && m = 1, \ldots, M, \\
& \sum_{k=1}^{K} z_k x_{kn} \leqq x_n, && n = 1, \ldots, N, \\
& z_k \geqq 0, && k = 1, \ldots, K\}.
\end{aligned} \tag{1}
$$

Thus the output set $P(x)$ consists of all output vectors $y \in \mathbf{R}_+^M$ that can be produced from the input vector $x \in \mathbf{R}_+^N$. The output set is then formed from the $M + N$ inequalities and the $k = 1, \ldots, K$ nonnegativity constraints above.

The technology defined in (1) satisfies some important properties:[4]

- Inputs are freely disposable, i.e., $x \geqq x'$ implies $P(x') \subseteqq P(x)$.
- Outputs are freely disposable, i.e., $y \leqq y' \in P(x)$ implies $y \in P(x)$.
- The output set $P(x)$ is convex, i.e., $y, y' \in P(x)$ and $0 \leqq \lambda \leqq 1$ imply $\lambda y + (1 - \lambda) y' \in P(x)$.
- The input set, i.e., $L(y) = \{x\colon y \in P(x)\}$ is convex.
- Constant returns to scale holds, i.e., $P(\lambda x) = \lambda P(x)$, $\lambda > 0$.

3. DEA is the abbreviation for data envelopment analysis, a phrase which was coined by Charnes, Cooper, and Rhodes (1978).

4. A proof can be found in Färe and Grosskopf (1996), pp. 41–44.

In addition to these five properties one can show that $P(0) = \{0\}$ and that $P(x)$ is a bounded set. It is also true that the graph $GR = \{(x,y) : y \in P(x),\, x \in \mathbf{R}_+^N\}$ is a closed set. Finally, the technology in (1) satisfies

- The law of diminishing returns, i.e., for scalar output, $\sup\{y : y \in P(x),\, x_1 \leqq \overline{x}_1,\, x_n \geqq 0,\, n = 2, \ldots, N\} < +\infty$, if x_1 is essential in the sense that $P(0, x_2, \ldots, x_N) = \{0\}$, for all $x_2, \ldots, x_N \geqq 0$.

The law of diminishing returns is, of course, a fundamental concept in economics, namely, it is one of our notions of scarcity. As we shall see our dynamic model also allows for diminishing returns.

To cast light on the idea that poor countries are less efficient than rich countries, we first introduce a static measure of inefficiency. The measure we use is the output-oriented Farrell measure of technical efficiency. This is defined as the reciprocal of Shephard's output distance function. Specifically, for a given country or activity k', we can calculate its technical efficiency as the solution to the following linear programming problem

$$
\begin{aligned}
(D_o(x^{k'}, y^{k'}))^{-1} = \max_{z,\theta}\ & \theta \qquad\qquad (2)\\
\text{s.t.}\ & \sum_{k=1}^{K} z_k y_{km} \geqq \theta y_{k'm}, \quad m = 1, \ldots, M,\\
& \sum_{k=1}^{K} z_k x_{kn} \leqq x_{k'n}, \quad n = 1, \ldots, N,\\
& z_k \geqq 0, \quad k = 1, \ldots, K.
\end{aligned}
$$

The measure $F_o(x^{k'}, y^{k'}) = (D_o(x^{k'}, y^{k'}))^{-1}$ can be interpreted as the ratio of maximum potential output to observed output, given the input bundle of country k'. Equivalently, it can be thought of as the ratio of maximum to observed average product, where average product is to be interpreted as a measure of total factor (rather than single factor) productivity. Intuitively, it tells us how far a country k' is from the part of the world production frontier consistent with its input levels and mix.

The world frontier is created from the $k = 1, \ldots, K$ input-output vectors (x^k, y^k) in accordance with (1), i.e., it is best practice based on the sample. This is very similar to the idea of the "meta-production function" used in Kim and Lau (1994). In our case if the value of $F_o(x^{k'}, y^{k'})$ equals one, then country k' is on the best practice frontier and is therefore efficient relative to the countries or activities in the sample. If the value of $F_o(x^{k'}, y^{k'})$ exceeds one, then it is relatively inefficient, i.e., it could have produced proportionally more output by adopting the world frontier technology at its observed input mix. We note that the evaluation of efficiency is in some sense local, since any observation is compared to the best practice frontier at its own observed input mix. It will be compared to countries with a similar input mix. Nonetheless, the presumption is that all countries, at least in principle, have access to the same technology, as in Kim and Lau.

We apply the model (2) to analyze the performance of countries in APEC (Asian-Pacific Economic Community). This includes Australia, Canada, Chile, China, Hong Kong, Indonesia, Japan, Korea, Malaysia, Mexico, New Zealand, the Philippines, Papua (New Guinea), Singapore, Taiwan, Thailand, and the United States. These countries have in common a border on the Pacific ocean. Otherwise they represent a fairly diverse group both economically and politically. The data are gleaned from the Penn World Tables, version 5.6. We follow Färe, Grosskopf, Norris, and Zhang (1994) and use real GDP as our output variable and employment and nonresidential capital stock as inputs. These are in international prices, base year 1984. The data are compiled for the 1975–1990 period.[5] These are the same variables used in Kim and Lau (1994).

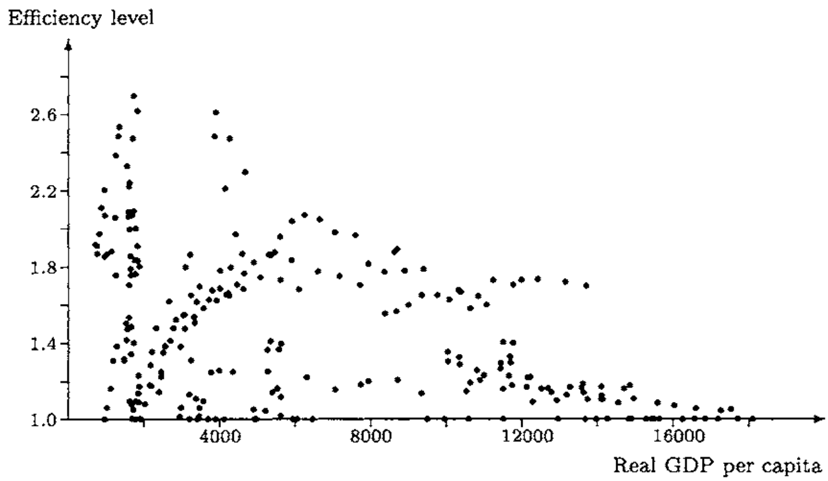


Fig. 1. Efficiency level vs. Y/P for 17 countries 1975–1989

5. Capital stock data are not available for China, Indonesia, Malaysia, Singapore, and New Guinea. We use the perpetual inventory method (benchmark year 1960, depreciation set at .10) to construct capital stock series for these countries based on investment data from PWT 5.6.

We compute technical efficiency for each country for each of the years 1975–1990. We plot the resulting efficiency measures against the corresponding per capita income in figure 1. Recalling that efficiency values in excess of one reflect *in*efficiency, the plotted values suggest a negative correlation between degree of *in*efficiency and per capita income, i.e., rich countries are relatively more efficient than poor countries. Not only is average efficiency lower in poor countries than rich, but there is a greater degree of dispersion in performance. Note that this is based on an entirely static model.

3. Comparative Statics: Catching-Up and Technical Change

In this section we turn to the comparative statics of performance: i.e., we compare performance across periods, but in a static framework. Our previous model provides a measure of the relative *level* of (total factor) productivity of a given country in a given period. In this section we compute the *change* in relative total factor productivity between periods. The index we use to compute total factor productivity change is the Malmquist productivity index first proposed by Caves, Christensen, and Diewert (1982) and operationalized in an activity analysis framework by Färe, Grosskopf, Lindgren, and Roos, initially in 1989. That approach allows us to identify which countries are shifting the best practice production frontier, as well as identifying which countries are catching up.

To get an intuitive feel for the Malmquist productivity index and its component measures, suppose one input is used to produce a single output. Moreover, suppose that the technology, represented by its graph GR, is known at two time periods t and $t+1$. Let (x^t, y^t) and (x^{t+1}, y^{t+1}) be two given input-output vectors, then the output-oriented Malmquist productivity index is easily illustrated, see figure 2. The two observations (x^t, y^t) and (x^{t+1}, y^{t+1}) belong to their own period technologies, i.e., GR^t and GR^{t+1}. The t period observation is also feasible at $t+1$, i.e., $(x^t, y^t) \in GR^{t+1}$, but (x^{t+1}, y^{t+1}) is not feasible in period t, i.e., technical progress has occurred.

We can measure the comparative static performance for these two observations by computing and comparing the corresponding distance functions. In terms of the distances on the y-axis, we observe that $D_o^t(x^t, y^t) = 0f/0e$ and that $D_o^{t+1}(x^{t+1}, y^{t+1}) = 0c/0a$. These measure the efficiency of (x^t, y^t) and (x^{t+1}, y^{t+1}), respectively. Thus the efficiency change is given is

$$EFFCH = \frac{D_o^{t+1}(x^{t+1}, y^{t+1})}{D_o^t(x^t, y^t)} = \frac{0c/0a}{0e/0f}. \tag{3}$$

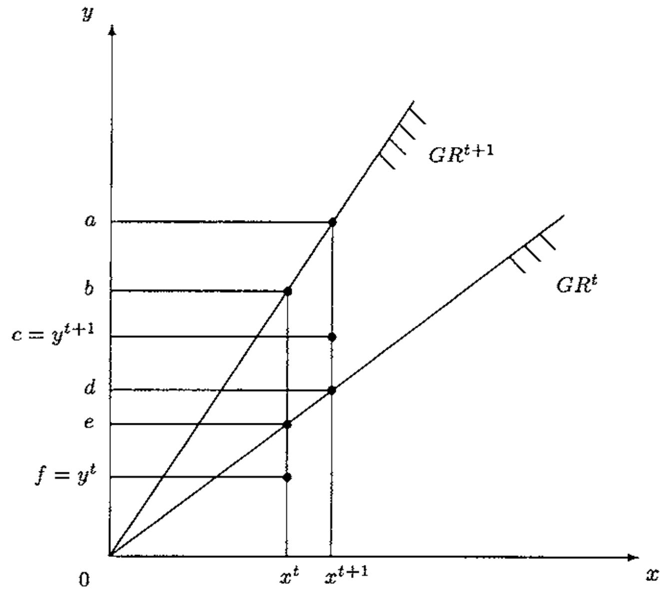


Fig. 2. The output oriented Malmquist productivity index

We use efficiency change to capture the notion of "catching up", i.e., how much closer to (farther from) the frontier a country has come from period t to $t+1$.

Following Färe, Grosskopf, Lindgren, and Roos (1994), we measure technical change as the geometric mean of the shift in the frontier evaluated at x^{t+1} and x^t. In terms of distances on the y-axis in figure 2

$$TECH = \left(\frac{0a/0d}{0b/0e}\right)^{1/2}. \tag{4}$$

In terms of output distance functions this becomes

$$TECH = \left(\frac{D_o^t(x^{t+1},y^{t+1})}{D_o^{t+1}(x^{t+1},y^{t+1})}\frac{D_o^t(x^t,y^t)}{D_o^{t+1}(x^t,y^t)}\right)^{1/2}. \tag{5}$$

We note that in (5) there are two "mixed period" distance functions, namely $D_o^{t+1}(x^t,y^t)$ and $D_o^{t+1}(x^t,y^t)$. In each case, the data being evaluated is from a different period than the technology relative to which it is being evaluated. To illustrate how these may be computed, in terms of the model (1), we may compute $D_o^t(x^{t+1}, y^{t+1})$ for observation k' as

$$(D_o^t(x^{k',t+1}, y^{k',t+1}))^{-1} = \min_{z,\theta} \theta \tag{6}$$

$$\begin{aligned} \text{s.t.} \ \sum_{k=1}^{K} z_k y_{km}^t &\geqq \theta y_{k'm}^{t+1}, \quad & m = 1, \ldots, M, \\ \sum_{k=1}^{K} z_k x_{kn}^t &\leqq x_{k'n}^{t+1}, \quad & n = 1, \ldots, N, \\ z_k &\geqq 0, \quad & k = 1, \ldots, K. \end{aligned}$$

If we multiply (3) by (5), we obtain the Malmquist productivity index proposed by Färe, Grosskopf, Lindgren, and Roos (1994) and used by among others Färe, Grosskopf, Norris, and Zhang (1994). This index is the geometric mean of the t and $t + 1$ period Malmquist indexes as originally suggested by Caves, Christensen, and Diewert (1982).

We compute the Malmquist productivity index as well as $EFFCH$ and $TECH$ for the 17 APEC countries for every pair of years over the 1975–1990 period. The variables used to specify technology are the same as those described in the previous section for the static model. We plot the individual indexes against per capita income.[6] These appear in figures 3–5. To interpret the results, we note that values of the indexes in excess of one are consistent with progress or improvements in performance, whereas values below one reflect declines in performance over the two periods being evaluated.[7] Beginning with the plot of the Malmquist index values, we note that the scatter plot shows slightly more dispersion at the low income end, but with no obvious slope. If we ignore the handful of low-income/low-productivity points that appear to be outliers, the pattern is very flat. The plots for EFFCH and TECH are quite similar. There is more dispersion in terms of the indexes at the low income end than the high income end, but otherwise, there is no obvious relation between per capita income and total factor productivity change and its components.

One obvious reason for the difference in patterns between the earlier static case and the comparative-static case considered here is that the former captured (relative) *levels* of total factor productivity, whereas the

6. We average the income over the two periods involved for each index. Thus, for country k, we plot the the EFFCH index for country k between 1975 and 1976 against their per capita income averaged over 1975 and 1976.

7. This is in contrast to the interpretation of the levels of efficiency discussed in the previous section. There, values in excess of one reflect inefficient performance. Values equal to one signal efficiency.

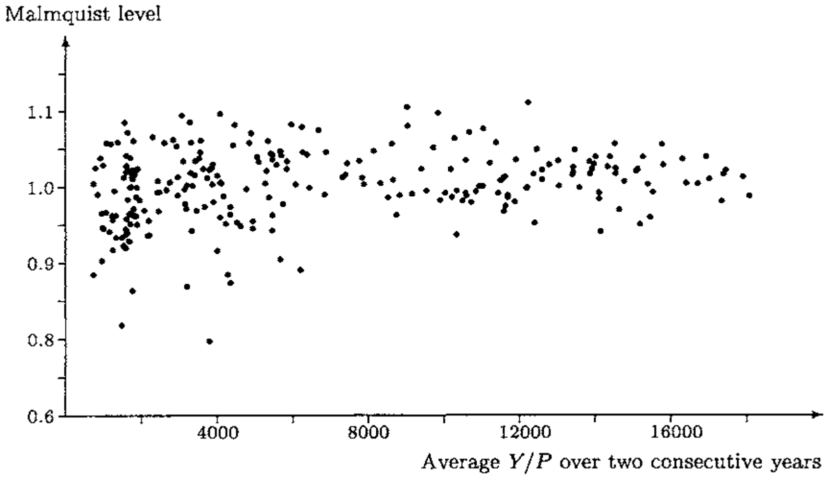


Fig. 3. Malmquist vs. average Y/P for 17 countries from 1975–1989

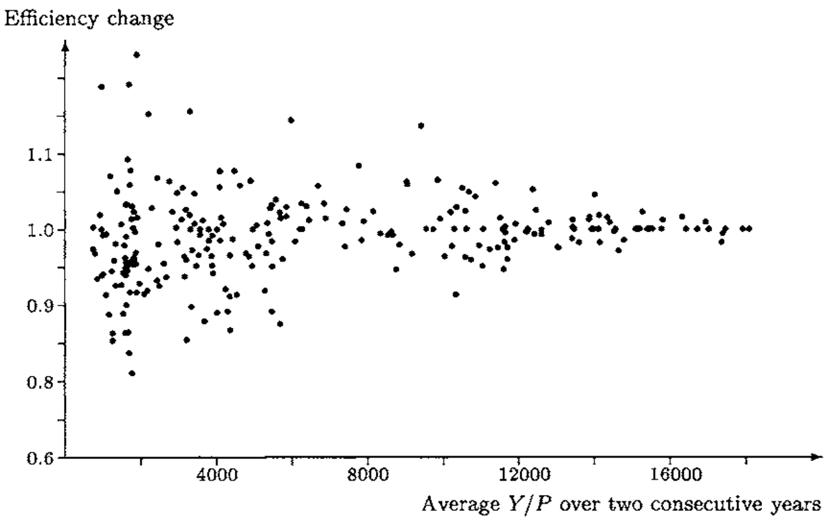


Fig. 4. Efficiency change vs. average Y/P for 17 countries

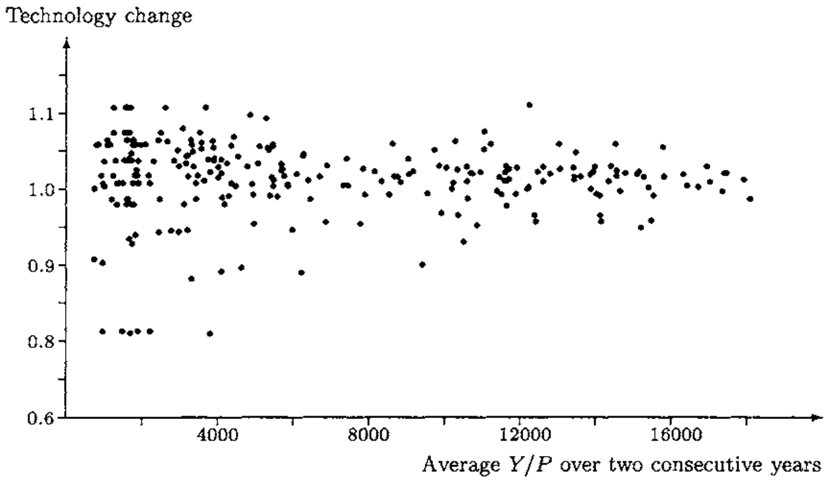


Fig. 5. Technology change vs. average Y/P for 17 countries

current results capture *changes* in total factor productivity over time. In a world of absolute convergence, however, one might expect to see a negative relationship between productivity growth and income. In particular, we might expect to see a negative relationship between catching up and per capita income, however, our plot of efficiency change and income does not provide strong visual support for that type of convergence.

In order to get some sense of the pattern by country, we also computed average annual values of the three components by country as well as the cumulated values. The cumulated values are the multiplicative sums of the individual indexes, i.e., they are the equivalent of a chained index and represent the total change between 1975 and 1990 for the individual countries.[8] These results appear in tables 1–2. For this sample on average, we find evidence of improved productivity: on average productivity change exceeded unity. We also observe across the board improvements in terms of technical change, both on average for each country and in terms of the cumulated values.[9]

8. These indexes do not satisfy the circular test, therefore these values are "path dependent", i.e., their values depend on the march of time.

9. We note that our measure of technical change captures shifts in the frontier, and therefore not the technical change actually realized by any individual country. One can, however, identify which countries are shifting the frontier: if TECH exceeds one and that country has $D_o^{t+1}(x^{t+1}, y^{t+1}) = 1$, then they are shifting the frontier. In

In terms of the cumulative productivity measure, the top performers in this sample are Hong Kong, Singapore, Korea, and perhaps surprisingly, Canada. These countries exhibit both catching up and technical progress. At the other end, the lowest cumulative performers included Indonesia, China, and Papua, New Guinea. For these countries, their adverse performance is due to "falling behind" the frontier as it shifts away from them, i.e., they are becoming relatively more inefficient over time. These three countries are also among the bottom four in terms of per capita income in 1990 in our sample.[10]

Country	Productivity	Efficiency Change	Technical Change
Canada	1.0147	1.0013	1.0133
Mexico	0.9958	0.9803	1.0158
United States	1.0062	1.0000	1.0062
Chile	0.9968	0.9799	1.0173
China	0.9827	0.9748	1.0081
Hong Kong	1.0436	1.0225	1.0206
Indonesia	0.9464	0.9389	1.0080
Japan	0.9978	0.9969	1.0009
Korea	1.0167	0.9934	1.0235
Malaysia	0.9936	0.9810	1.0128
Philippines	0.9976	0.9847	1.0131
Singapore	1.0230	1.0141	1.0088
Taiwan	1.0028	0.9941	1.0087
Thailand	0.9986	0.9875	1.0112
Australia	1.0103	0.9970	1.0134
New Zealand	0.9959	0.9866	1.0095
Papua	0.9923	0.9781	1.0145
Grand Mean	1.0007	0.9887	1.0121

Table 1. Malmquist output based productivity: Average annual changes: 1975–1990

this sample, the U.S., Chile, and Mexico are the off and on shifters over the 1975–1980 period. From 1982 to 1990, the U.S. and Hong Kong are the most frequent technology shifters.

10. The lowest per capita incomes in our sample for 1990 include China ($1324), New Guinea ($1425), the Philippines ($1763), and Indonesia ($1974). These data are from PWT 5.6.

Country	Cumulated Productivity	Cumulated Efficiency Change	Cumulated Technical Change
Canada	1.2443	1.0203	1.2194
Mexico	0.9393	0.7424	1.2652
United States	1.0973	1.0000	1.0973
Chile	0.9531	0.7373	1.2925
China	0.7701	0.6822	1.1288
Hong Kong	1.8968	1.3966	1.3581
Indonesia	0.4376	0.3881	1.1275
Japan	0.9677	0.9542	1.0141
Korea	1.2827	0.9059	1.4158
Malaysia	0.9077	0.7501	1.2100
Philippines	0.9641	0.7930	1.2158
Singapore	1.4064	1.2333	1.1403
Taiwan	1.0422	0.9152	1.1388
Thailand	0.9796	0.8286	1.1822
Australia	1.1664	0.9558	1.2202
New Zealand	0.9406	0.8166	1.1518
Papua	0.8904	0.7170	1.2417

Table 2. Disaggregated cumulative results: 1975–1990

4. A Dynamic Activity Analysis Model

Our static and comparative static results suggest that inefficiency is present in the low income countries in our sample. We would also like to analyze their performance in a dynamic activity analysis framework. That would allow us to analyze the role of investment, which is ignored in our static and comparative static framework.

In the dynamic activity analysis model we allow for intermediate outputs, which serve to provide the link between (discrete) time periods. That is, outputs in say period t can be inputs in period $t+1$. In contrast to the classical Ramsey (1928) model, we allow for many outputs rather than a single final output.

To put the static, comparative static activity analysis models from sections 2–3 into perspective with our proposed dynamic model, we first turn to a schematic drawing. Figure 6 includes a sketch of our static model of technology in terms of the output set. At period t, inputs x^t are used in P^t to produce final outputs y^t. The comparative static model from section 3 can also be illustrated in a simple figure, see figure 7. The difference between figures 6–7 is that in 7 inputs and outputs at $t+1$ are related to the t period technology, and (x^t, y^t) is related to the $t+1$ technology. This is how the two mixed period distance functions, $D_o^t(x^{t+1}, y^{t+1})$ and $D_o^{t+1}(x^t, y^t)$, are generated. These were used in the productivity and technical change indexes.

The dynamic model can also be illustrated in a similar figure, see

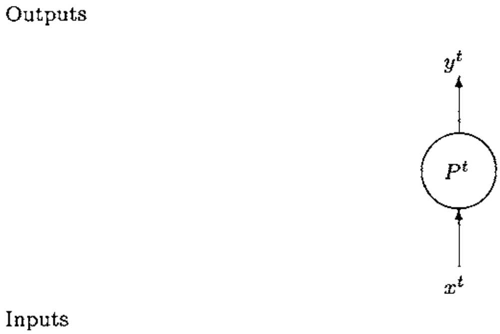


Fig. 6. The static technology

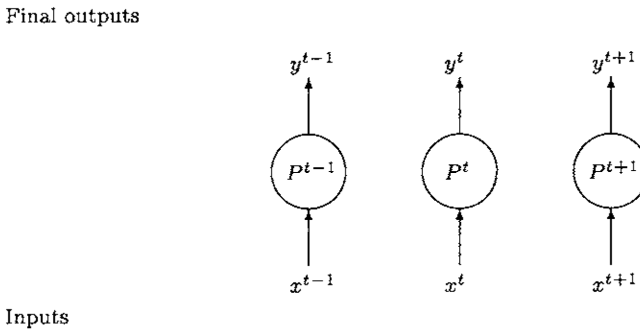


Fig. 7. The comparative static technology

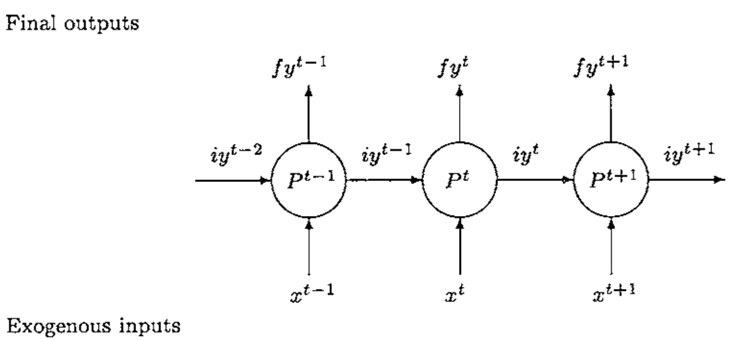


Fig. 8. The dynamic model

figure 8. Suppose that there are three periods, $t-1, t$ and $t+1$ with a different technology at each period, P^τ, $\tau = t-1, t, t+1$. As in the static model P^τ is modelled using observed data on inputs and outputs. At each period, there are two types of inputs, namely x^τ, $\tau = t-1, t, t+1$ which are exogenous to the dynamic technology like labor in the Ramsey model, and iy^τ, $\tau = t-2, t-1, t$, which are endogenous to the technology, like investment.[11]

The endogenous inputs iy^τ are output from the previous period and together with the final outputs like consumption, fy^τ, $\tau = t-1, t, t+1$, constitute total production. For example, the total output at period t equals $(fy^t + iy^t)$. A simple example would be corn production, where fy^t_m is the amount of corn that is consumed and iy^t_m is the amount that is used as seed corn or input in the next period. In this model the "dynamics" i.e., the time interdependence arises from the endogenous inputs $iy^\tau = (iy^\tau_1, \ldots, iy^\tau_M)$, $\tau = t-1, t, t+1$. This model can be easily translated into a form that closely resembles the (discrete) form of the Ramsey model. This could be accomplished by evaluating the final outputs in each period by a corresponding utility function. If we maximize the sum of these utilities, the result is of the standard Ramsey form.[12]

One of the advantages of the dynamic model illustrated in figure 8 is that it can be implemented as a dynamic activity analysis model. Recall that $k = 1, \ldots, K$ denotes the individual observations. In our case we have the same number of observations in every period, i.e., we have a balanced panel of data. We can write the output set consistent with figure 8 for three periods as

$$P(x^{t-1}, x^t, x^{t+1}, iy^{t-2}) = \{(fy^{t-1}, fy^t, (fy^{t+1} + iy^{t+1}) : \tag{7}$$

$$\begin{aligned}
fy^{t-1}_m + iy^{t-1}_m \ &\leqq \ \sum_{k=1}^{K} z^{t-1}_k (fy^{t-1}_{km} + iy^{t-1}_{km}), && m = 1, \ldots, M, \\
&\sum_{k=1}^{K} z^{t-1}_k iy^{t-2}_{km} \leqq iy^{t-2}_m, && m = 1, \ldots, M, \\
&\sum_{k=1}^{K} z^{t-1}_k x^{t-1}_{kn} \leqq x^{t-1}_n, && n = 1, \ldots, N, \\
&z^{t-1}_k \geqq 0, && k = 1, \ldots, K,
\end{aligned}$$

11. In practice, the initial period value of iy^τ is given.

12. That is, we would $\max U^{t-1}(fy^{t-1}) + U^t(fy^t) + U^{t+1}(fy^{t+1})$ given the technologies P^τ and given iy^{t-2}, x^τ, $\tau = t-1, t, t+1$.

$$
\begin{aligned}
fy_m^t + iy_m^t \ \leqq \ & \sum_{k=1}^{K} z_k^t (fy_{km}^t + iy_{km}^t), && m = 1, \ldots, M, \\
& \sum_{k=1}^{K} z_k^t iy_{km}^{t-1} \leqq iy_m^{t-1}, && m = 1, \ldots, M, \\
& \sum_{k=1}^{K} z_k^t x_{kn}^t \leqq x_n^t, && n = 1, \ldots, N, \\
& z_k^t \geqq 0, && k = 1, \ldots, K,
\end{aligned}
$$

$$
\begin{aligned}
fy_m^{t+1} + iy_m^{t+1} \ \leqq \ & \sum_{k=1}^{K} z_k^{t+1} (fy_{km}^{t+1} + iy_{km}^{t+1}), && m = 1, \ldots, M, \\
& \sum_{k=1}^{K} z_k^{t+1} iy_{km}^{t} \leqq iy_m^{t}, && m = 1, \ldots, M, \\
& \sum_{k=1}^{K} z_k^{t+1} x_{kn}^{t+1} \leqq x_n^{t+1}, && n = 1, \ldots, N, \\
& z_k^{t+1} \geqq 0, && k = 1, \ldots, K\}.
\end{aligned}
$$

One can prove, see Färe and Grosskopf (1996), that if each period technology, P^{t-1}, P^t, and P^{t+1} satisfies the conditions listed in section 2, then the dynamic model in (7) satisfies them as well.

For a given country or observation k', we can compute its dynamic efficiency by solving a linear programming problem, namely

$$
\begin{aligned}
& (D_o(x^{k',t-1}, x^{k',t}, x^{k',t+1}, iy^{k',t-2}))^{-1} = \max_{z,\theta,iy} \theta \qquad (8) \\
& \text{s.t.} \quad \theta(fy^{k',t-1}, fy^{k',t}, (fy^{k',t+1} + iy^{k',t+1})) \in \\
& \qquad P(x^{k',t-1}, x^t, x^{k',t+1}, iy^{k',t-2}).
\end{aligned}
$$

We would like to compute a dynamic model of the general sort specified by Färe and Grosskopf and described in (8), but which we can apply to our data and compare to our static and comparative static results. As a consequence, instead of scaling all periods by the same θ, we modify (8) to the following problem:

$$
\begin{aligned}
& (D_o(x^{k',1975}, \ldots, x^{k',1990}, iy^{k',1975}))^{-1} = \max_{z,\theta^\tau,iy^\tau} \sum_{\tau=1975}^{1990} \theta^\tau \qquad (9) \\
& \text{s.t.} \quad \theta^\tau (fy^{k',\tau}) \in P(x^{k',1975}, \ldots, x^{k',1990}, iy^{k',1975}), \\
& \qquad \tau = 1975, \ldots, 1990.
\end{aligned}
$$

In order to make this model comparable to our earlier models, we need to be explicit about what we mean by iy and fy. Our earlier model specifies output for country k in period t, i.e., (y_k^t) as real GDP for that country and year.[13] In this context we now define

$$y_k^t = fy_k^t + iy_k^t, \quad k = 1, \ldots, K; \quad t = 1975, \ldots, 1990. \tag{10}$$

A schematic appears in figure 9. In our programming problem, however, iy^t becomes endogenous starting in 1976. We treat iy^t as investment. This means, of course, that it is related to the capital stock variable we use in the static and comparative static models. In particular, we have

$$c_k^t = c_k^{t-1}(1 - \delta) + iy_k^{t-1}, \quad k = 1, \ldots, K; \quad t = 1976, \ldots, 1990, \tag{11}$$

where c_k^t denotes the capital stock in country k in period t, and δ is the depreciation rate. In the earlier models the capital stock was denoted as x_{k2}^t. We now use c_k^t for capital stock and x_k^t (note that there is no n subscript) to represent employment in country k in period t.

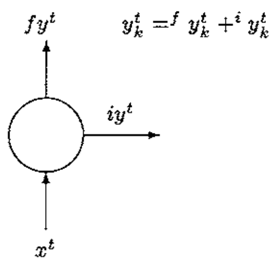


Fig. 9. Final and intermediate output

Using this notation, we can write out our programming problem for country k' as follows

$$(D_o(x^{k',1975}, \ldots, x^{k',1990}, iy^{k',1975}))^{-1} = \max_{z^\tau, \theta^\tau, iy^\tau} \sum_{\tau=1975}^{1990} \theta^\tau \tag{12}$$

subject to:

13. We suppress the m subscript since $M = 1$ in our application.

First Period: $\tau = 1975$

$$\begin{aligned}
\theta^{\tau} f y_{k'}^{\tau} + i y^{\tau} &\leqq \sum_{k=1}^{K} z_k^{\tau} y_k^{\tau}, \\
\sum_{k=1}^{K} z_k^{\tau} c_k^{\tau} &\leqq c_{k'}^{\tau} = (i y_{k'}^{\tau-1} + c_{k'}^{\tau-1}(1-\delta)), \\
\sum_{k=1}^{K} z_k^{\tau} x_k^{\tau} &\leqq x_{k'}^{\tau}, \\
\theta^{\tau} &\geqq 1, \\
z_k^{\tau} &\geqq 0, \quad k = 1, \ldots, K.
\end{aligned}$$

Middle Periods: $\tau = 1976, \ldots, 1989$

$$\begin{aligned}
\theta^{\tau} f y_{k'}^{\tau} + i y^{\tau} &\leqq \sum_{k=1}^{K} z_k^{\tau} y_k^{\tau}, \\
\sum_{k=1}^{K} z_k^{\tau} c_k^{\tau} &\leqq c_{k'}^{\tau-1}(1-\delta) + i y^{\tau-1}, \\
\sum_{k=1}^{K} z_k^{\tau} x_k^{\tau} &\leqq x_{k'}^{\tau}, \\
\theta^{\tau} &\geqq 1, \\
z_k^{\tau} &\geqq 0, \quad k = 1, \ldots, K.
\end{aligned}$$

End Period: $\tau = T$ (1990)

$$\begin{aligned}
\theta^{\tau} y_{k'}^{\tau} &\leqq \sum_{k=1}^{K} z_k^{\tau} y_k^{\tau}, \\
\sum_{k=1}^{K} z_k^{\tau} c_k^{\tau} &\leqq c_{k'}^{\tau-1}(1-\delta) + i y^{\tau-1}, \\
\sum_{k=1}^{K} z_k^{\tau} x_k^{\tau} &\leqq x_{k'}^{\tau}, \\
\theta^{\tau} &\geqq 1, \\
z_k^{\tau} &\geqq 0, \quad k = 1, \ldots, K.
\end{aligned}$$

Note that we are restricting the individual θ^ts to be greater than or equal to one in order to prevent countries from reducing past production. We set the initial capital stock equal to observed capital stock, i.e., investment in 1974 is assumed to be exogenous. In later periods, investment is endogenous, and it is possible to compare the optimal investment path with the actual path whenever this is of interest.

We run this programming problem for each of the 17 countries in our APEC sample. To our knowledge, this is the first time this dynamic model has been estimated. One of the real advantages of this approach is that solutions are found using linear programming; we do not need to use dynamic programming or optimal control techniques. Since this programming problem allows more choice than our original static period efficiency problems, we expect the annual values of the θ^t's to be no smaller than those computed in the static case. This is confirmed by the plot of the individual solution values against per capita income in figure 10. If we compare this to figure 1 (which plots the static efficiency results) against per capita income, we see that the range of inefficiency is greater in the dynamic than the static case, as expected. The general relationship between efficiency and per capita income persists in the dynamic model (although perhaps not as pronounced): high income is associated with high levels of technical efficiency. Low income is associated with a much greater range in terms of performance. The most inefficient observations are low income.

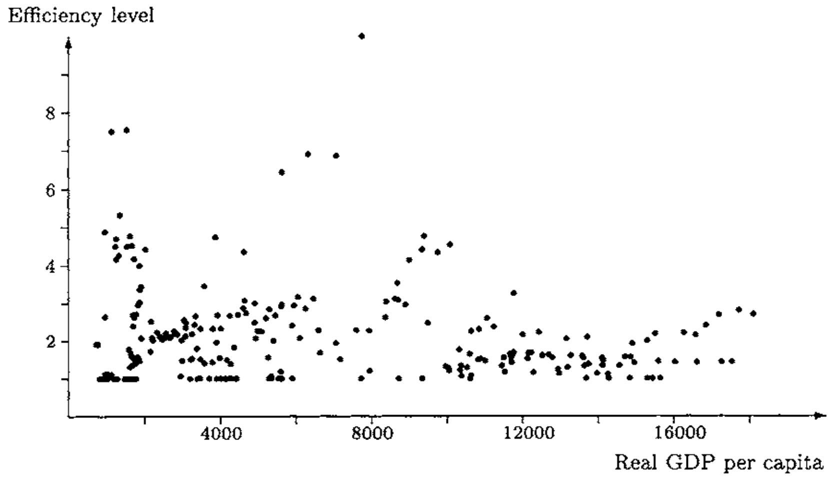


Fig. 10. Efficiency level vs. Y/P for 17 countries 1975–1989

If we ignore the extremely inefficient points in figure 10, say those with scores of 4 or greater, the distribution becomes more uniform than we observe in figure 1 using the static approach. One possible interpretation of this is that changes in investment patterns can benefit countries at a wide range of income levels. The "outlier" points at the low income levels suggest extremely large potential gains from changes in investment for very low income countries.

5. Summary

One of the goals of this chapter was to investigate Mancur Olson's claim that low income countries that do not exhibit the "catching up" that is predicted by traditional growth theory have inefficient institutions and economic policies. We proceed by looking at the relationship between efficiency and income for the sample of countries in the Asian-Pacific Economic Community (APEC) over the 1975–1990 period. The notion of efficiency we use is technical efficiency, which we compute as the reciprocal of a Shephard-type output distance function. The computational approach we use is activity analysis. Intuitively, we construct a best practice technology from the data we have on inputs and output for our sample. Technical efficiency is computed as deviations from the best practice frontier.

We begin by specifying efficiency in the standard static model. Technical efficiency is computed annually for each country in our sample. In this framework, each year has its own technology. Our results suggest that rich countries are relatively efficient. The poorest countries have lower efficiency on average, and a greater variation in performance than rich countries in the static framework.

We next consider a comparative static case. Here we compute intertemporal performance using ratios of distance functions, in particular, we compute Malmquist productivity and its components, technical change and efficiency change. Here we find greater dispersion in performance among poor countries than rich countries, i.e., both the highest and lowest productivity changes are observed at low levels of income. Those countries with the lowest productivity growth are those with the lowest incomes. The three countries with the highest cumulated productivity growth (Hong Kong, Singapore, and Korea) all have per capita incomes below $6000 in the base year, 1975. Real per capita income more than doubled between 1975 and 1990 in these three countries. Hong Kong and Singapore have the highest cumulated efficiency improvements, whereas Korea has the highest cumulated technical change in the sample.

Finally, we develop a simple dynamic activity analysis model to measure dynamic efficiency. This simple model allows for allocation over time through investment. The dynamic performance measures result in even greater dispersion at the low income end than the static measures. Dispersion is relatively low at the high income end. Again we see the pattern of higher average efficiency at high incomes, and relatively low average efficiency (with high variance) at the low end. Nonetheless, we do find considerable potential gains to high income countries from changes in investment.

Our evidence is, of course, only suggestive. We would like to enhance the empirical and theoretical model by explicitly including human capital. Future empirical work would presumably focus on an enriched specification of technology to include multiple outputs, disaggregated inputs adjusted for quality, and an expanded number of countries and time periods. As suggested by the editors, it would also be of considerable interest to look at the relationship between efficiency and trade and efficiency and growth explicitly. We would also like to use our efficiency and productivity measures to pursue formal hypothesis tests. Finally, from a policy perspective, it is important to identify the *sources* of inefficiency.

Acknowledgement. This chapter was prepared while the authors were visiting fellows at the Department of Agricultural and Resource Economics at the University of Maryland. We would like to thank the Center for Agricultural and Natural Resources for their support and Bob Chambers and Marc Nerlove for helpful discussions while we were writing the manuscript. We are also grateful to Bill Barnett for his careful reading of and insightful comments on our chapter. Thanks are also due to Bjarne S. Jensen and Kar-yiu Wong for their extensive comments and suggestions. The authors are listed alphabetically. Please direct correspondence to Shawna Grosskopf.

References

Caves, Douglas, Laurits Christensen, and W. Erwin Diewert (1982). "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity." *Econometrica* 50, 1393–1414.

Charnes, Abe, William W. Cooper, and Ed Rhodes (1978). "Measuring the Efficiency of Decision Making Units." *European Journal of Operations Research* 2, 429–444.

Färe, Rolf, and Shawna Grosskopf (1996). *Intertemporal Production Frontiers: With Dynamic DEA*. Boston: Kluwer Academic Publishers.

Färe, Rolf, Shawna Grosskopf, Björn Lindgren, and Pontus Roos (1994). "Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach." In A. Charnes, W.W. Cooper, A. Lewin, and L. Seiford (eds.), *Data Envelopment Analysis: Theory, Methodology and Applications.* Boston: Kluwer Academic Publishers.

Färe, Rolf, Shawna Grosskopf, Mary Norris, and Zhongyang Zhang (1994). "Productivity Growth, Technical Progress and Efficiency Change in Industrialized Countries." *American Economic Review* 84, 66–83.

Kim, Jong-Il, and Laurence J. Lau (1994). "The Sources of Economic Growth of East Asian Newly Industrialized Countries." *Journal of Japanese and International Economics* 10, 235–271.

Malmquist, Sten (1953). "Index Numbers and Indifference Surfaces." *Trabajos de Estadistica* 4, 209–242.

Olson, Mancur, Jr. (1996). "Distinguished Lecture on Economics in Government: Big Bills Left on the Sidewalk: Why some Nations are Rich, and Others Poor." *Journal of Economic Perspectives* 10, 3–24.

Ramsey, Frank P. (1928). "A Mathematical Theory of Saving." *Economic Journal* 38, 543–559.

Shephard, Ronald W., and Rolf Färe (1980). *Dynamic Theory of Production Correspondences.* Cambridge, Mass.: Gunn and Hain Publishers, Inc.

Van, Pham Hoang, and Henry Y. Wan, Jr. (1997) "Interpreting East Asian Growth." (this volume).

CHAPTER 9

Interpreting East Asian Growth

Pham Hoang Van and Henry Y. Wan, Jr.

1. Introduction

We study in this chapter the East Asian economies, in particular, the Newly Industrialized Economies (NIEs) of Korea, Taiwan, Hong Kong, and Singapore. These economies have attracted much attention both for their own sake and for the valuable implications embodied in their remarkable experiences. Their performances are subjects of recent raging debates.[1] The development of the NIEs are characterized by:

1. Their sustained rapid growth in per capita real income of 6% per annum for 30 years;
2. Their export expansion which is spectacular in both volume and variety, including high-tech exports for advanced economies;
3. Their macro-economic stability is manifested in: (i) low unemployment and low annual inflation, (3% or less in both for some NIEs); (ii) low income inequality; and (iii) a balanced government budget and balanced international payments.

The debates concerning the NIEs are focused upon four aspects:

1. Is *trade* essential in the development of the NIEs? On this Young (1994) expressed the novel view that trade has only a once-and-for-all impact on technology;
2. Is *State guidance* indispensable for their development schemes? On this Amsden (1989) for example, holds that State guidance plays an indispensable role;
3. Is *technical progress* present in their evolution? On this Kim and Lau (1994a) and Young (1995) claim that, unlike five OECD economies, the NIEs have made no technical progress of statistical significance;
4. Is *collapse* Soviet-style the eventual fate of the NIEs? On this Krugman (1994) is sure.

1. See, for example, Page (1994).

We argue in this study that in the development of the NIEs, that (i) trade has been playing a continuous, indispensable role; state guidance has not; and (ii) there is continuous gain in technology, in the sense as both laymen understand it and what policy-makers care about. This may be so even if the NIEs' performance may be unremarkable by some derived measure (like the "Solow residual"); and (iii) the tempo of their growth is likely to taper off as has happened to Japan, but there is no reason to expect a nose-dive, Soviet style.

Our analysis is based upon two fundamental properties of the NIEs. First, in their mechanism for growth, there is a *generalized complementarity* among (i) capital accumulation, (ii) technology gain, and (iii) international trade. Their growth would be impossible if any one of the three is lacking. This view generalizes the complementarity hypothesis of Kim and Lau (1994b), between physical and human capital accumulation.

Second, about their mode of development, the contagion model of Findlay (1978) captures the essence of the matter. The NIEs acquire technology by close association with the developed world where in such "emulative growth", they rely on "borrowed technology", not domestic research and development (R&D). Their rapid progress is secured by side-stepping the risky and costly R&D investments, yet, the avoidance of which comes at a price: the foregoing of the innovators' profits. Thus, without such profits, the NIEs must devote much more inputs than the advanced economies in earning the same unit of value on the world market.

Analytically, we do not assume the existence of any particular form of an aggregate production function, which can vary under technical progress only in some specific manner (e.g., output-augmenting or input-augmenting). In the development context, what may count for an aggregate production function is contingent upon different trade regimes and varies at a pace which depends on the evolution of the economy. Our stance follows the position of Solow stated in his review of Hicks' *Capital and Growth* thirty years ago. To him, the aggregate production function which he employed with such virtuosity is but an artifice, to be used wherever it is useful and in the absence of superior alternatives. Notably, Solow never for once applied such a tool in the context of economic development. To address development concerns, we fashion an analytic apparatus from an extension of the surrogate production function of Samuelson (1962).

Following the tradition of economic analysis on many issues (e.g., the backward-bending labor supply curve), our study is motivated by casual observations, followed through with deductive reasoning and illustrated with specific examples. Its ultimate validation must come from empirical evidence of a formal or casual nature.

Returning to the controversies we alluded to earlier, we believe that the schematic chart shown in figure 1 provides a bird's eye view of what shapes the rapid growth of the NIEs.

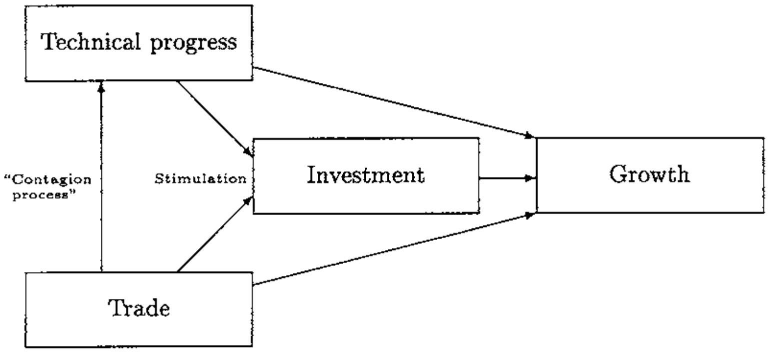


Fig. 1. Causal links

The rest of the chapter is organized as follows. In section 1, we shall set forth our basic notions about technology: its nature, acquisition, economic representation, and implications on the innovation process among the OECD states and the emulation process of the NIEs. In section 3, we focus on the relationship between technology acquisition and capital accumulation, from both the supply and demand sides of the coin. Section 4 examines the crucial role played by trade in technology acquisition by the NIEs. We conclude with comments on the outlook of the NIEs.

2. Technology, Innovation, and Emulation

We focus attention on the acquisition of technology by a developing economy which has little influence on world market prices. It is useful now to digress on the concept of *technology* we use in this study.

Here, technology is identified as the vast body of tacit and specific knowledge regarding what is to be done under various eventualities in production. It resembles an expert system program in computer science, or a "policy" in dynamic programming. As illustrated in the well known case of "Sexing the Chicken", a group of Japanese farmers found no way to instruct their American hosts about how to decide the gender of new-born chickens by inspection; the latter gained proficiency only by watching the Japanese practising their art [see Biederman and Shiffrar (1987)]. Thus, in real life, oftentimes, technology cannot be acquired fully by merely taking possession of a set of operating manuals or blueprints, nor by undergoing certain types of formal training, though all these may help.

Among many alternatives, three approaches for acquiring technology deserve attention: (i) learning by doing, in which one progresses by both "trial and error" and serendipity; (ii) organized R&D, which is heavily relied upon by developed economies in innovation; and (iii) observations on how the "informed" parties act (sometimes under the formal tutelage of the latter), which is important to the less developed economies in the process of "emulation".

In developed economies, costly and risky R&D is often necessary to achieve appreciable productivity gain: their production methods are already close to the current best practice. For the developing economies, reverse-engineering is often far more attractive than re-inventing the wheel. Because it is "self-financing", learning by doing is useful to all economies, yet its effect exhibits diminishing returns at some point [see Young (1993) on bounded learning].

For our purpose, we shall adopt the following simplifying assumptions.

Assumption 1. *The vector of world prices, p, is given.*

Assumption 2. *Constant returns prevail in all production processes, with labor, capital, and possibly some intermediate goods as inputs, according to given proportions.*

A production process, $\mathbf{P}$, is characterized by the ordered triplet, (x_P, K_P, L_P), where x_P is a vector with positive signs for outputs and negative signs for inputs and K_P and L_P are respectively the capital and labor inputs associated with that process. Write

$$V_P = p\,x_P$$

as the value added of $\mathbf{P}$. Clearly only processes with $V_P > 0$ will be carried out in real life.

Remark. In contrast to the traditional treatment, there is no need at this point to rule out joint outputs. Nor must traded inputs or non-traded outputs receive any special treatment. These only influence the characteristics of particular processes. ∇

By Assumptions 1–2, we need only consider any *normalized process* **P** with unity as its value-added. For our present purpose, we focus on (L_P, K_P), its ordered pair of labor and capital requirements, which is the projection of **P** on $\mathbf{R}^2$.

Definition 0. **C** *is the (closed) convex hull of process-specific input pairs for all* **P**. *Its lower boundary is shown as AA' in figure 2 and is the economic representation of the technology open to an economy at a specific point in time. In such a representation, one abstracts from descriptive details of what collection of goods are producible in an economy and what collection of processes are available to produce a particular good. Here, it is natural to introduce the following definition.*

Definition 1. *A technology gain means a set-theoretic enlargement of* **C** *over time.*

We now adopt:

Assumption 3 [Atkinson and Stiglitz (1969); Lapan and Bardhan (1973)]. *Any gain in technology is localized.*

In figure 2, the process of innovation which reduces the input requirement pair of a process from M to N leads to an improvement of technology to the new envelope: $ATNT'A'$, with the area $TNT'M$ representing the gain.

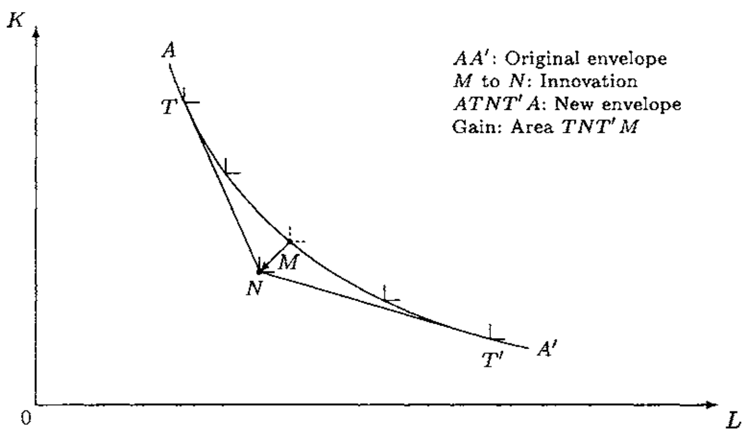


Fig. 2. Innovation

Emulation by a developing economy (depicted in figure 3) differs from innovation by a developed economy in that it represents a sequence of movements of the BB' envelope toward the AA' envelope, ending on some position along CC'. The main difference between innovation and emulation is that for innovation, the nature of the next break-through is unpredictable. In contrast, emulation is an activity with a ready template, the AA' envelope. At this point, it is natural to introduce a concept of technology gap below:

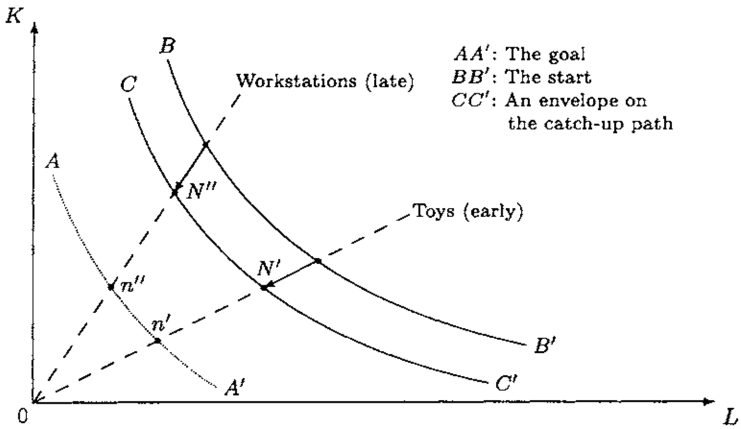


Fig. 3. Emulation

Definition 2. *If the aggregate input pair of an economy is (L, K) with an output value V and if another economy can produce the same output V with an input pair, $[(1-g)\,L, (1-g)\,K]$, then,*

$$g = \text{the technology gap.}$$

For simplicity, we adopt further that:

Assumption 4. *AA', BB', and CC' in figure 3 belong to a family of homothetic loci, with CC' lying between AA' and BB'.*

The reason that CC' is likely to lie somewhere above AA' will be discussed later.

Next, we shall postulate,

Assumption 5. *At any point in time, a developing economy can catch-up with the advanced economies in allocative efficiency, over all processes not exceeding the critical "degree of complexity", z, which is positively associated with the capital-labor ratio, $k(z)$. This marks the most complex process which an economy can master by emulation.*

If z is a real-valued index for human capital (or, knowledge capital), then the functional relationship $k(z)$ is the manifestation of the complementarity hypothesis of Kim and Lau.

Definition 3. *$C(z)$ is that convex hull when the best practice has been acquired for all processes up to and including any technique at complexity z.*

Definition 4. *$\partial C(z)$, the boundary of $C(z)$, is a unit-value isoquant characterizing a surrogate production function $F(K, L; z)$* [see Samuelson (1962)].

Definition 5. *$f(k; z) = f(K/L; z) = F(K, L; z)/L$ is the associated per worker production function.*

Remark. $f(k; z)$ may have a kink at the point, $(k(z), f(k(z); z))$. ∇

Remark. z represents the ability of an economy to "domesticate" a complex process. As is emphasized in the theory of endogenous growth, such "industrial competence" is a nonrival good, accumulated through the external effects of the actions by the individual firms and persons. The process of its accumulation over time is considered in a later section. ∇

Example. In figure 4, success in emulation comes to toys before coming to workstations. ∇

With the production of toys (resp. workstations) successfully emulated, the required input pairs for all processes no more capital intensive than toys (resp. workstations) are on the CC' locus, while those for more capital-intensive processes remain on the BB' locus. This implies an envelope of $C'N'T'B$ (resp. $C'N'N''B$).

We can also represent the situation in terms of the surrogate production function as shown in figure 5. The developed economies possess advanced technology depicted by a surrogate production function which is everywhere above the surrogate production function of the developing economy which uses backward technology.

For a developing economy which has successfully emulated techniques for toy production but not those for production of more complex goods, the surrogate production function would be the darker segments

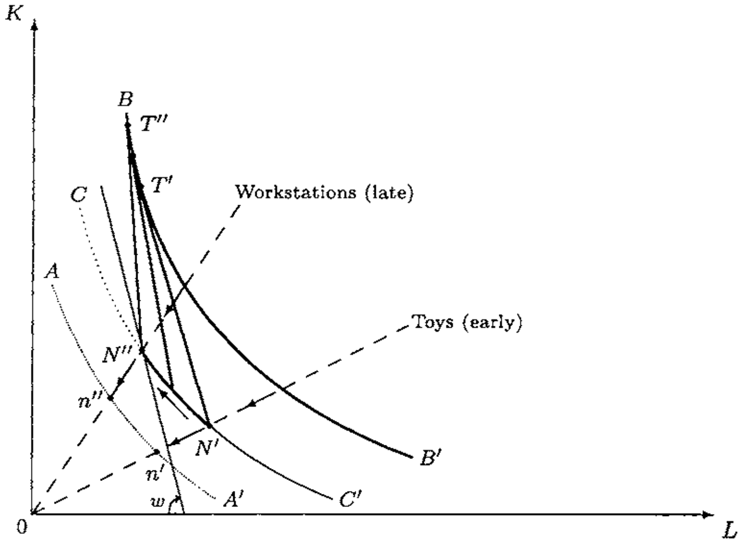


Fig. 4. Evolution of envelopes

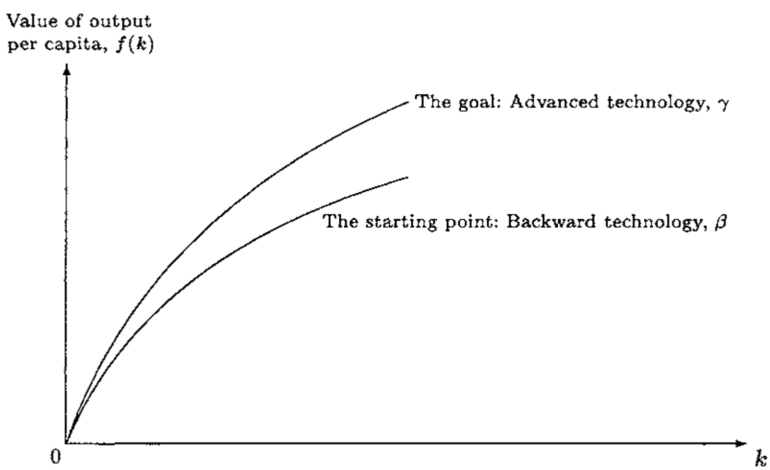


Fig. 5. The two worlds

in figure 6a with a kink at point n. As the economy gains more technical expertise, the kink point would move up along the higher production function, to n'' (shown in figure 6c) at the point where the production of workstations has been mastered.

The steady rise of the capital intensity of outputs is well documented for the NIEs and undisputed. The questions at issue are three: (i) what *causes* such an evolution: capital accumulation or technological gain? (ii) how is a new output introduced to the NIEs? and (iii) what measures the technical capability of the NIEs?

To both Kim and Lau (1994a,b) and Young (1995), capital accumulation is the essence of growth for the NIEs. The introduction of a new product to these economies only involves the simple act of buying the needed equipment. For the scenario we have described in figure 4, Kim-Lau and Young would represent the technical capability of the NIEs with the same CC' path over the entire period of observation. Thus, in contrast with the advanced economies, the measure for technology gap:

$$g = 1 - (\overline{On}'/\overline{ON}') = 1 - (\overline{On}''/\overline{ON}'')$$

stays at a constant value, with no tendency of narrowing.

This is at variance with our interpretation, where the prime mover is the sequential acquisition of various product-specific technologies. Over time, the unit value envelope for input requirements moves gradually from $C'N'T'B$ to $C'N''T''B$. By comparing Definitions 1–2, we can now state the following proposition.

Proposition 1. *Under Assumptions 1–5, an economy may enjoy a sequence of technological gains without reducing the technology gap.*

By our interpretation, CC' is the historical path traced out over time. Its entire length does not represent the technical capability of an economy at any instant, and therefore, the measure g is by no means a perfect measure of the differences in technology between two economies. The reasoning backing our interpretation is elaborated in the next section.

Replace Assumption 1 with:

Assumption 1$'$. *World prices are given to individuals in the LDCs, including the NIEs, but innovators in the DCs can set prices monopolistically.*

We can now consider the situation depicted in figure 5 where the absence of innovation in the NIEs (a consequence of their low levels of R&D efforts) means the absence of the profit component in output

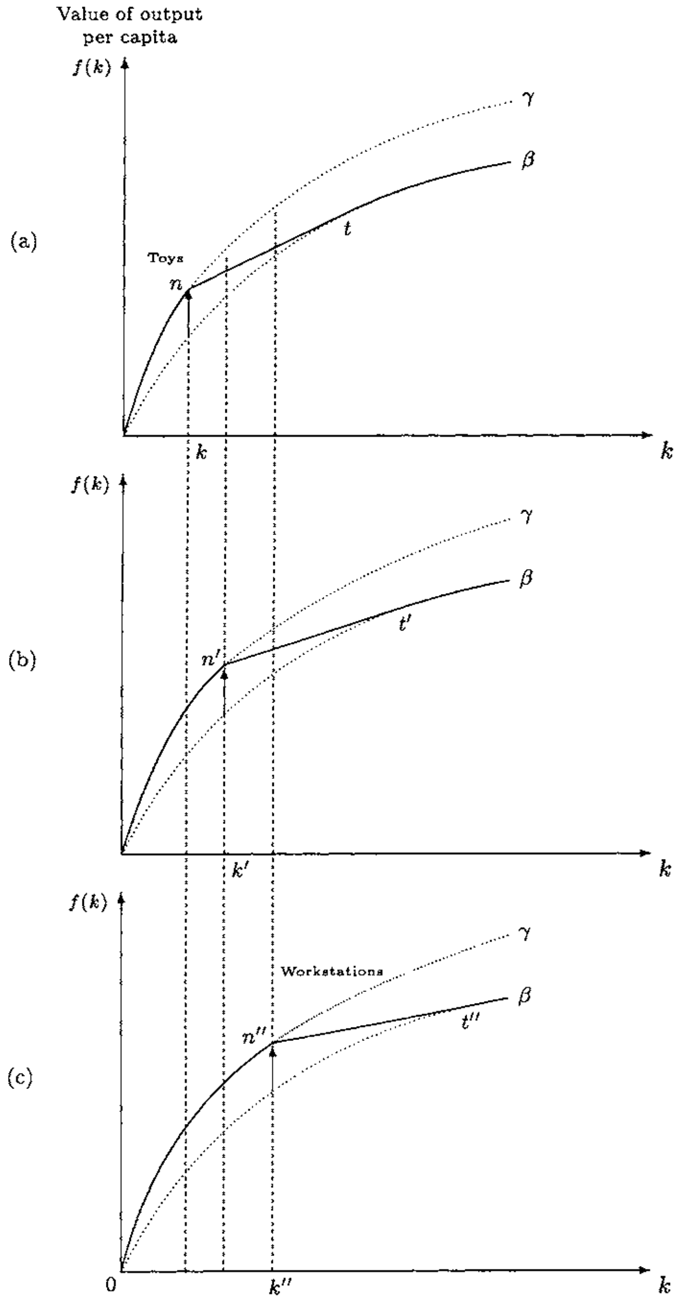


Fig. 6. Staged technology transfer

values. Hence more inputs are required to obtain the same aggregate output value. This may be formally stated as,

Proposition 2. *Under the aforementioned assumptions, the lack of R&D efforts in the NIEs implies that they will persistently face $g > 0$.*

The separation of AA' from CC' in figures 3–4 follows naturally the above result.

3. The Role of Technology Acquisition

In the scenario discussed in the last section, both capital accumulation and gains in technology are at work. For Kim and Lau as well as for Young, there is not much technical progress in the NIEs, and it is the extraordinary rates of capital accumulation which fuel their growth in per worker output. In contrast, to us, technological advance is the ultimate cause. Two systematic factors have probably affected the results of Kim-Lau and Young.

The first concerns the problem of identification. Again, return to our simplified scenario. CC' in figures 3–4 would be identified by them as an isoquant of the aggregate production function (or as lying on the surface of the *same* production set). To us, each observation lies on a different production set. In principle, one may conduct an empirical test which could conceivably clarify the issue. Consider the situation in figure 4, and suppose that the best practice for producing the workstation has just been acquired. Note that the upper contour set $\mathbf{C}$ has a kink at N''. The tangent of CC' at N'' has a slope which is their presumptive marginal rate of factor substitution (viz., when CC' is treated as an isoquant of the aggregate production function). This can be either equal to or less steep than the market ratio of wage to the user cost of capital. At N'' in figure 4, the line with slope w only has to be a supporting line of the upper contour set $\mathbf{C}$, not of the envelope, CC'. Should the magnitudes of the estimated slopes of CC' be repeatedly less than the corresponding observed w's, then clearly CC' is not lying on an unchanged production set.

The second factor concerns the "product-mix effect", namely, the impact of the changing product-mix on the net accumulation for the NIEs. It is known that a rapid succession of products are important in the NIEs [e.g., Young (1992)]. For example, the output mix of the Handok Co., Korea was dominated by wigs early on, but by computers a dozen years or so later [Kim and Leipziger (1993)]. Since equipment for wig-making is hardly appropriate for producing computers, these economies have no recourse but to make continuously heavy gross invest-

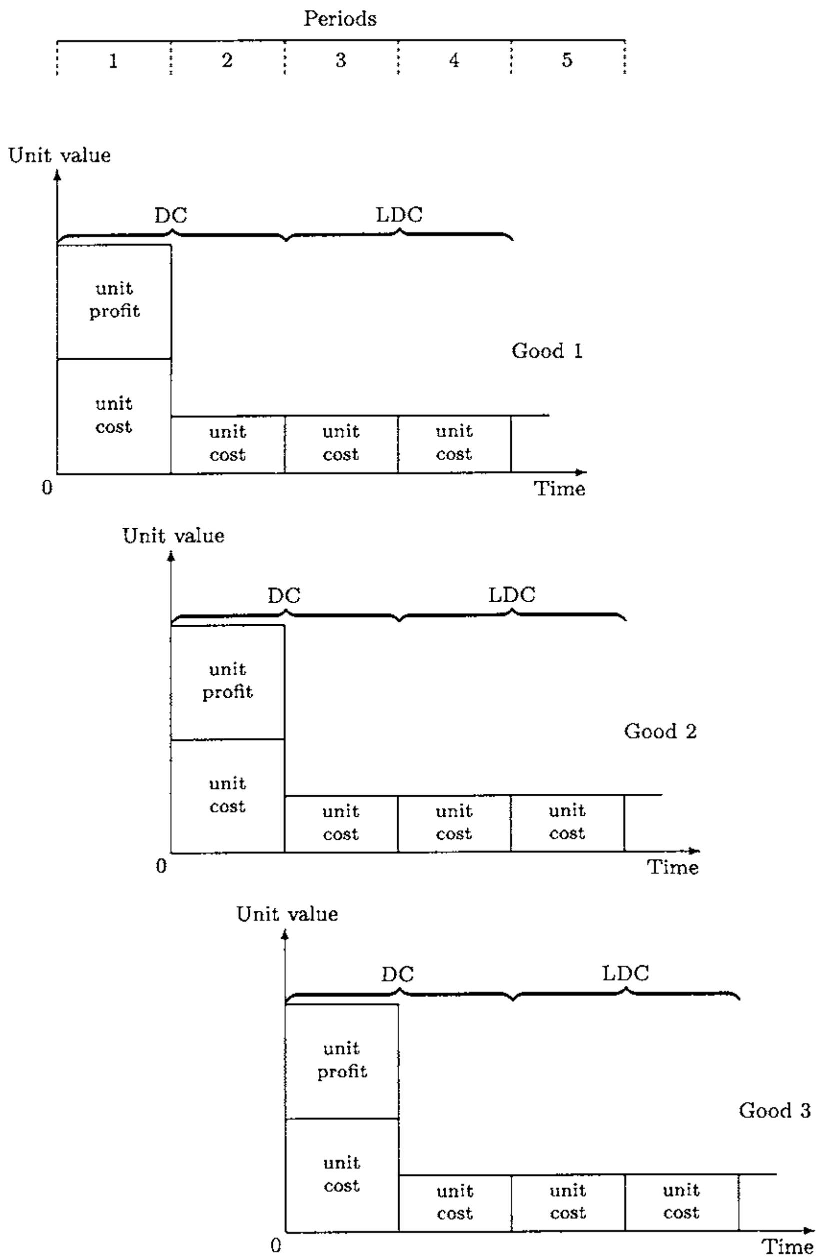


Fig. 7. Product cycles

ments in equipment and sustain higher rates of economic obsolescence than otherwise.

It should be noted that the principal argument behind our view is conceptual. In market economies, the rate of capital accumulation by rational agents can never be a causal determinant for growth. It is endogenously determined by profit prospects. These can never be separated from the acquisition of product technology (as well as the opening to trade). The above reasoning remains true both when an economy has access to the international capital market and when it has not. The relevance of this argument is attested by the experience of the NIEs.

To illustrate the above points, we shall embed now the concept of technology in a sketch of an open-economy Ramsey model specified below.

Assumption 6. *Labor endowment is constant over time. By normalization, $L(t) \equiv 1$ and the capital good is fully malleable.*

Remark. Capital being malleable, Assumption 6 implies there is no product-mix effect. ∇

Assumption 7. *All persons are identical, facing an infinite horizon with rational expectations, a smooth, increasing and concave felicity index $u(\cdot)$ dependent on a scalar c and a time preference rate which is normalized to unity by the choice of unit for time.*

Assumption 8. *Capital can be lent and borrowed at the world market rate of interest $r(t)$.*

Now denote:

- the initial wealth as a;
- the balance of net assets abroad as $b(t)$;
- the total consumption spending as $c(t)$;
- the indirect utility index with respect to spending c and price p as $u(c;p) = u(c)$.

Then the representative individual faces an open economy Ramsey model [see e.g., Bardhan (1965) or Wan (1971)],

$$\begin{aligned}
&\max \int_{t=0}^{\infty} e^{-t} u[c(t)]dt \\
&\text{s.t.} \quad dk/dt = f[k(t); z(t)] + r(t)\, b(t) - c(t) - h(t), \\
&\qquad\quad db/dt = h(t), \\
&\qquad\qquad a = k(0) + b(0).
\end{aligned}$$

Next, introduce:

Definition 6. $\partial f[k(z); z]$ *is the subgradient at the "kinked" point:* $(k(z), f[k(z); z])$, *that is, the set of slopes of the support lines at that point.*

Introducing the Hamiltonian format with adjoint variables λ and m corresponding to the state variables k and b, the Euler-Lagrange first-order necessary conditions for optimization imply that along the equilibrium time-path,

$$\begin{aligned} e^{-t}u'(c) &= \lambda(t), \\ \lambda(t) &= m(t), \\ d\lambda/dt &\in -\lambda\,\partial f(k; z), \\ dm/dt &= -\lambda\, r(t). \end{aligned}$$

By the last three conditions,

$$r(t) \in \partial f\left[k(t); z(t)\right].$$

Thus, the domestic capital stock (hence its rate of accumulation, dk/dt) is endogenously determined by the time-path of the world interest rate $r(t)$ and parametrically dependent upon the evolution of technology $z(t)$. Thus, accumulation can in no way be regarded as an ultimate source for growth. On the other hand, to gauge the role of technology acquisition, one can consider what happens in its absence, that is, in the case of technological stagnation,

$$z(t) \equiv z(0), \quad \forall t.$$

If $r(t)$ stays constant, so must $k(t)$, and stagnation reigns. In figure 8 this stagnation is shown at k, and only with technical progress can the economy increase the capital-labor ratio to k''.

For real life relevance, one may note that in recent years, Hong Kong and Singapore have become regional financial centers and the perpetual trade surpluses of Taiwan are matched by the frequent trade deficits in Korea. Thus, in all these four NIEs, their *domestic investment*, dk/dt, need not equal *domestic savings*, $f\left[k(t); z(t)\right] - c(t)$. In fact, neither item can possibly be said to be a determinant (hence cause) of the growth rate of domestic output. Instead, both must be regarded as determined by the growth potential in investors' expectations. For further direct evidence, see also the valuable study of Itoh (1996) on the direct investment policy of Sony. The latter assigns different types of production facilities among Singapore, Malaysia, and Indonesia, based upon the relative wage rates and skill levels (mainly acquired *on the job*) at these three alternative locations.

But even in the absence of international capital mobility, the above reasoning is essentially intact: technology acquisition determines capital accumulation and not vice versa. To see this, we first replace Assumption 8 with:

Assumption 8′. *There is no international lending or borrowing:* $b(t) \equiv 0$.

We now have,

Lemma [Cass and Shell (1976)]. *Capital accumulation reaches a standstill if:*

$$1 \in \partial f[k(z); z].$$

To demonstrate that technology acquisition matters to accumulation, let:

- $z(0)$ be the level of acquiring the best practice for toy-making, and
- z^* be the same for workstations;
- $k = k[z(0)]$ and
- $k' = k(z^*)$.

Next consider two alternative scenarios:

(i) $z(t) \equiv z(0)$;

(ii) $z(t)$ increases at least to z^*.

Under (i), the above lemma implies that stagnation prevails at k. Under the latter, we have a time-dependent Ramsey model. Qualitatively speaking, once $z(t)$ reaches any level,

$$z(T) \in (z(0), z^*)\,,$$

such that,

$$1 \notin \partial f\,[k[z(T)]; z(T)]$$

then,

$$dk/dt > 0.$$

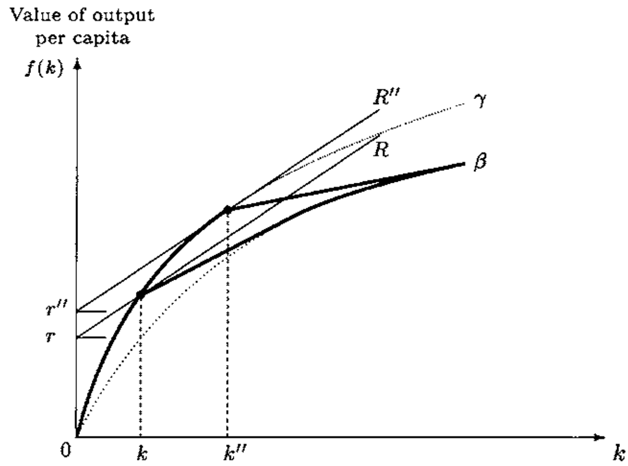


Fig. 8. Accumulation is decided by technology

This is graphically shown in figure 8.

In the historical context, in the earlier period, capital was not internationally mobile for all the NIEs. Yet, it is known that at the eve of their trade reform around 1960, both the Korean and Taiwanese economies stagnated under the classical import substitution regime, with rampant excess capacities in their industries.

Summing up, we have:

Proposition 3. *With or without international lending or borrowing, there are cases where the presence or absence of further technological gain decides whether there will be further capital accumulation.*

4. The Trade-Growth Nexus

From both conceptual reasoning and empirical evidence, trade has a dynamic effect on the growth performance of the NIEs in at least four different ways.

First, there is the production gains from trade. This is very important as intermediate goods (including capital goods, parts, and components) are prominent among the NIE trade. The comparative advantage of the NIEs are in assembling or simple fabrication but not in sophisticated fabrication or science-based material processing. Thus, when exporting to America, Korea ships Hyundai cars carrying Japanese engines by Mitsubishi and Taiwan supplied electronic components function inside Casio calculators from Japan. If Korea and Taiwan had to produce cars and calculators completely by themselves, they would not be cost-competitive for a long time to come.

In our framework, consider only productive processes involving no import nor export of any intermediate goods. Construct C^*, the convex hull of their unit-value input pairs. C^* is a shrinkage from $C: C \supset C^*$. So is ∂C^* an outward shift of ∂C, the unit-value isoquant. This will cause a downward-shift of the function $f(k)$. Together with Assumption $8'$, this would in all likelihood discourage domestic accumulation.

Second, there are consumption gains from trade which is very important for the NIEs. They enjoy the advantages to be both small [see Chipman (1965) on Mill's Paradox] and having rapidly changing comparative advantages. Thus, take Singapore for example. Freed from the need of seeking food self-sufficiency, it can develop the expertise to supply the major portion of world's demand for hard disks.

Again, under Assumption $8'$, the rate of domestic accumulation is endogenous in a Ramsey sense. Trade allows the NIEs to pursue whatever is the most profitable, without a shadow of concern about whether it is adhering to "balanced growth". Such profit should encourage accumulation.

In both of the above aspects, the "static trading gain" tends to have more than a once-and-for-all effect through its encouragement for accumulation. In the next two instances, the causal chain links trade to growth through technical progress (see figure 1): the trade gain is thus dynamic. To show how such issues may be treated formally, we sketch a model for the evolution of the generalized human capital index $z(t)$ after introducing the following notations.[2] Let:

- $d(t)$ be the vector of goods consumed under spending $c(t)$ and price $p(t)$;
- $D(t)$ be its time integral;
- $x(t)$ be the production at t;
- $X(t)$ be its time integral;
- $E(t) = (d(t), D(t), x(t), X(t))$.

Within $E(t)$, one can express the trade vector (net imports) as $d(t) - x(t)$, with its time integral being $D(t) - X(t)$. Thus various hypotheses

2. Note that here, like in Lau and Wan (1993), generalized human capital may effect different outcomes on different productive sectors.

for learning (and the principle of bounded learning) may be expressed in terms of:

Assumption 9. $dz(t)/dt = \Lambda[E(t)]$.

Remark. For example, the learning from producing output x_1 may be represented by:

$$\partial\Lambda/\partial x_1 \geq 0.$$

The principle of bounded learning may be conveyed with:

$$\partial\Lambda/\partial X_1 \leq 0, \qquad \lim_{x_1\to\infty}(\partial\Lambda/\partial x_1) = 0. \qquad \nabla$$

Third, it is reasonable to regard that the pace of learning is decided by both the capability and the intention of the learner. Capability depends on access (the basis of Findlay's contagion theory), so that *trading* with advanced economies may help the process of catching up.

Finally, the intention to learn is the strongest if one has to meet new and stringent demands to *export* to the affluent societies – the domestic customers of the developing economies are rarely demanding[3]. Moreover, once familiar to what they demand, there will be little more to be learned, by the principle of bounded learning. For a developing economy in autarky, goods affordable to the buyers are produced with mature technologies and hence providing no useful experience.

Some intuitive discussion here may help. Like Lucas (1988), most economists find it shocking that the per-capita income of one economy (say, Japan) can be higher than that of another (say, China) by a multiple of 50 times or more. After all, for most goods both produce, like rice or wine, one can hardly expect the per worker output of one to be ten times higher than the other. Of course, as in the "water and diamond" paradox, it is the "marginal goods" that matter. What Chinese can produce and Japanese cannot, most are collectors' items at best. What Japanese can produce and Chinese cannot include items most Chinese would pay a lot for. Substitute "the North" for Japan and "the South" for China, and our statements still hold, by and large. Products made and bought by "the North" typically contain "more" desirable characteristics – item per item: finer, more durable, better fabricated, of higher quality, etc. – whatever these precisely mean. In short, "Southern sensibility" aside, in the present context, more $z(t)$ means a higher ability

3. Analytically speaking, this means different income elasticities for different goods, a phenomenon which can be conveniently captured by a Stone-Geary preference as used in Basu and Van (1996).

to supply goods more universally desired (i.e. more acceptable to "the North"). Hence, the ability to export to advanced economies is a reasonable proxy for human capital. It follows then that exporting to the North can promote human capital formation. By Amsden (1977) and Morawetz (1981), the experience that counts is what is gained from competing in the markets of advanced economies.[4] Additional insight can be obtained from the following case illustrating the contagion theory.

Watanabe (1980) described the case of an American firm which assembled radios in Hong Kong. The local supervisors learned quality control and similar skills, on the job, then quit to open their own shop assembling digital watches.

Thus, learning to assemble radios is not just a matter of turning screwdrivers to put together the various imported parts, but also a matter of quality control, and so on. These make the operation competitive and thus profitable.

Those former supervisors assemble the digital watches only because they have the knowledge capital of quality control. That body of relevant information is acquired when participating in another export-oriented activity (for instance, the assembly of radios). That proves that both the transfer of technology was successful and what is transferred is more than product-specific.

The fact that those radios were exported to advanced economies is relevant because the clients in the advanced economies have tight tolerances, so that only those practising quality control can be economically viable.

By the principle of bounded learning of Young, what counts is the experience of mastering what is new, and hence typically the more challenging. This is the most needed experience of the producers of the South. It is gained by producing those goods affordable only in the North, but not the South. This is precisely why neither the tariff, nor the customs union among developing economies seem to be of much help.

As shown in the analytic study of Van and Wan (1997), the fact that those former supervisors operate assembly shops for digital watches but not radios is what encourages the foreign investor to set up the radio shop in the first place. It may suggest the transfer of technology is incomplete, and therefore "incentive compatible".

4. Specifically, in competitive markets, firms learn to supply goods of acceptable quality and at a promised date. This is true for machine tool producers by Amsden and for apparel suppliers by Morawetz. The latter showed that it makes little difference whether Columbians sell to customers at home or in Venezuela, since buyers are too forgiving in either market.

5. Concluding Remarks

This chapter is part of a larger inquiry into endogenous growth and the East Asian experience. Whereas Van and Wan (1997) scrutinizes the micro-foundations of North-South technology diffusion, the current study explores the macro-implications of this diffusion on the trade-growth-accumulation nexus.

We have reviewed the quantitative studies of East Asian growth by Kim and Lau as well as Young, drawing insights from their influential work, for example, concerning the technological gap highlighted by Kim and Lau, the concept of bounded learning by Young, etc. In fact, the *complementarity hypothesis* of Kim and Lau has played a major role in our analysis. Regarding our differences in interpreting East Asian development, it is hoped that resolutions can be found through future empirical studies of a disaggregated nature.

Perhaps this is an occasion to compare the development patterns of the OECD economies, the Asian NIEs and the former Soviet economy. The OECD economies introduce new products to the world, so did the former Soviet Union, but not the Asian NIEs. The latter fabricate products with proven market appeal, either as subcontractors for multinationals or on their own account. The NIEs do not enjoy the innovative profit but nor do they face the risk of the Soviets in investing in new capital goods of antiquated design to produce outputs which can be readily supplanted by goods of OECD origin. In a world of differentiated products, the size of market matters. By the Allyn Young externality, the OECD economies together enjoy a higher degree of division of labor than the former Soviet Union. Using the same input dosage, the OECD economies can produce not only more goods (as treated in usual textbooks), but also better-designed goods which can out-compete anything the isolated Russians could come up with. The point is, if the Allyn Young effect only implies higher productivity, its competitor can offset it by accepting lower factor rewards. But if the competitive advantage comes in the form of better design (like a Volkswagen against a Trabant) then there can be no effective defense. Because the Asian NIEs emulate OECD designs, we do not expect they can fail in the way of the Soviets.

Acknowledgement. Earlier versions of this chapter have been presented at the Chengchi University, Taipei; the Economic Institute, Academia Sinica, Nankang; the Hong Kong University of Science and Technology, Hong Kong; the Tinbergen Institute, Amsterdam; and the Denmark conference on Dynamics, Economic Growth, and International Trade. We acknowledge the helpful comments received in these sessions as well as the thoughtful discussions by Professor Bo Södersten and comments from an anonymous referee. The second author has benefited from repeated discussions with Simone Clemhout and his earlier co-operative work with Man-lui Lau.

References

Amsden, Alice H. (1989). *Asia's Next Giant: South Korea and Late Industrialization.* NY: Oxford University Press.

—— (1977). "The Division of Labor is Limited by the Type of Market: The Case of the Taiwanese Machine Tool Industry." *World Development* 5, 217–233.

Atkinson, Anthony B., and Joseph E. Stiglitz (1969). "A New View of Technological Change." *Economic Journal* 79, 573–578.

Bardhan, Pranab K. (1965). "Optimum Accumulation and International Trade." *Review of Economic Studies* 32, 241–244.

Basu, Kaushik, and Pham Hoang Van (1996). "The Economics of Child Labor." Cornell University Working Paper.

Biederman, Irving, and Margaret M. Shiffrar (1987). "Sexing Day-Old Chicks: A Case Study and Expert Systems Analysis of a Difficult Perceptual-Learning Task." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13, 640–645.

Cass, David, and Karl Shell (1976). "Introduction to Hamiltonian Dynamics in Economics." *Journal of Economic Theory* 12, 1–10.

Chipman, John S. (1965). "A Survey of the Theory International Trade: Part 1, The Classical Theory." *Econometrica* 33, 447–519.

Findlay, Ronald (1978). "Relative Backwardness, Direct Foreign Investment, and the Transfer of Technology: A Simple Dynamic Model." *Quarterly Journal of Economics* 92, 1–16.

Itoh, Motoshige (1997). "Emulative Development through Trade Expansions: East Asian Evidence". In John Piggott and Alan Woodland (eds.), *International Trade Policy and the Pacific Rim.* Basingstoke, Hampshire: MacMillan.

Kim, Jong-Il, and Laurence J. Lau (1994a). "The Sources of Economic Growth of East Asian Newly Industrialized Countries." *Journal of the Japanese and International Economies* 8, 235–271.

Kim, Jong-Il, and Laurence J. Lau (1994b). "The Sources of Economic Growth of the Newly Industrialized Countries on the Pacific Rim". In L.R. Klein and C.-T. Yu (eds.), *Economic Development of ROC and the Pacific Rim in the 1990s and Beyond.* Singapore: World Scientific Publishing.

Kim, Kihwan, and Danny M. Leipziger (1993). Korea: A Case of Government-Led Development, Monograph, Washington DC: World Bank.

Krugman, Paul R. (1994). "The Myth of Asia's Miracle." *Foreign Affairs* 73, 62–78.

Lapan, Harvey, and Pranab K. Bardhan (1973). "Localized Technical Progress." *Journal of Economic Theory* 6, 585–595.

Lau, Man-lui, and Henry Y. Wan, Jr. (1993). "Laissez Faire and Industry Targeting." *Journal of International and Comparative Economics* 2, 25–47.

Lucas, Robert E., Jr. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22, 3–42.

Morawetz, David (1981). *Why the Emperor's New Clothes are not Made in Colombia: A Case Study of Latin American and East Asian Manufacturing Exports*, NY: Oxford University Press.

Page, John M. (1994). "The East Asian Miracle: An Introduction." *World Development* 22, 615–625.

Samuelson, Paul A. (1962). "Parable and Realism in Capital Theory: The Surrogate Production Function." *Review of Economic Studies* 29, 193–206.

Solow, Robert M. (1966). "Review of Capital and Growth", *American Economic Review* 56, 1257–1260.

Van, Pham Hoang, and Henry Y. Wan, Jr. (1997). "Emulative Development through Trade Expansions: East Asian Evidence." In John Piggott and Alan Woodland (eds.), *International Trade Policy and the Pacific Rim.* Basingstoke, Hampshire: MacMillan.

Wan, Henry Y., Jr. (1971). "A Simultaneous Variational Model for International Capital Movement. In J.N. Bhagwati, R.W. Jones, R.A. Mundell, and J. Vanek (eds.), *Trade, Balance of Payments, and Growth: In Honor of Charles P. Kindleberger.* NY: North-Holland.

Watanabe, Susumu (1980). "Multinational Enterprises and Employment Oriented Appropriate Technologies in Developing Countries", ILO Working Papers 14 (Multinational Enterprise Programme).

Young, Alwyn (1995). "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." *Quarterly Journal of Economics* 110, 641–680.

— (1994). "Lessons from the East Asian NICS: A Contrarian View", *European Economic Review* 38, 964–973.

— (1993). "Invention and Bounded Learning by Doing." *Journal of Political Economy* 101, 443–472.

— (1992). "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore." In O.J. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual 1992, 13–54. Cambridge: MIT Press.

Part V

Dynamics, Economic Growth, and International Factor Movement

CHAPTER 10

Endogenous Growth and International Labor Migration: The Case of a Small, Emigration Economy

Kar-yiu Wong

1. Introduction

The present chapter raises and analyzes two issues related to economic growth and international labor migration that so far have received very little attention in the literature. The first one is the relation between economic growth and international labor migration. This issue involves some very important questions to the government planners and economists: How may emigration affect the growth of the local economy? How may emigration affect education and formation of skilled workers? How may emigration affect domestic income distribution in the long run? How may growth have a feedback effect on emigration?

While the major part of the literature on international labor migration focuses on static models, some work on analyzing international labor migration in a dynamic context has been made. For example, Rodriguez (1976), Findlay and Rodriguez (1981), and Blomqvist (1986) examine the determination of the education level; Galor (1986) investigates the effects of time preferences on the direction of labor movement; Galor and Stark (1990) compare the saving behavior of natives and foreign workers in an overlapping-generations model; Galor and Stark (1994) show how immigration may reverse the adjustment of an economy; and Shea and Woodfield (1996) derive the optimal immigration policy in a dynamic setting. However, none of these papers consider the growth effects of international labor migration, or try to answer the questions mentioned above.

The second issue raised in this chapter is how workers choose the timing and length of emigration. A worker may choose to emigrate while young or old. Alternatively, an unskilled worker may choose to emigrate now or to get education in the source country first and then emigrate as a skilled worker. A worker may also determine how long she will stay in the host country.

By explicitly considering the above decisions of workers, this chapter endogenizes the choice of the type of international labor migration. By the type of migration, we refer to: ***permanent migration***, ***brain drain***, and ***temporary migration***. Migration is said to be permanent if the migrants are permitted and plan to stay in the host country permanently, with no intention of migrating back to the source country in the foreseeable future. Temporary migrants are those who go to the host country but expect to return back to the source country in the near future, either voluntarily or involuntarily (as in the case of guest workers). Brain drain refers to the outflow of skilled and professional workers, usually permanently. These three types of migration are related not just to where people work, but also to where people get education.[1]

In the labor migration literature, usually one type of migration is analyzed at a time. This is not a complete analysis, since it is important to investigate not only whether an individual would choose to migrate, but also *when* the individual would choose to migrate and, after migration, whether to return back sometime later. Yet, as shown later in this chapter, there are cases in which an individual finds several types of migration preferable to no migration, but one of them may dominate the other types. Furthermore, this chapter also shows that as the source and host economies grow over time, the type of migration that migrant chooses may switch from one type to another.

To investigate the issues of the relationship between economic growth and international labor migration, and the endogeneity of migration decision, this chapter extends a popular education model (Uzawa, 1965; Lucas, 1988) in the recent endogenous growth literature.[2] In this model, an economy can grow perpetually without technological progress due to the unbounded accumulation of human capital. With education, workers can acquire skill and change from unskilled workers to skilled ones. If the economy is closed, workers have to get education at home, but if international labor migration is allowed, they have the option of receiving foreign education abroad. The skilled workers then have the options of working at home or in a foreign country. Different types of migration then refer to the locations where the workers choose to receive education and work, and these decisions are analyzed in a unified framework so that we can investigate how different types of labor migration may affect economic growth, and how workers choose one type of migration

1. There are other types of labor migration not considered in this chapter; for example, people that move with physical capital, and migrants that save but do not get education. Some people may choose to migrate to another country for the purpose of getting political asylum, or for retirement.

2. For a recent survey of the endogenous growth literature, see Long and Wong (1997).

over the other ones. One implication of the present model is that policies that affect emigration may have growth effects.[3]

The present chapter is organized as follows. Section 2 examines a closed economy with physical and human capital accumulation determined endogenously. The existence, uniqueness, and stability of a balanced growth path are analyzed. In section 3, international labor migration is introduced, and some of its basic features are explained. Sections 4–6 analyze the relationship between economic growth and each of the three types of migration: permanent migration, brain drain and temporary migration. Section 7 endogenizes the choice of the type of migration and shows the possibility of switching from one type of migration to another as the source country grows. Concluding remarks are given in section 8.

2. A Closed Economy

Consider an overlapping-generations economy with two factors, labor and capital, and two sectors, the production sector and the education sector. In the production sector, a homogeneous good is produced by a large number of competitive firms. In period t, $t = 0, \ldots, \infty$, the technology of the production sector can be described by the following production function,

$$Q_t = F(K_t, L_t), \tag{1}$$

where Q_t is the output, and K_t and L_t are the capital and effective labor inputs, respectively. Using subscripts to denote partial derivatives and for the time being dropping the time subindex, the properties of the production functions are:

Assumption 1. *The production function $F(K, L)$ satisfies the following conditions:*

(a) It is twice differentiable, concave, increasing, and linearly homogeneous in inputs.

(b) For all $K > 0$, F_L approaches infinity as L approaches zero.

(c) For all $K > 0$ and $L \geq 0$, F_K is bounded from below.

Part (a) of assumption 1 is standard for a neoclassical production function, and part (b) ensures that the wage rate is sufficiently large

3. Barro and Sala-i-Martin (1995, Chapter 9) suggest other models in which migration may have growth effect.

when the population approaches zero. By part (c), there always exists some incentive to save.[4]

Perfect price flexibility implies full employment of factors so that factor inputs are equal to the available stocks of factors in the economy. Define $k_t \equiv K_t/L_t$ as the capital-labor ratio. Linear homogeneity implies that the production function can be written in an alternative form:

$$Q_t = L_t f(k_t), \tag{1'}$$

where $f(k_t) \equiv F(k_t, 1)$. Cost minimization implies that factors are paid their marginal products. Thus the rental rate and wage rate are given respectively by

$$r_t = r(k_t) \equiv f'(k_t), \tag{2.1}$$

$$w_t = w(k_t) \equiv f(k_t) - k_t f'(k_t). \tag{2.2}$$

Dual to the production function is the unit cost function, $c(w_t, r_t)$, which is twice differentiable, concave, and linearly homogeneous in factor prices. Therefore $c(w_t, r_t) = 1$ describes the factor price frontier (FPF) of the economy. By Shepherd's lemma, the slope of the frontier is equal to

$$\left.\frac{\mathrm{d}r_t}{\mathrm{d}w_t}\right|_{\mathrm{FPF}} = -\frac{\partial c/\partial w_t}{\partial c/\partial r_t} = -\frac{L_t}{K_t} < 0.$$

The FPF is negatively sloped and convex to the origin.[5]

Factor services are supplied by two distinct groups of individuals: capitalists supplying capital services and workers supplying labor services. These two groups of individuals and their supplies of services are described as follows.

2.1 The Workers

Each worker lives two periods, and is called young, "unskilled" worker in the first period, and old, "skilled" worker in the second period.[6] In period t, $t = 0, \ldots, \infty$, there are N_t^u young, unskilled workers and N_t^s old, skilled workers. When the economy is closed with no international

4. An example of production function that satisfies assumption 1 is $F(K, L) = AK + K^\gamma L^{1-\gamma}$, $A > 0$, and $\gamma \in (0, 1)$.

5. For more details about the properties of the unit cost function, see Wong (1995, Chapter 2).

6. The terms "unskilled" or "skilled" are used in terms of the skill level of a worker in different periods. In each period, all young and old workers have the same skill level.

labor migration, the population is fixed over time so that $N_t^u = N_t^s = \overline{N}$ for all t, where $\overline{N}$ is a large number. Normalize the unit of time so that each individual is endowed with one unit of nonleisure time in each period. A representative young worker in period t, right after birth, inherits the average level of general knowledge (human capital) in the economy, x_t.[7] Thus all unskilled workers possess x_t efficiency units of labor. Each unskilled worker divides her time in that period between work and education. Denote the amount of time she chooses to spend on education by τ_t, and the amount of time spent on working is $1 - \tau_t$. She also takes the prevailing wage rate per efficiency unit of labor w_t as given. Education is provided for free by the government. For simplicity, no bond market is considered, and no saving by the workers exists. Thus the budget constraint of a representative unskilled worker when young is

$$C_t^y \leq (1 - \tau_t)x_t w_t, \tag{3.1}$$

where C_t^y is her consumption when young.

How human capital accumulates is the key to growth in the present economy. Following Uzawa (1965) and Lucas (1988), we postulate that the efficiency level of labor each unskilled worker has in the next period, x_{t+1}, depends positively on three factors: the initial level of average labor efficiency, x_t, the total number of educators employed in the education sector, e_t, and the time the worker spends on education, τ_t. Specifically the production of education is assumed to be

$$x_{t+1} = h(\tau_t, e_t)x_t. \tag{4}$$

Assumption 2. *Function $h(\tau, e)$ has the following properties (subscripts used to denote partial derivatives, and the time subindex being dropped for simplicity):*

(a) *$h(\tau, e) = 1$ if τ or e is zero, and $h(\tau, e) > 1$ for all $\tau, e > 0$;*

(b) *for all $e > 0$, $h_\tau, h_e > 0$, $\lim_{\tau \to 0} h_\tau > 1/\rho$, where ρ is the time discount factor for a representative worker;*

(c) *for all $e > 0$, $h_{\tau\tau} < 0$.*

7. That every young worker inherits the current average skill level is assumed for simplicity. "Depreciation" in human capital can be allowed by assuming that every young worker inherits only a fraction (less than unity) of the current average skill level. If this fraction is close to unity, perpetual growth of the economy can still exist. Unless this fraction is made endogenous, allowing depreciation of human capital does not add too much to the model.

Function (4) combines the education functions in Uzawa (1965) and Lucas (1988): the former emphasizes the educator-student ratio, while the latter considers only the fraction of time each worker spends on education.[8] However, unlike Uzawa (1965), we include the number of educators instead of the educator-student ratio in (4). The reason is that we want to capture the scale and external effects in education and productivity that have been emphasized in the migration literature; see, for example, Grubel and Scott (1966), Johnson (1967), and Miyagiwa (1991).[9] Similar approaches to modelling the education sector have also been used by Azariadis and Drazen (1990), Caballé and Santos (1993), and Bond, et al. (1996).

In period $t+1$, the representative worker earns a wage of w_{t+1} per efficiency unit of labor, but has to pay an income tax of ad valorem rate equal to ϕ_{t+1}. Therefore her budget constraint in this period is equal to

$$C_{t+1}^{o} \leq x_{t+1} w_{t+1}(1-\phi_{t+1}), \tag{3.2}$$

where C_{t+1}^{o} is her consumption when old.

All workers have identical preferences. Denote the utility of the representative worker over these two period by $u(C_t^y, C_{t+1}^o)$, which is differentiable, increasing and quasi-concave in the two consumption bundles. The worker chooses the consumption bundles in the two periods, subject to the budget constraints given by (3.1)–(3.2), to maximize her utility. The first-order condition with an interior solution is

$$\frac{u_{y,t}}{u_{o,t+1}} = \frac{h_{\tau,t} w_{t+1}(1-\phi_{t+1})}{w_t}, \tag{5}$$

where $u_{y,t}$ and $u_{o,t+1}$ are the workers' marginal utilities when young and old, respectively. Condition (5) can be solved for the optimal time devoted to education, τ_t^o, which depends on the current wage and tax rates and the next-period wage rate.

Because the workers have identical preferences, all unskilled workers in period t choose the same amount of time for education, implying that

8. Contrary to the Lucas education model in which no real resources are required (except the opportunity cost of time), papers like Razin (1972a, 1972b), Manning (1975, 1976), Hu (1976), and Rodriguez (1976) all emphasize the need of educators in education. Similarly, Ohyama (1991), and Galor and Stark (1994) specify the use of resources in education and Bond et al. (1996) assume the use of both human and physical capital in education. However, except for the models in Ohyama (1991) and Bond et al. (1996), all others do not have sustained growth.

9. Including both the number of educators and the student-educator ratio in the education function does not change our results qualitatively.

all have the same human capital level in the next period, which is given by (4). The human capital level each of the workers has is also the average level for the economy. Condition (4) thus gives the growth rate of human capital from period t to period $t+1$ as:

$$G_{x,t} = h(\tau_t, e_t) - 1.$$

2.2 The Capitalists

The next group of individuals in the economy, called capitalists, own capital, save, consume, but, under all relevant economic conditions, choose not to supply any labor services. There are M (a large number) identical capitalists in each period, each of whom is endowed with y_0 units of capital in the beginning of period 0. In period t, each capitalist earns a rental income of $r_t y_t$, saves s_t of it and consumes the rest, $c_t = r_t y_t - s_t$. The capital stock each capitalist owns in period $t+1$ is equal to

$$y_{t+1} = (1-\delta)y_t + s_t, \tag{6}$$

where δ is the rate of depreciation, which is assumed to be constant over time for all capitalists. Denote the per-period utility of each capitalist by $v(c_t)$ and the constant discount rate by $\beta \in (0,1)$.

Assuming parents with perfect bequest, the problem of each capitalist in period 0 is to choose the stream of savings to maximize the sum of discounted utilities of hers and her future generations':

$$\begin{aligned} &\max \sum_{t=0}^{\infty} \beta^t v(r_t y_t - s_t) \\ &\text{s.t.} \;\; y_{t+1} = (1-\delta)y_t + s_t, \qquad \text{for all} \;\; t = 0, \ldots, \infty. \end{aligned} \tag{7}$$

To solve problem (7), define the following Bellman equation,

$$V(y_t, t) = \max_{s_t} v(r_t y_t - s_t) + \beta V[(1-\delta)y_t + s_t, t+1], \tag{8}$$

where the investment constraint (6) has been used. The first-order condition with respect to s_t, after rearranging terms, is

$$v'_t = \beta V'_{t+1}, \tag{9}$$

where $V'_{t+1} \equiv \partial V(y_{t+1}, t+1)/\partial y_{t+1}$. Condition (9) has a nice interpretation. It is well known that V'_{t+1} represents the current shadow price of the capital stock in period $t+1$, or that of saving made in period t. The term on the left-hand side of the condition, v'_t, is the marginal utility of consumption, or the marginal disutility of saving, in period t. Condition (9) thus implies that at the optimal point, the marginal benefit of saving is equal to the marginal cost of saving. Differentiate both sides of (8) with respect to y_t to give

$$V_t' = r_t v_t' + \beta(1-\delta)V_{t+1}'. \tag{10}$$

Combining (9) and (10) together gives

$$V_{t+1}' = V_t'[\beta(1+r_t-\delta)]^{-1}, \tag{10$'$}$$

which describes how the shadow price of capital changes over time. Conditions (9) and (10) also give

$$\frac{v'(r_{t+1}y_{t+1}-s_{t+1})}{v'(r_t y_t - s_t)} = \frac{1}{\beta(1+r_{t+1}-\delta)}. \tag{11}$$

The transversality condition is

$$\lim_{t\to\infty} \beta^t V_t' y_t = 0, \tag{12}$$

which states that at $t=\infty$ at least one of the discounted shadow price of capital and the optimal capital stock is zero.

Because the capitalists are identical, the total capital stock available to the economy in period t is equal to

$$K_t = M y_t. \tag{13}$$

Conditions (6) and (13) give the growth rate of the capital stock in the economy:

$$G_{K,t} = G_{y,t} = \frac{s_t}{y_t} - \delta. \tag{14}$$

2.3 The Government and the Education Sector

The government hires skilled workers to provide education for free to all unskilled workers. For simplicity, we assume throughout this chapter that the government chooses a fixed educator-student ratio, $\alpha = e_t/N_t^u < 1$, for all $t = 0, \ldots, \infty$. This implies that in the absence of any international labor movement, the total labor supply in period t, in terms of efficiency units of labor, is equal to

$$L_t = N_t^u(1-\tau_t-\alpha)x_t + N_t^s x_t = \overline{N}(2-\tau_t-\alpha)x_t, \tag{15}$$

where the last equality is due to $N_t^u = N_t^s = \overline{N}$ for a closed economy with a constant population.

The government pays each educator the on-going wage rate for skilled workers. Therefore the education expenditure in period t is $w_t x_t e_t = \alpha w_t x_t N_t^u$. This expenditure is financed by imposing an income tax rate of ϕ_t in the same period on all skilled workers, including the educators.

A balanced government budget for period t requires that $w_t x_t e_t = \phi_t x_t w_t N_t^s$, or that

$$e_t = \phi_t N_t^s. \tag{16}$$

In other words, for a balanced government budget in that period, the government hires a fraction of the skilled workers equal to ϕ_t to educate the young workers. If the number of skilled workers is equal to the number of unskilled workers, condition (16) reduces to

$$\phi_t = \alpha, \qquad \text{for all } \ t = 0, \ldots, \infty. \tag{16$'$}$$

2.4 Balanced Growth

We now consider a balanced growth path of the economy which is defined as one on which the capital-labor ratio and thus the factor prices remain stationary, while the consumption of workers and capitalists, the physical and human capital stocks, and outputs grow over time with fixed rates.

To do that, we specify special forms of the utility functions of the workers and capitalists. First, the utility function of a worker has the form of $u(C_t^y, C_{t+1}^o) = \ln C_t^y + \rho \ln C_{t+1}^o$.[10] Utility maximization condition (5) reduces to

$$\frac{(1-\tau_t)h_\tau(\tau_t, e_t)}{h(\tau_t, e_t)} = \frac{1}{\rho}. \tag{17}$$

Let τ_t^o denote the solution to the problem in (17), which can be expressed as a function of the number of unskilled workers:

$$\tau_t^o = \tau(N_t^u; \alpha, \rho). \tag{18}$$

Lemma 1. *Given condition C2, $\tau_t^o \in (0,1)$ and is unique.*

10. Caballé and Santos (1993) show that in the Uzawa-Lucas type models, to have a balanced growth the elasticity of intertemporal substitution must be constant. The present log-linear utility function is a special case that satisfy the Caballé-Santos condition: the elasticity of intertemporal substitution is equal to unity.

Proof. See the Appendix. □

Lemma 1 and condition (17) implies that in a closed economy with a fixed population and educator-student ratio, the growth rate of human capital and thus that of the labor force do not depend on the growth rate of physical capital stock. Because of (18), for convenience, we define a reduced form of the education function: $\widetilde{h}(N_t^u;\alpha) \equiv h(\tau(N_t^u), \alpha N_t^u)$.

The dependence of τ_t^o on N_t^u can be illustrated graphically in figure 1 by schedule EE, and algebraically can be obtained by differentiating (17) and rearranging terms (the time index being dropped for simplicity):

$$\left.\frac{\mathrm{d}\tau}{\mathrm{d}N^u}\right|_{EE} = \frac{\alpha(hh_{\tau e} - h_\tau h_e)}{(1+\rho){h_\tau}^2 - hh_{\tau\tau}}. \tag{19}$$

The second-order condition for maximizing a representative worker's welfare implies that $(1+\rho){h_\tau}^2 - hh_{\tau\tau} > 0$, which is implied by assumption 2.[11]

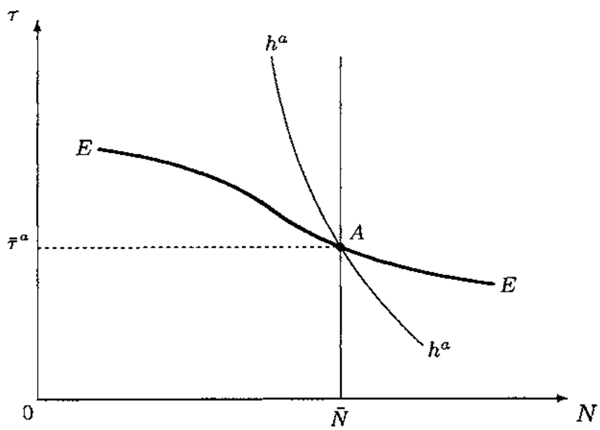


Fig. 1. Balanced path of a closed economy

Condition C1. $h_{\tau e}$ *is sufficiently small.*

11. The second-order condition is

$$u_{\tau\tau} = \frac{\rho[hh_{\tau\tau} - (1+\rho){h_\tau}^2]}{h^2} < 0.$$

Condition C1 means that the cross-effects in the education function are sufficiently small.[12] This condition and the second-order condition for optimal education time imply that $\mathrm{d}\tau/\mathrm{d}N^u$ is negative or that schedule EE is negatively sloped, as shown in figure 1. Alternatively, this means that a shrinkage of population will induce unskilled workers to spend more time on education.

When the economy is closed with a fixed population and if the government chooses a fixed educator-student ratio, the optimal education time of each unskilled worker remains constant over time:

$$\overline{\tau}^a = \tau(\overline{N}; \alpha, \rho), \tag{18$'$}$$

where "a" represents the autarkic value and a bar denotes a balanced path value of a variable. In figure 1, the balanced path is depicted by the point of intersection, A, between schedule EE and a vertical line corresponding to $\overline{N}$.

Assumption 3. $1 - \tau_t - N_t^u \tau_t' - \alpha \in (0, 1)$.

From (15), $1 - \tau_t - N_t^u \tau_t' - \alpha$ is the partial effect of an increase in the number of unskilled workers on the labor force. By assumption 3, this effect is positive.

Using the value of $\overline{\tau}^a$ obtained in (18′), the growth rate of the human capital stock and that of the efficiency units of labor are equal to

$$G_L = G_x = h(\overline{\tau}^a, \alpha\overline{N}) - 1. \tag{20}$$

By lemma 1 and assumption 2, $G_L = G_x > 0$.

We now turn to the capitalists. The per-period utility function of the capitalists is assumed to be $v(c_t) = \ln c_t$, but the capitalists and workers have different preferences if their time discount rates are different. Using this function, condition (11) reduces to

$$\frac{r_t y_t - s_t}{r_{t+1} y_{t+1} - s_{t+1}} = \frac{1}{\beta(1 + r_{t+1} - \delta)}, \tag{11$'$}$$

or, in terms of consumption, reduces to

$$c_{t+1} = c_t \beta(1 + r_{t+1} - \delta), \tag{21}$$

which describes how a representative capitalist's consumption grows over time.

12. In the present chapter, the difference between an assumption and a condition is that the former is made throughout this chapter while the latter is made for a particular result because other cases and results can usually be analyzed easily.

In a balanced path, $s_t/y_t = s_{t+1}/y_{t+1}$, and the rental rate remains constant at $\bar{r}^a$. Thus, condition (11′) reduces to

$$G_K = G_y = \beta(1 + \bar{r}^a - \delta) - 1. \tag{11''}$$

Conditions (10′) and (11″) imply that both V_t' and y_t grow at the same rate. The transversality condition as given by (12) is satisfied since $\beta < 1$.

In a balanced path, both labor and capital grow at the same rate, implying that $G_L = G_K$, or that

$$h(\bar{\tau}^a, \overline{N}) = \beta(1 + \bar{r}^a - \delta). \tag{22}$$

Proposition 1. *For the present economy under autarky, a unique balanced path with sustained growth exists.*

Proof. See the Appendix. □

2.5 Transitional Dynamics

We now examine the dynamics of the economy and determine whether its balanced path is stable. First, the consumption of a representative capitalist grows according to (21). The change over time in the capital stock owned by a capitalist as given by (6) can be rewritten as

$$y_{t+1} = (1 + r_t - \delta)y_t - c_t, \tag{23}$$

where the budget constraint of the capitalist has been used. The change in the human capital stock is obtained from (4):

$$x_{t+1} = \widetilde{h}(\overline{N})x_t, \tag{24}$$

where $\widetilde{h}(\overline{N}) \equiv h(\tau(\overline{N}), \alpha\overline{N})$, which means that when the workers choose the optimal time devoted to education, the education function can be expressed as a function of the population. Note that with given population, (24) describes the growth of human capital, while the workers choose their optimal consumption accordingly.

Using (23) and (24), the capital-labor ratio grows according to the following equation:

$$\begin{aligned} k_{t+1} &= \frac{My_{t+1}}{\overline{N}[2 - \tau(\overline{N}) - \alpha]x_{t+1}} \\ &= \frac{1 + r_t - \delta - c_t/y_t}{\widetilde{h}(\overline{N})}k_t. \end{aligned} \tag{25}$$

Equations (21), (23), (24), and (25) describe the dynamics of the system. Because of the block-recursive nature of these equations, they can be analyzed separately. First, let us define $\theta_t \equiv y_t/c_{t-1}$. Combining equations (21) and (23), with c_t expressed as a function of c_{t-1} in the former equation, the adjustment of θ_t is equal to

$$\theta_{t+1} = \frac{\theta_t}{\beta} - 1. \tag{26}$$

By equation (26), the steady-state value of θ is equal to

$$\overline{\theta} = \frac{\beta}{1-\beta}. \tag{27}$$

Equation (26) also reveals the fact that the adjustment of θ_t is not stable.[13] However, the adjustment of y_{t+1} and c_t is saddle-path stable. This requires that all capitalists choose (c_0, y_1), (c_1, y_2), and so on, so that $y_{t+1}/c_t = \overline{\theta}$ as given by (27), for all $t = 0, \ldots, \infty$.[14] Under this rule, the value of θ_t is constant over time.

Substituting the value of θ_t according to the saddle-path stable rule into (24), we get

$$y_{t+1} = \beta[1 - r(k_t) - \delta]y_t. \tag{23$'$}$$

Note that the adjustment of y_{t+1} depends on y_t and the rental rate. Making use of conditions (23′) and (24), the growth of the capital-labor ratio is

$$k_{t+1} = \frac{\beta[1 + r(k_t) - \delta]}{\widetilde{h}(\overline{N})} k_t. \tag{25$'$}$$

Differentiate both sides of (25′) and evaluate it in a small region close to the balanced path to yield

$$\mathrm{d}k_{t+1} = \left[1 + \frac{\beta r_t \varepsilon_r}{\widetilde{h}(\overline{N})}\right] \mathrm{d}k_t, \tag{28}$$

where ε_r is the elasticity of the rental rate. Note that $\beta r_t < \widetilde{h}(\overline{N})$ in the neighborhood close to the balanced path. Three different cases may exist.

13. I am indebted to Koji Shimomura for pointing out this result.

14. As long as the discount rate is not too small, choosing θ_t to be equal to $\overline{\theta}$ is the optimal consumption stream of a capitalist. It θ_t is unstable, either the capitalist's capital stock approaches zero (when k_t approaches zero) or the consumption approaches zero (when k_t approaches infinity). Either case is not optimal for them.

Case (a): $0 > \varepsilon_r > -h/(\beta r_t)$. This implies that k_t adjusts monotonically toward its balanced-path value. Case (b): $-2h/(\beta r_t) < \varepsilon_r < -h/(\beta r_t)$, in which k_t oscillates around its balanced-path value but moves asymptotically toward the latter. Case (c): $\varepsilon_r < -2h/(\beta r_t)$. In this case, the balanced path with respect to k_t is not stable.

For the purpose of the present analysis, we rule out case (c), meaning that we assume that $\varepsilon_r > -2h/(\beta r_t)$. Note that if the production function is of the Cobb-Douglas type, then we must have case (a). We summarize the above result in the following proposition:

Proposition 2. *The autarkic balanced path is saddle-path stable with respect to y_t and c_t, and stable with respect to k_t, if $0 > \varepsilon_r > -2h/(\beta r_t)$ in the neighborhood close to the balanced path. If the production function is of the Cobb-Douglas type, k_t adjusts monotonically toward its balanced-path value.*

3. International Labor Migration

To analyze international migration, let us consider two economies which have the structures described in the previous section. These two economies are labelled the source (emigration) country and the host (receiving) country, with labor possibly flowing from the former country to the latter country (hence their names). Since the focus of the present chapter is on the effects of emigration, we assume that the source country is small in the sense that the amount of labor movement is too small to affect the autarkic equilibrium of the host economy.

International labor migration from the source country to the host country is allowed by both governments in the beginning of period 0. We further assume that before international labor migration is allowed, both countries are on their balanced paths. The variables of the host country are denoted by an asterisk while that of the source country have no asterisks; for example, $\overline{w}^*$ and $\overline{e}^*$ denote the wage rate and the number of educators in the host country, respectively.

International labor migration can take place costlessly, and the decision to migrate depends exclusively on what life-time utility a worker gets from either country. To focus our analysis on labor migration, we follow the tradition in the literature and assume no movement of capital or capitalists.

The following three different types of labor emigration are considered in the present chapter:

1. *Permanent Migration* – Unskilled workers in the source country move to the host country. They then receive education, work and

then die there.

2. *Brain Drain* – Unskilled workers in the source country receive education and work when young in their country but move abroad after graduation and work in the host country when old (and die there).

3. *Temporary Migration* – Unskilled workers in the source country move to the host country when young and get education there. When old, they return to the source country and work there as skilled workers.

Young workers in the source country inherit the prevailing average human capital stock in their country. If they decide to receive education in the host country, they will bring this human capital stock with them and receive education there. In the next period they become skilled workers, and then choose where they work.

If unskilled workers determine to receive education in the host country, they may choose a different amount of time for education, which can be expressed in terms of the following condition

$$\overline{\tau}^* = \tau^*(\overline{e}^*; \rho), \tag{18''}$$

where $\overline{e}^*$ is the number of educators in the host country.

Education is financed in the host country in the following way. It is provided for free by the host government and financed by a constant ad valorem income tax of ϕ^* on skilled workers. This assumption, which is consistent with the observations in many countries, means that if an unskilled worker in the source country migrates permanently to the host country, she will pay an education tax in the same way as a worker in the host country.[15] If, however, the migration is temporary, she will come back after graduation and then pay an education tax in the source country. If brain drain occurs, the worker receives a free education in the source country but has to pay an education tax when working in the host country.

Consider a representative unskilled worker in period t in the source country facing the options of permanent migration, brain drain, temporary migration, and no migration. Denote the corresponding normalized (with respect to the effective labor of a young worker) utility of the

15. Education is usually heavily subsidized in many countries. In some countries such as the United States, governments provide virtually free basic education even to the children of legal and illegal immigrants. The advantage of this assumption is that we do not have to consider explicitly how an individual finances her education.

worker if one of these options is chosen by U_t^p, U_t^b, U_t^t, and U_t^o, respectively. These four utility levels are defined as follows:

$$U_t^p = \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho\ln[\overline{h}^*(1-\phi^*)\overline{w}^*], \tag{29.1}$$
$$U_t^b = \ln[(1-\tau(N_t^u))w_t] + \rho\ln[h(\tau(N_t^u), e_t)(1-\phi^*)\overline{w}^*], \tag{29.2}$$
$$U_t^t = \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho\ln[\overline{h}^*(1-\phi_t)w_{t+1}], \tag{29.3}$$
$$U_t^o = \ln[(1-\tau(N_t^u))w_t] + \rho\ln[h(\tau(N_t^u), e_t)(1-\phi_t)w_{t+1}], \tag{29.4}$$

where $\overline{h}^* = h^*(\overline{\tau}^*, \overline{e}^*)$.

These three types of labor migration are described in detail in the next three sections. For simplicity, we first assume that only one type of migration is allowed, meaning that workers in the source country can choose between one type of migration and no migration. This assumption will be relaxed in section 7. Note that by the nature of permanent migration, U_t^p depends on the variables in the host country only, and similarly U_t^o depends on variables in the source country only.

4. Permanent Migration

This section analyzes the effects of permanent migration, assuming that brain drain and temporary migration are not permitted. This means that the governments of the source and host countries allow movement of unskilled workers, but not skilled workers, from the source country to the host country. Those who move then become citizens of the host country and are not allowed to return back. The no-return assumption, which will be relaxed later, allows us to focus on examining the features of permanent migration.

4.1 Features of Permanent Migration

Suppose that in the beginning of period 0 both countries are growing along their balanced paths, and that permanent migration of unskilled workers from the source country is allowed by both governments. Using the definition of the utility levels defined in (29.1)–(29.4), permanent migration exists in periods $t \geq 0$ if and only if $U_t^o < U_t^p$, or

$$\begin{aligned}&\ln[(1-\tau_t)w_t] + \rho\ln[h(\tau_t, e_t)(1-\phi_t)w_{t+1}] < \\ &\ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho\ln[\overline{h}^*(1-\phi^*)\overline{w}^*].\end{aligned} \tag{30}$$

In this section, we assume that condition (30) is satisfied in period 0 before any movement of labor.

An alternative, but perhaps more intuitive, way of expressing condition (30) is

$$\ln\{(1-\tau_t)\,[(1-\phi_t)h(\tau_t,e_t)]^\rho\} + \ln w_t + \rho \ln w_{t+1} < \\ \ln\{(1-\overline{\tau}^*)[(1-\phi^*)\overline{h}^*]^\rho\} + (1+\rho)\ln \overline{w}^*. \tag{30'}$$

Condition (30′) highlights the causes of permanent migration: a better and more effective education and/or higher wages.

Permanent migration has the following features. First, it represents a shrinkage of the population size because migrated workers do not come back. The population of the source country then becomes an endogenous variable to be determined in the present model. The change in the country's population has substantial effects on the economy, because both the physical stock, saving, output, factor prices, and the growth rates of capital stocks are generally affected by migration. Second, since only unskilled workers are allowed to move out, the number of unskilled workers in the source country in any period is usually less than the number of skilled workers. This affects the amount of education provided by the source government. In particular, the education finance burden shared by skilled workers tends to fall.

We now formally analyze these features of permanent migration. Denote the numbers of unskilled workers (after emigration) and skilled workers (including educators) in the source country in period t by N_t^u and N_t^s, respectively. The non-emigrating unskilled workers will spend part of their nonleisure time on working and the rest on education. The dependence of these workers' optimal education time on the number of unskilled workers is again given by (18). All these unskilled workers become skilled workers in the next period. Therefore $N_t^u = N_{t+1}^s$. The effective labor force in period t is equal to

$$\begin{aligned} L_t &= [N_t^s + N_t^u(1-\tau(N_t^u)-\alpha)]x_t \\ &= [N_{t-1}^u + N_t^u(1-\tau(N_t^u)-\alpha)]x_t. \end{aligned}$$

The capital-labor ratio is equal to

$$k_t = \frac{My_t}{[N_{t-1}^u + N_t^u(1-\tau(N_t^u)-\alpha)]x_t},$$

which implies the following way of adjustment of the capital-labor ratio over time:

$$k_{t+1} = \frac{[N_{t-1}^u + N_t^u(1-\tau(N_t^u)-\alpha)]\beta[1+r(k_t)-\delta]}{[N_t^u + N_{t+1}^u(1-\tau(N_{t+1}^u)-\alpha)]\widetilde{h}(N_t^u)}\,k_t, \tag{31}$$

where it is assumed that physical capital accumulates along the saddle path as described above.

The cost of education is equal to $e_t w_t x_t = \alpha N_t^u w_t x_t$, with the educator-student ratio given. This cost is financed by imposing an income tax on the skilled workers, which generates a revenue of $N_t^s w_t x_t \phi_t$. A balanced budget means that

$$\phi_t = \frac{\alpha N_t^u}{N_t^s} = \frac{\alpha N_t^u}{N_{t-1}^u}. \tag{32}$$

4.2 Balanced Growth

We first analyze the balanced path of the source country before turning to its dynamics. In a new balanced path, the economy is smaller than before in terms of population, but the outflow of workers stops.[16] Denote its new number of unskilled and skilled workers in each period by $\overline{N}^p$. Assume that the source government adopts the same education policy (such as the educator-student ratio) before and after labor migration. Thus the number of educators in each period is equal to $\alpha\overline{N}^p$.

The equilibrium condition for permanent migration is

$$\begin{aligned}\ln[(1-\overline{\tau}^p)\overline{w}^p] + \rho\ln[h(\overline{\tau}^p, \alpha\overline{N}_t^p)(1-\overline{\phi})\overline{w}^p] = \\ \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho\ln[\overline{h}^*(1-\phi^*)\overline{w}^*],\end{aligned} \tag{33}$$

which states that the marginal unskilled worker in the source country gets the same utility whether she migrates or not. In condition (33), since the population is stationary, $\overline{\phi} = \alpha$.

Permanent migration also affects the education time the remaining unskilled workers choose to spend, as lemma 1 suggests, and it is given by condition (18′):

$$\overline{\tau}^p = \tau(\overline{N}^p). \tag{34}$$

Making use of (34), condition (20) gives the growth rate of human capital. The growth rate of physical capital is obtained from (11″). In a balanced growth path, both human capital and physical capital grow at the same rate. This gives the growth-rate-equalization condition:

$$\widetilde{h}(\overline{N}^p) = \beta(1 + \overline{r}^p - \delta). \tag{35}$$

16. If the source country has a positive growth rate of population (instead of zero), then the rate of emigration under a new balanced path can be positive, but it cannot be greater than the population growth rate.

The wage rate and rental rate are on the factor price frontier of the economy:

$$c(\overline{w}^p, \overline{r}^p) = 1. \tag{36}$$

Equations (33) to (36) can be solved for the four unknowns: $\overline{\tau}^p$, $\overline{N}^p$, $\overline{w}^p$, and $\overline{r}^p$.

To determine this new balanced growth more explicitly, let us solve the factor price frontier (FPF) to give $w = \zeta(r)$ (time subindices dropped for simplicity). Substitute the value of $\overline{r}^p$ obtained from condition (35) into the wage function to give:

$$w = \zeta[\widetilde{h}(N)/\beta + \delta - 1],$$

where $\zeta' \equiv \mathrm{d}w/\mathrm{d}r < 0$ is the reciprocal of the slope of the FPF.

Let us define the following functions:

$$\begin{aligned} P(\tau, N) &= \ln(1-\tau) + \rho \ln[(1-\phi)h(\tau, \alpha N)] \\ &\quad + (1+\rho)\ln\{\zeta[h(\tau, \alpha N)/\beta + \delta - 1]\}, \end{aligned} \tag{37.1}$$

$$P^* = \ln(1-\tau^*) + \rho \ln[(1-\phi^*)\overline{h}^*] + (1+\rho)\ln \overline{w}^*. \tag{37.2}$$

Function $P(\tau, N)$ represents the normalized (with respect to the effective labor of a young worker) life-time welfare of a representative worker in a balanced path (with stationary wage rate and equalization of the growth rates of both types of capital). In the autarkic balanced path, the value of this function is $P(\overline{\tau}^a, \overline{N})$. The variable P^* represents the normalized welfare a worker gets from the host country if she chooses to migrate.

The partial derivatives of function $P(\tau, N)$ are equal to

$$\begin{aligned} P_\tau &= \frac{(1+\rho)\zeta' h_\tau}{\beta w} < 0, \\ P_N &= \frac{\alpha \rho h_e}{h} + \frac{\alpha(1+\rho)\zeta' h_e}{\beta w}, \end{aligned}$$

where in calculating P_τ it is assumed that τ is always chosen optimally by the workers. The sign of P_N is ambiguous.

Condition C2. *$|\zeta'(r)|$ is sufficiently large.*

If condition C2 holds, $P_N < 0$. Function $P(\tau, N)$ can be represented by different schedules in figure 2. The slope of a schedule such as PP (with τ close to the optimal education time) is equal to

$$\left.\frac{\mathrm{d}\tau}{\mathrm{d}N}\right|_{PP} = -\frac{P_N}{P_\tau} = -\frac{\alpha[\beta w \rho + (1+\rho)h\zeta']h_e}{(1+\rho)h\zeta' h_\tau}. \tag{38}$$

The sign of the slope of schedule PP is ambiguous, but if condition C2 is satisfied, the schedule is negatively sloped. In the diagram, schedule PP gives the combinations of (τ, N) that satisfy

$$P(\tau, N) = \ln[(1 - \overline{\tau}^*)\overline{w}^*] + \rho \ln[(1 - \phi^*)\overline{h}^*\overline{w}^*],$$

which represents the normalized welfare of a worker under a balanced path when permanent migration is allowed.

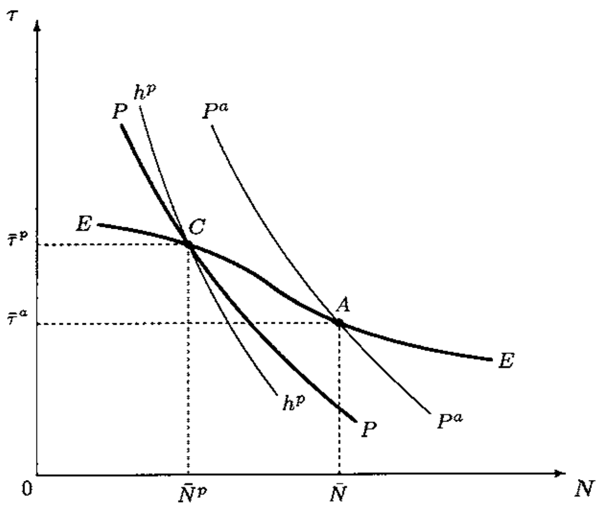


Fig. 2. Balanced path with permanent migration

Figure 2 also shows schedule EE which represents function $\overline{\tau} = \tau(N^u)$ as given by (18). The intersecting point, point B, between schedules PP and EE thus represents the values of N and τ that satisfy conditions (33) to (36), and thus represents the balanced path of the source country in the presence of permanent migration.

Proposition 3. *Given condition (30), a balanced path under permanent migration with a positive population in the source country exists. If conditions C1 and C2 are further satisfied, the balanced path is unique.*

Proof. See the Appendix. □

For the sake of comparison, the autarkic balanced growth, which is represented by point A, is also shown in figure 2. Since permanent migration causes a drop in the population of the source country, $\overline{N}^p < \overline{N}$. This means that point A is below point B. As long as schedule EE is negatively sloped (as under condition C1), permanent migration leads to not just a shrinkage of population, but also an increase in the time unskilled workers choose to spend on education. The reason for the rise in education time in the present model is that with a smaller population and a smaller size of the student body, the government will employ less skilled workers to educate the students, causing the unskilled workers to spend more time on education. It is clear from the diagram that conditions C1 and C2 are stronger than what is needed for a unique balanced path.

We now examine how permanent migration may affect the growth rate and factor prices. Since the growth rate of human capital stock is equal to $h(\tau, \alpha N) - 1$, different growth rates can be shown by different iso-growth-rate contours in figure 2. The diagram shows two schedules passing through points A and B, $h^a h^a$ and $h^p h^p$, respectively. The slope of a representative schedule hh is equal to

$$\left.\frac{\mathrm{d}\tau}{\mathrm{d}N}\right|_{hh} = -\frac{\alpha h_e}{h_\tau} < 0. \tag{39}$$

By conditions (19) and (39), if $h_{\tau\tau} < 0$ and if condition C1 holds, then both schedules EE and hh are negatively sloped, but the latter is steeper, as shown in figure 2.

The effect of a change in population on the growth rate of the source country can be measured by

$$\begin{aligned}\frac{\mathrm{d}h(\tau, N)}{\mathrm{d}N} &= h_\tau \left.\frac{\mathrm{d}\tau}{\mathrm{d}N}\right|_{EE} + \alpha h_e \\ &= \alpha \frac{h h_\tau h_{\tau e} + \rho h_e (h_\tau)^2 - h h_e h_{\tau\tau}}{(1+\rho)(h_\tau)^2 - h h_{\tau\tau}}.\end{aligned} \tag{40}$$

The total effect of the population change on function $h(\tau, N)$ is in general ambiguous. However, if condition C1 is satisfied, the expression in (40) is positive. In terms of figure 2, this means that $h^p h^p$ and EE are both negatively sloped with the former schedule being steeper, and a decrease in population due to permanent emigration hurts the growth of the source country.

The intuition behind this result is that permanent migration has generally two opposing effects on the growth rate: a positive one due to an increase in the time unskilled workers spent on education, and a negative one due to a decrease in population and thus the number of educators the government hires to educate. The negative effect is based on the assumption that the government also keeps a constant educator-student ratio. The government can reduce or even reverse the negative effect of permanent migration by increasing the educator-student ratio, but such a policy is not considered in the present chapter.

The new rental rate can be obtained from condition (35). A decrease in the growth rate thus lowers the rental rate, which in turn implies a rise in the wage rate and the capital-labor ratio. It is interesting to note that qualitatively the effects of permanent migration on factor prices and the capital-labor ratio obtained in the present model are similar to those derived from a static model. These results are summarized by the following proposition.

Proposition 4. *Suppose that condition C1 holds. A drop in the population due to permanent emigration slows down the growth of the source country, lowers the rental rate but raises the wage rate, capital-labor ratio and education time.*

4.3 Transitional Dynamics

We now examine the dynamics of permanent migration. As under autarky, the adjustment of the capitalists' consumption and their capital stocks is saddle-path stable. So we assume that they follow the saddle path as described above. The adjustment of the rest of the source economy depends on the outflow of unskilled workers. We first consider the case in which in each period unskilled workers in the country are allowed to move out costlessly and instantaneously to equalize the effective income in both countries. This means that labor flows in each period until condition (30) or (30′) with an equality is satisfied. This condition and condition (31) then describe a system of second-order difference equations in N_t^u and k_t.

To examine local stability, we linearize equations (30) and (31) around the balanced path and define $z_{t+1} = N_t^u = N_{t+1}^s$. Let a "tilde" denote the deviation of a variable from its balanced-path value; for example, $\widetilde{k}_t \equiv k_t - \overline{k}$. Differentiating equations (30), with an equality, and (31) gives

$$\begin{bmatrix} A & B & C \\ 0 & 0 & 1 \\ F & 0 & G \end{bmatrix} \begin{bmatrix} \widetilde{k}_{t+1} \\ \widetilde{N}_{t+1}^u \\ \widetilde{z}_{t+1} \end{bmatrix} = \begin{bmatrix} D & 0 & E \\ 0 & 1 & 0 \\ H & 0 & J \end{bmatrix} \begin{bmatrix} \widetilde{k}_t \\ \widetilde{N}_t^u \\ \widetilde{z}_t \end{bmatrix}, \tag{41}$$

where the coefficients, when evaluated close to the balanced path, are equal to

$$A = \overline{N}^p(2 - \overline{\tau}^p - \alpha)\overline{h}^p,$$

$$
\begin{aligned}
B &= \overline{k}^p\overline{h}^p(1 - \overline{\tau}^p - \overline{N}^p\overline{\tau}^{p\prime} - \alpha), \\
C &= \overline{k}^p\overline{N}^p(2 - \overline{\tau}^p - \alpha)\overline{h}^{p\prime} + \overline{k}^p\overline{h}^p(\overline{\tau}^p + \overline{N}^p\overline{\tau}^{p\prime} + \alpha), \\
D &= \overline{N}(2 - \overline{\tau}^p - \alpha)(\overline{h}^p + \beta\overline{k}^p\overline{\tau}^{p\prime}), \\
E &= \overline{k}^p\overline{h}^p, \\
F &= \rho\overline{w}^{p\prime}/\overline{w}^p, \\
G &= \alpha\rho(\overline{h}^p_e/\overline{h}^p - 1/[(1-\alpha)\overline{N}^p]), \\
H &= -\overline{w}^{p\prime}/\overline{w}^p, \\
J &= -\alpha\rho/[(1-\widetilde{\phi})\overline{N}^p],
\end{aligned}
$$

where $\overline{\tau}^{p\prime} \equiv \mathrm{d}\tau_t/\mathrm{d}N_t^u$ (slope of schedule EE), $\overline{h}^{p\prime} \equiv \mathrm{d}\widetilde{h}_t/\mathrm{d}N_t^u$, and $\overline{w}^{p\prime} \equiv \mathrm{d}\overline{w}_t/\mathrm{d}k_t$. All the endogenous variables in the above expressions are evaluated at the balanced path, with their values distinguished by a bar and a superscript "p."

Inverting the matrix in (41) and rearranging terms give

$$
\begin{bmatrix} \widetilde{k}_{t+1} \\ \widetilde{N}^u_{t+1} \\ \widetilde{z}_{t+1} \end{bmatrix} = \mathcal{H} \begin{bmatrix} \widetilde{k}_t \\ \widetilde{N}^u_t \\ \widetilde{z}_t \end{bmatrix} \tag{42}
$$

with

$$
\mathcal{H} = \begin{bmatrix} H/F & -G/F & J/F \\ D/B - AH/BF & AG/BF - C/B & E/B - AJ/BF \\ 0 & 1 & 0 \end{bmatrix}.
$$

Proposition 5. *A balanced path under permanent migration is locally stable with respect to k_t and N_t^u if all the eigenvalues of the matrix in (42) are less than unity in magnitude.*

The proof of proposition 5 is based on (42) and is omitted. To find more explicit conditions in terms of technologies for a stable balanced path is more difficult. However, one possible case of instability is given below.

Condition C3. *$w'(k)$ is sufficiently large.*

Condition C3 means that the wage rate is very sensitive to a change in the capital-labor ratio, or to the labor movement.[17] This condition is satisfied if the production function is of the Cobb-Douglas type and if

17. For example, consider the production function given in footnote 3: $F(K, L) = AK + K^{\gamma}L^{1-\gamma}$. Condition C3 is satisfied if the capital-labor ratio is small.

the capital-labor ratio is sufficiently small. The latter condition is more relevant for developing countries with a small capital stock.

Let us evaluate the determinant of the matrix in (42). It is equal to

$$\frac{DJ - EH}{BF} = -\frac{1}{\rho(1 - \overline{\tau}^p - \overline{N}^p \overline{\tau}^p - \alpha)} - \frac{\alpha \overline{w}^p (2 - \overline{\tau}^p - \alpha)(\overline{h}^p + \beta \overline{k}^p \overline{r}^{p\prime})}{\overline{k}^p \overline{h}^p (1 - \phi)(1 - \overline{\tau}^p - \overline{N}^p \overline{\tau}^p - \alpha) \overline{w}^{p\prime}}.$$

If condition C3 is satisfied and if $\tau^{p\prime}$ is sufficiently small, then the determinant given above is greater than unity. In other words, at least one of the eigenvalues of the matrix in (42) is greater than unity in magnitude, meaning that the system is not stable.

If the system is not stable when the unskilled workers flow out in the way described above, the government can direct the economy toward the new balanced path by regulating emigration. We now introduce one way. Suppose that the government permits outflow of unskilled workers according to the following formula: $N_t^u = b^{t+1}\overline{N}$ where $\overline{N}$ is the initial population of unskilled workers and $0 < b < 1$ but is close to unity. Substitute the number of unskilled workers into (31), and after simplifying terms we have

$$\begin{aligned} &b\{1 + b[1 - \tau(b^{t+2}\overline{N}) - \alpha]\}\widetilde{h}(b^{t+1}\overline{N})k_{t+1} = \\ &\quad \{1 + b[1 - \tau(b^{t+1}\overline{N}) - \alpha]\}\beta[1 + r(k_t) - \delta]k_t. \end{aligned} \tag{43}$$

Equation (43) represents a first-order difference equation in k_t. By this equation, $\mathrm{d}k_{t+1}/\mathrm{d}k_t > 1$ if r_t' is not too significant when w_t' is sufficiently large, as with a Cobb-Douglas production function and a small capital-labor ratio. This means that the balanced path is not stable as long as the source government follows the above emigration rule.

However, the source government will not follow this emigration rule forever. When the remaining number of unskilled workers is equal $\widetilde{N}_p$, no more emigration is allowed. In this case, the population of the source country stays stationary. Proposition 2 states that with a fixed population and if r_t' is not too significant (as in the case with a Cobb-Douglas production function), the system is stable, at least in the neighborhood of the balanced path.[18]

The adjustment of the capital-labor ratio under the above emigration rule can be illustrated in figure 3. Point A is the initial equilibrium point under autarky. If emigration is now allowed under the above rule,

18. In the present framework, it is difficult to determine whether the balanced path is stable starting from the point at which emigration is first prohibited.

the relationship between k_{t+1} and k_t is described by schedule ABC. The path with arrows represents the adjustment of the capital-labor ratio, but this adjustment path is not a stable one. Suppose that when point B is reached, the population of the source country drops down to $\overline{N}^p$. The government stops the emigration. Because the population is fixed, the capital-labor ratio could then adjust along a path as represented by schedule DBE until point P is reached.

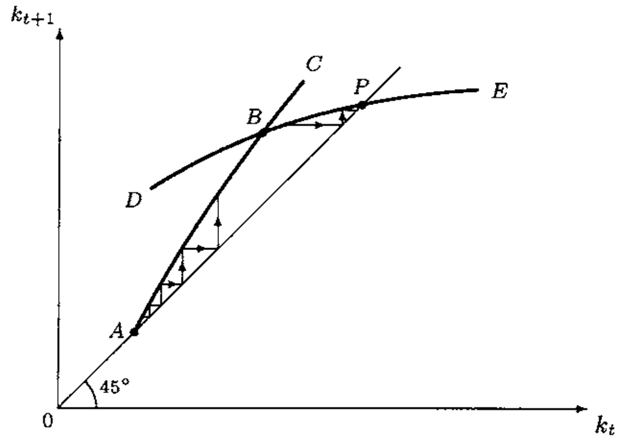


Fig. 3. Adjustment under permanent migration

5. Brain Drain

This section analyzes the effects of brain drain. Once again, we follow the above strategy to simplify the present analysis in this section by assuming that only brain drain, but not permanent migration or temporary migration, is allowed.

5.1 Features of Brain Drain

In period $t \geq 0$, brain drain is allowed. It exists if and only if $U_t^b > U_t^o$. Define $W_b^* \equiv (1 - \phi^*)\overline{w}^*$, which is the after-tax income for a skilled worker in the host country. Therefore the necessary and sufficient condition for the existence of brain drain reduces to

$$(1 - \phi_t)w_t < W_b^*. \tag{44}$$

In the present dynamic model, brain drain has several features that we usually do not find in a static setting. First, the rise in the domestic wage rate increases the cost of education. Second, there are less skilled workers to support the education which is provided for free by the government of the source country. Third, since the emigrants leave and never return, their children will be the residents of another country. There is thus a shrinkage in the population of the source country, as in the case of permanent migration. Fourth, brain drain may affect the growth of the source/host country.

Let us analyze these features rigorously. Define N_t^u and N_t^s as the numbers of unskilled and non-emigrating skilled workers (including educators), respectively, in the source country in period t. With brain drain, generally $N_t^u \neq N_t^s$. The stock of effective labor in period t is equal to

$$L_t = \{N_t^s + N_t^u[1 - \tau(N_t^u) - \alpha]\}x_t. \tag{45}$$

Since the unskilled workers in period t are the children of the skilled worker that have not emigrated in the previous period, $N_t^u = N_{t-1}^s$. Condition (45) can be written as

$$L_t = \{N_t^s + N_{t-1}^s[1 - \tau(N_{t-1}^s) - \alpha]\}x_t, \tag{45'}$$

which describes how the effective labor stock depends on the change in the population over time. The number of emigrants in period t is equal to $N_{t-1}^u - N_t^s$.

The movement of skilled workers affects the government budget. The education expenditure in period t is $e_t w_t x_t = \alpha N_t^u w_t x_t$, while the revenue is $N_t^s w_t x_t \phi_t$, where a constant educator-student ratio is assumed. The condition for a balanced budget is

$$\phi_t = \frac{\alpha N_t^u}{N_t^s}. \tag{46}$$

Thus when some skilled workers flow out, making the number of remaining skilled workers less than that of the unskilled workers, we must have at least one of the followings: (i) an increase in the tax rate, ϕ_t; (ii) a decrease in the educator-student ratio; and (iii) a government budget deficit. The first outcome has the effect of encouraging even more outflow of skilled workers, the second outcome lowers the education quality and the rate of human capital accumulation, and the third outcome could imply economic, social and political costs.

Substitute (46) into (44) to give

$$\left(1 - \frac{\alpha N_t^u}{N_t^s}\right) w_t < W_b^*, \tag{44'}$$

which shows the condition for brain drain.

Condition (45′) implies that the capital-labor ratio in period t is equal to

$$k_t = \frac{My_t}{\{N_t^s + N_{t-1}^s[1 - \tau(N_{t-1}^s) - \alpha]\}x_t}. \tag{47}$$

5.2 Balanced Growth

We now focus on a balanced path of the source country under brain drain. In this case, a balanced path is defined as one along which the factor prices and population are stationary. This requires that outflow of skilled workers no longer exists, giving the same number of skilled and unskilled workers in each period, and that both human capital and physical capital are growing at the same rate. Distinguish the variables in a balanced path by a bar and a superscript "b." For example, the number of skilled or unskilled workers in a balanced path is denoted by $\overline{N}^b$. The equilibrium condition under brain drain is then

$$(1 - \overline{\phi})\overline{w}^b = W_b^*. \tag{48}$$

Note that with the same number of unskilled and skilled workers, the income tax rate is $\overline{\phi} = \alpha$. The optimal education time is obtained from (18′):

$$\overline{\tau}^b = \tau(\overline{N}^b; \rho, \alpha). \tag{49}$$

The corresponding growth rate of the human capital stock is equal to $h(\overline{\tau}^b, \alpha\overline{N}^b) - 1$. In a balanced path, physical capital and human capital grow at the same rate, implying that

$$h(\overline{\tau}^b, \alpha\overline{N}^b) = \beta(1 + \overline{r}^b - \delta). \tag{50}$$

The factor price frontier again gives the relationship between the wage rate and the rental rate:

$$c(\overline{w}^b, \overline{r}^b) = 1. \tag{51}$$

Equations (48) to (51) can be used to solve for the number of unskilled and skilled workers, the optimal education time, and the factor prices along a balanced path.

Proposition 6. *Given (44) under autarky and condition C1, a unique balanced growth under brain drain exists.*

Proof. See the Appendix. □

Proposition 6 can be illustrated in figure 4. In terms of figure 4, let schedule $h^b h^b$ represent the locus $h(\tau, N) = h(\overline{\tau}^b, \overline{N}^b)$. By (44), schedule $h^b h^b$ is below (above) schedule EE when $N = \overline{N}$ ($N = 0$). Continuity of the schedules implies that they cut each other at least once. By condition C1, both schedules are negatively sloped, but schedule $h^b h^b$ is steeper at the point of intersection. So they intersect only once, at point B in the figure.

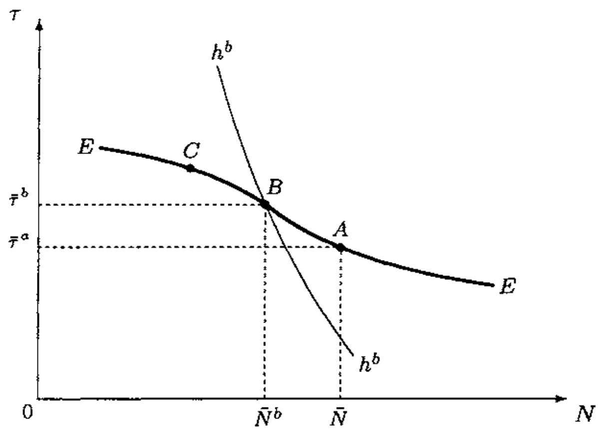


Fig. 4. Balanced path with brain drain

We now determine how brain drain may affect the growth rate of the source country. The effect of a change in population on the growth rate is given by condition (40) or figure 4.

Proposition 7. *Given condition C1, a drop in the population due to brain drain slows down the growth of the source country, lowering the rental rate but raising both the wage rate, capital-labor ratio and education time.*

Proof. See the Appendix. □

5.3 Transitional Dynamics

The adjustment of the system depends on the rate of emigration. We assume that in each period, skilled workers in the source country can move instantaneously and costlessly to the host country to equalize the after-tax incomes in the two countries, i.e., to satisfy condition (44′) with an equality:

$$\left(1 - \frac{\alpha N^s_{t-1}}{N^s_t}\right) w(k_t) = W^*_b. \tag{52}$$

Condition (52) describes the relationship between the capital-labor ratio and the change in population. The change in the capital-labor ratio along the saddle path described earlier can be obtained from (47), (23′), and (24):

$$\begin{aligned} &k_{t+1}\{N^s_{t+1} + N^s_t[1 - \tau(N^s_t) - \alpha]\}\widetilde{h}(N^s_{t-1}) = \\ &\quad k_t\{N^s_t + N^s_{t-1}[1 - \tau(N^s_{t-1}) - \alpha]\}\beta[1 + r(k_t) - \delta]. \end{aligned} \tag{53}$$

Equations (52) and (53) describe the changes in k_t and N^s_t over time. Alternatively, using (52) k_t can be expressed in terms of N^s_t and N^s_{t-1}, and k_{t+1} in terms of N^s_{t+1} and N^s_t, which can be substituted into (53). The latter becomes a second-order difference equation in N^s_t.

We now consider local stability. Let us use a "tilde" to denote the deviation of a variable from its balanced-path value; for example, $\widetilde{N}^s_t \equiv N^s_t - \overline{N}^b$ and $\widetilde{k}_t \equiv k_t - \overline{k}^b$. Linearizing condition (53) around the balanced path, we have

$$\widetilde{A}\widetilde{k}_{t+1} + \widetilde{B}\widetilde{N}^s_{t+1} + \widetilde{C}\widetilde{N}^s_t + \widetilde{D}\widetilde{k}_t + \widetilde{E}\widetilde{N}^s_{t-1} = 0, \tag{54}$$

where the coefficients, evaluated at the balanced path, are

$$\begin{aligned} \widetilde{A} &= \overline{N}^b(2 - \overline{\tau}^b - \alpha)\overline{h}^b, \\ \widetilde{B} &= \overline{k}^b\overline{h}^b, \\ \widetilde{C} &= -\overline{k}^b\overline{h}^b(\overline{\tau}^b + \overline{N}^b\overline{\tau}^{b\prime} + \alpha), \\ \widetilde{D} &= -\overline{N}^b(2 - \overline{\tau}^b - \alpha)[\overline{h}^b + \beta\overline{k}^b\overline{r}^{b\prime}], \\ \widetilde{E} &= \overline{k}^b\overline{h}^{b\prime}\overline{N}^b(2 - \overline{\tau}^b - \alpha) - \overline{k}^b\overline{h}^b(1 - \overline{\tau}^b - \overline{N}^b\overline{\tau}^{b\prime} - \alpha). \end{aligned}$$

Equation (52) can be linearized in the same way:

$$\widetilde{k}_t = \widetilde{F}(\widetilde{N}^s_t - \widetilde{N}^s_{t-1}), \tag{55}$$

where $\widetilde{F} = -\alpha\overline{w}^b[(1-\alpha)\overline{w}^{b\prime}\overline{N}^b]^{-1}$. Substitute (55), for both $\widetilde{k}_{t+1}$ and $\widetilde{k}_t$, into (54) to yield

$$a_0\widetilde{N}^s_{t+1} + a_1\widetilde{N}^s_t + a_2\widetilde{N}^s_{t-1} = 0, \tag{56}$$

where $a_0 \equiv \widetilde{A}\widetilde{F} + \widetilde{B}$, $a_1 \equiv \widetilde{C} - \widetilde{A}\widetilde{F} + \widetilde{D}\widetilde{F}$, and $a_2 \equiv \widetilde{E} - \widetilde{D}\widetilde{F}$. Equation (56) is a second-order difference equation in $\widetilde{N}_t^s$. In order to have a stable balanced path in terms of N_t^s, the roots of the corresponding quadratic complementary equation must be less than unity in magnitude, and by the Schur theorem, these roots are less than unity in magnitude if and only if the following conditions are satisfied: (i) $a_0^2 > a_2^2$; and (ii) $(a_0 + a_2 - a_1)(a_0 + a_1 + a_2) > 0$.[19]

Proposition 8. *Suppose that condition C3 holds, and $\widetilde{h}'$ is sufficiently small. The two Schur conditions are satisfied, and the balanced path of the source country is saddle-path stable, with (c_t, y_{t+1}) adjusting along the saddle path. Furthermore, the population of skilled workers in the source country decreases over time until the balanced path is reached.*

Proof. See the Appendix. □

In proposition 8, $\widetilde{h}'$ is sufficiently small if (a) there is a sufficiently small scale effect in education, when education depends mainly on the educator-student ratio which is fixed by the government independent of the population, and (b) human capital accumulation is insensitive to the time spent by each unskilled worker on education. In this case, labor emigration has negligible effect on human capital accumulation. On the other hand, if $\overline{w}'$ approaches infinity, then the change in the wage rate is very sensitive to labor outflow, meaning that a small emigration is needed to achieve condition (52), or that brain drain occurs slowly over time.

As the Appendix shows, both the product and the sum of the two roots of the complementary function of (56) are negative. This means that the two roots are of opposite signs, but the dominating root is positive. So emigration takes place gradually and the skilled worker population decreases monotonically over time.

The two conditions mentioned in proposition 8, however, are strong and may not be satisfied in general. To guarantee stability, the government can regulate labor outflow so that only a small number of skilled workers can emigrate in each period until the balanced path is reached in a way similar to that under permanent migration.

6. Temporary Migration

We now turn to temporary migration. Again for the time being, only this type of migration is considered: no permanent migration or brain drain

19. See Chiang (1974, p. 599).

is allowed, and the temporariness of migration may be involuntary.[20]

6.1 Features of Temporary Migration

As we did before, we suppose that starting from the beginning of period 0, with their countries initially in their balanced paths, both governments permit temporary migration from the source country to the host country. Unskilled workers in the source country have the option of going to the host country, work and receive education there. After receiving education, they become skilled workers and they are required to return back to the source country and work.

Temporary migration of the unskilled workers to the host country is attractive if $U_t^o < U_t^t$, or if

$$\ln[(1-\overline{\tau}_t)w_t] + \rho\ln[h(\overline{\tau}_t, e_t)] < \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho\ln[\overline{h}^*]. \tag{57}$$

We assume that this condition is satisfied, at least in period 0.

Temporary migration has four important features that distinguish it from the other two types of migration. First, because workers who receive education in the host country later return to the source country, they and their children remain citizens of the source country. This means that the population of skilled workers in the source country is not affected by such labor movement. Second, the workers that receive education in the host country generally have a different human capital level, and upon their return they affect directly the average human capital level in the source country. Third, migration may continue to exist even in a new balanced path of the source country. In other words, in a balanced path of the country the number of unskilled workers may not be the same as that of the skilled workers. Fourth, even though how the education burden is shared by the skilled workers is affected by the outflow of unskilled workers, it does not affect directly the decision of the unskilled workers about migration, because they have to come back when old. As a result, condition (57) does not depend on the income tax rate in the source country.

Following the notation introduced earlier, N_t^u denotes the number of unskilled workers in the source country who have not moved out in period t. Since the unskilled workers going out must return when old,

20. Because in the present section the unskilled workers are not allowed to stay in the host country after graduation, the return to the source country may be an involuntary one, and the migration is similar to a guest worker system. In the next section, both temporary migration and permanent migration are allowed, and if unskilled workers does not stay in the host country after graduation, the return to the source country is a voluntary one.

the number of skilled workers in the source country remains stationary. This implies that the effective labor force in period t is equal to

$$L_t = \{\overline{N} + N_t^u[(1 - \tau(N_t^u) - \alpha]\}x_t. \tag{58}$$

The capital-labor ratio is then

$$k_t = \frac{My_t}{\{\overline{N} + N_t^u[(1 - \tau(N_t^u) - \alpha]\}x_t}. \tag{59}$$

These two equations show how the outflow of unskilled workers changes the effective labor force and the capital-labor ratio.

The outflow of unskilled workers also affects directly human capital accumulation and the growth of the source country. In the beginning of period t, all newly born workers inherit a human capital level of x_t. Of these workers, $\overline{N} - N_t^u$ move out to the host country and get education there while N_t^u stay behind. In the beginning of next period, those who moved out return with a human capital level of $\overline{h}^* x_t$, while those who stay behind possess a human capital level of $h(\overline{\tau}, \alpha \overline{N}^u)x_t$ after having spent a time of $\overline{\tau}$ on education. The average human capital level in this period, which will be inherited by a new generation, is a weighted average of these two levels,

$$x_{t+1} = \left[\lambda_t \widetilde{h}(N_t^u) + (1 - \lambda_t)\overline{h}^*\right]x_t, \tag{60}$$

where $\lambda_t = \overline{N}_t^u/\overline{N}$ and $\overline{h}^*$ is the education output in the host country, given the time chosen by the emigrants on education. Thus those unskilled workers who are educated abroad bring back with them a new, possibly higher, skill which will be added to the existing human capital stock in the source country.

The fact that international labor migration can be a medium for human capital transfer has not received much attention in the literature. For example, if the average human capital in the host country is higher than that in the source country, then the returning migrants would bring back a higher level of skill. This in turn would raise the average skill level and thus the education effectiveness in the source country. This way of transferring human capital is analogous to technology transfer in the case of foreign direct investment.

Using conditions (59), (60), and the assumption that c_t and y_t adjust along a saddle path as described in section 2, the adjustment of the capital-labor ratio is given by

$$k_{t+1} = \frac{\beta[1 + r(k_t) - \delta][\overline{N} + N_t^u(1 - \tau(N_t^u) - \alpha)]}{[\overline{N} + N_{t+1}^u(1 - \tau(N_{t+1}^u) - \alpha)][\lambda_t \widetilde{h}(N_t^u) + (1 - \lambda_t)\widetilde{h}^*]}k_t. \tag{59'}$$

6.2 Balanced Path

We first examine the balanced path of the source country. Let a bar and a superscript "t" denote the balanced-path value of an endogenous variable. The equilibrium condition of temporary migration is

$$\ln[(1-\overline{\tau}^t)\overline{w}^t] + \rho \ln[h(\overline{\tau}^t, \alpha \overline{N}^{ut})] = \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho \ln \overline{h}^*, \tag{61}$$

where $\overline{N}^{ut}$ is the number of unskilled workers in the source country along a new balanced path. Condition (61) states that the unskilled workers remaining in the source country are indifferent to moving out and staying in the economy.

Those unskilled workers who determine not to migrate choose the education time as given by condition (18): $\overline{\tau}^t = \tau(\overline{N}^{ut})$.

Condition (60) then gives the rate of growth of human capital:

$$G_x = \overline{\lambda}^t h(\overline{\tau}^t, \alpha \overline{N}^{ut}) + (1-\overline{\lambda}^t)\widetilde{h}^*(\overline{\tau}^*, \overline{e}^*) - 1, \tag{62}$$

which depends not only on the domestic and foreign education, but also on the number of temporary migrants and the human capital they bring back.

In the presence of temporary migration, the labor force in the production sector in period t is equal to

$$L_t = [\overline{N}^{ut}(1-\overline{\tau}^t) + (1-\alpha)\overline{N}]x_t. \tag{63}$$

With the population and the number of emigrants constant under a balanced path, the labor force and the human capital stock grow at the same rate.

The growth of the physical capital stock in the source country is described by condition (11′), which can be used to give another condition for a balanced path: the equalization of the growth rates of physical and human capital, i.e.,

$$\overline{\lambda}^t h(\overline{\tau}^t, \alpha \overline{N}^{ut}) + (1-\overline{\lambda}^t)\overline{h}^* = \beta(1+\overline{r}^t - \delta), \tag{64}$$

where $\overline{r}^t$ is the balanced-path rental rate. Condition (18) gives the optimal education time. Lastly, the factor price frontier describes the relationship between the factor prices:

$$c(\overline{w}^t, \overline{r}^t) = 1. \tag{36′}$$

Conditions (61), (64), (18), and (36′) are solved for the balanced growth $\overline{\tau}^t$, $\overline{N}^{ut}$, $\overline{w}^t$ and $\overline{r}^t$.

The effects of temporary migration can be analyzed as follows. First, the factor price frontier (36′) and the growth rate equalization condition (64) can be combined together to give

$$w = \zeta(r) = \zeta\big([\overline{\lambda}^t h(\overline{\tau}, \alpha\overline{N}^{ut}) + (1 - \overline{\lambda}^t)\overline{h}^*]\beta^{-1} + \delta - 1\big).$$

Making use of this wage function, we can define the following:

$$\begin{aligned} T(\tau, N^u) &= \ln(1-\tau) + \rho \ln h(\tau, \alpha N^u) \\ &\quad + \ln \zeta\Big([\lambda h(\overline{\tau}, \alpha\overline{N}^u) + (1-\lambda)\overline{h}^*]\beta^{-1} + \delta - 1\Big), \\ T^* &= \ln[(1-\overline{\tau}^*)\overline{w}^*] + \rho \ln \overline{h}^*. \end{aligned}$$

Function $T(\tau, N^u)$ is a measure of the normalized welfare of a representative worker in the source country (less the after-tax wage income when old).[21] Condition (57) can be written in an alternative form:

$$T(\tau^a, \overline{N}) < T^*. \tag{57′}$$

The partial derivatives of function $T(\tau, N^u)$ are

$$\begin{aligned} T_\tau &\equiv \frac{\partial T}{\partial \tau} = \frac{\lambda\zeta' h_\tau}{\beta w} > 0, \\ T_N &\equiv \frac{\partial T}{\partial N^u} = \frac{\alpha\rho h_e}{h} + \frac{\zeta'[e h_e - (h^* - h)]}{\beta\overline{N}w}, \end{aligned}$$

where in evaluating T_τ, τ is assumed to be chosen optimally, and for T_N, it is noted that $\lambda\alpha\overline{N} = \overline{e}$. The sign of T_N is in general ambiguous.

Function $T(\tau, N^u)$ can be shown by different schedules in figure 5; for example, schedule TT that has the value equal to T^* is shown in the diagram. The schedule shows different combinations of (τ, N^u) in the source country that will give welfare to those workers staying behind the same as that of those that choose to migrate to the host country temporarily.

The slope of schedule TT is equal to

$$\left.\frac{\mathrm{d}\tau}{\mathrm{d}N^u}\right|_{TT} = -\frac{T_N}{T_\tau} = -\frac{e h_e - (h^* - h) + \alpha\rho\beta w\overline{N} h_e (h\zeta')^{-1}}{\lambda\overline{N} h_\tau}.$$

The sign of this slope is generally ambiguous. If condition C2 is satisfied, then schedule TT is positively sloped.

21. The after-tax wage income of a worker when old is not included in the definition of $T(\tau, N^u)$ because this income is not relevant in determining the decision of a worker in terms of temporary migration.

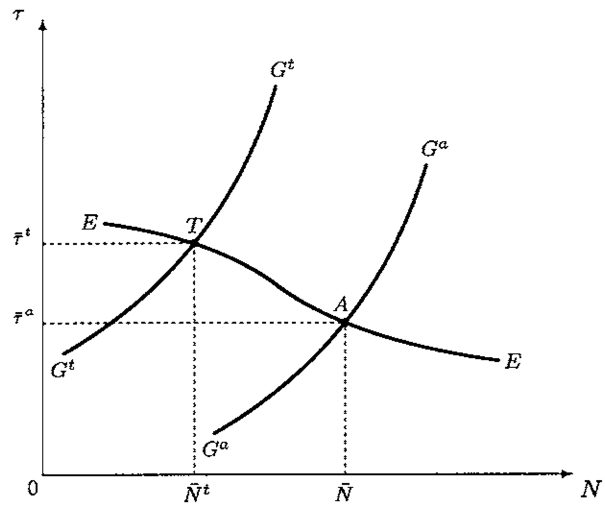


Fig. 5. Balanced path with temporary migration

Figure 5 also shows schedule EE, which describes how the optimal education time is dependent on the number of unskilled workers in the source country. As before, this schedule is given by condition (18). The intersecting point, D, between schedules EE and TT in the diagram thus gives the values of τ and N^u that represent the optimal education time for workers when taking the unskilled worker population as given, and this point satisfies the temporary-migration equilibrium and growth rate equalization conditions.

Proposition 9. *Given conditions (57′) and C1 to C2, a unique balanced path exists under temporary migration with a positive quantity of unskilled workers.*

Proof. See the Appendix. □

Figure 5 shows the equilibrium point, point D. For comparison, we also show the autarkic equilibrium point A. By making use of the diagram, we conclude that temporary migration lowers the population of the unskilled workers but induces those staying behind to spend more time on education.

How would temporary migration affect the growth rate of the source country? This is the question we now turn to. The growth rate in the presence of returning temporary migrants is

$$G^t(\tau, N^u) = \lambda h(\tau, \alpha N^u) + (1-\lambda)h^*(\tau^*, \overline{e}^*) - 1.$$

Its derivatives are (denoted by subscripts)

$$\begin{aligned} G^t_\tau &= \lambda h_\tau > 0, \\ G^t_N &= -\frac{h^* - h(1+\sigma)}{\overline{N}}, \end{aligned}$$

where $\sigma \equiv e h_e / h > 0$ is the elasticity of $h(\tau, e)$ with respect to e. The effect of an increase in τ on the growth rate is positive, $G^t_\tau = \lambda h_\tau > 0$, but that of an increase in N^u is ambiguous. However, if the education system in the host country is much more efficient than in the source country in the sense that $h^* > h(1+\sigma)$, then $G^t_N < 0$.[22]

The total effect of a reduction in the population of the unskilled workers because of temporary migration on the growth rate is equal to

$$\begin{aligned} \frac{\mathrm{d}G^t}{\mathrm{d}N^u} &= G^t_\tau \frac{\mathrm{d}\tau}{\mathrm{d}N^u}\bigg|_{EE} + G^t_N \\ &= \frac{\lambda\alpha[h h_\tau h_{\tau e} + \rho h_e h_\tau^2 - h h_e h_{\tau\tau}]}{(1+\rho)h_\tau^2 - h h_{\tau\tau}} - \frac{h^* - h}{\overline{N}}. \end{aligned} \tag{65}$$

Based on the sign of G^t_τ, a sufficient condition under which $\mathrm{d}G^t/\mathrm{d}N^u$ is negative is that $G^t_N < 0$. Thus, we conclude that if the education system in the host country is sufficiently more efficient than that in the source country so that $h^* > h(1+\sigma)$, and if condition C1 is satisfied (so that schedule EE is negatively sloped), then $G^t_N < 0$ and temporary migration increases the growth rate of the source country. Furthermore, an increase in its growth rate will lead to, by condition (64), an increase in the balanced-path rental rate but a drop in the balanced-path wage rate. These results are summarized by the following proposition:

Proposition 10. *If condition C1 is satisfied and if $h^* > h(1+\sigma)$, then temporary migration increases the growth rate and rental rate, but lowers the wage rate in the source country.*

Two remarks about this result can be made. First, if a temporary migration increases the growth rate of the source country as described by proposition 10, the wage rate per efficiency unit of labor in a new balanced path is lower than that under autarky. The reason is that as the return of the temporary migrants substantially increases the domestic

22. Alternatively, $G^t_N < 0$ if $h^* > h$ and h_e is sufficiently small.

human capital stock (when the education system in the host country is more efficient), the domestic labor force is increased. In fact, the increase in the growth rate lowers the balanced-path capital-labor ratio. This lowers that the domestic wage rate per efficiency unit of labor.

Second, if the conditions listed in proposition 10 are not satisfied, then the growth effect of temporary migration is ambiguous. The exact value of the growth effect can be determined by using (65). However, the expression is complicated, and it is not easy to derive simple and intuitive conditions for an increase in the economy's growth rate.

6.3 Transitional Dynamics

The adjustment of the economy of the source country depends on the rate of emigration of unskilled workers. We assume that, in each period, unskilled workers emigrate until the incentive to flow out disappears for that period. This means that condition (57) is satisfied with an equality in all periods. This condition together with (59′) describe a system of first-order difference equations in N_t^u and k_t.

We consider the linearized equations around the balanced path. Again we let $\widetilde{k}_t \equiv k_t - \overline{k}^t$ and $\widetilde{N}_t^u \equiv N_t^u - \overline{N}$ be the deviations of the variables from their balanced-path values. Equation (59′) gives

$$\widehat{A}\,\widetilde{k}_{t+1} + \widehat{B}\widetilde{N}_{t+1}^u = \widehat{C}\widetilde{k}_t + \widehat{D}\widetilde{N}_t^u, \tag{66.1}$$

where these coefficients, when evaluated close to the balanced path, are equal to

$$\begin{aligned}
\widehat{A} &= [\overline{N} + \overline{N}^{ut}(1 - \overline{r}^t - \alpha)]\Phi,\\
\widehat{B} &= \overline{k}^t\Phi[1 - \overline{r}^t - \overline{N}^{ut}\overline{r}^{t\prime} - \alpha],\\
\widehat{C} &= \beta(1 + \overline{r}^t + \overline{k}^t\overline{r}^{t\prime} - \delta)[\overline{N} + \overline{N}^{ut}(1 - \overline{r}^t - \alpha)],\\
\widehat{D} &= \overline{k}^t\Phi(1 - \overline{r}^t - \overline{N}^{ut}\overline{r}^{t\prime} - \alpha) -\\
&\quad\ \overline{k}[\overline{N} + \overline{N}^{ut}(1 - \overline{r}^t - \alpha)][\overline{\lambda}^t\widetilde{h}^{t\prime} + (\widetilde{h}^t - h^*)]/\overline{N}],\\
\Phi &= [\overline{\lambda}^t\widetilde{h}^t + (1 - \overline{\lambda}^t)\widetilde{h}^*] = \beta(1 + \overline{r}^t - \delta).
\end{aligned}$$

Equation (61) can be linearized in a similar way, giving

$$\widetilde{N}_t^u = \widehat{E}\widetilde{k}_t, \tag{66.2}$$

where $\widehat{E} = -\overline{h}^t\overline{w}^{t\prime}(\alpha\rho\overline{w}^t h_e^{t\prime})^{-1}$. Substitute (66.2), for both $\widetilde{N}_t^u$ and $\widetilde{N}_{t+1}^u$, into (66.1) to give

$$\widetilde{k}_{t+1} = \frac{\widehat{C} + \widehat{D}\widehat{E}}{\widehat{A} + \widehat{B}\widehat{E}}\widetilde{k}_t. \tag{67}$$

Lemma 2. *The adjustment of k_t as given by the difference equation in (67) is stable if*

1. *$\overline{w}^{t\prime}$ is sufficiently small and if the rental rate is inelastic; or*
2. *$\overline{w}^{t\prime}$ is sufficiently large and $[\overline{\lambda}^t\widetilde{h}^{t\prime} + (\widetilde{h}^t - h^*)]/\overline{N}] < 0$.*

Proof. It follows directly from condition (61). □

Proposition 11. *The system is saddle-path stable if the conditions in lemma 2 are satisfied.*

Proof. If the difference equation in $\widetilde{k}_t$ is stable under the conditions stated in lemma 2, then by (66.1)–(66.2), so is the adjustment of $\widetilde{N}_t^u$. The adjustment of c_t and y_t are saddle-path stable, as described earlier.

□

7. Endogenous International Migration

So far, we have analyzed each of the three types of international migration separately under the assumption that only one type of migration is allowed. We now relax this assumption. To focus our analysis more on the source country, we do keep the simplifying assumption that only labor is allowed to move only from the source country to the host country.

Endogenizing these types of international migration is to allow workers in the source country to choose where to stay when young and old, where to receive education, and where to work. The decision of the worker therefore depends on the following two factors: the effectiveness of education and the after-tax wage rates. Let us measure the effectiveness of education by the following variable:

$$Z = (1-\tau)[(1-\phi)h(\tau, \alpha N^u)]^\rho.$$

Variable Z depends on the time not spent on education, $1-\tau$, and the discounted "after-tax" returns on education, $[(1-\phi)h(\tau,\alpha N^u)]^\rho$. Therefore education is said to be effective if it (a) requires less time; (b) implies a smaller income tax; and/or (c) produces more human capital. The corresponding variable for the host country can be similarly defined:

$$Z^* = (1-\tau^*)[(1-\phi^*)\overline{h}^*]^\rho.$$

The values of Z and Z^* are measures of education effectiveness in the source and host countries, respectively.

Focusing the analysis in the present section on balanced path, or the cases in which the source economy is close to a balanced path, we examine the effects of labor emigration on the wage rate and effectiveness of education in the source country. Dynamic analysis of switching between different types of labor migration, however, is beyond the scope of this chapter.

Let us measure respectively the welfare level of a worker in the source country in the presence of permanent migration, brain drain, and temporary migration by the respective functions:

$$\begin{aligned}\widetilde{P}(Z,w) &= \ln Z + (1+\rho)\ln w,\\ \widetilde{B}(w) &= \ln w,\\ \widetilde{T}(Z,w) &= \ln Z + \ln w,\end{aligned}$$

where all endogenous variables are measured near the balanced path. It is easy to see that each of these functions is strictly increasing in their arguments. These three functions are appropriate indicators of the welfare levels of workers remaining in the source country when different types of international migration are allowed. Denote the welfare levels of the workers who move in the form of permanent migration, brain drain, and temporary migration by $\widetilde{P}^*$, $\widetilde{B}^*$, and $\widetilde{T}^*$, respectively, which are defined as

$$\begin{aligned}\widetilde{P}^* &= \ln Z^* + (1+\rho)\ln w^*,\\ \widetilde{B}^* &= \ln w^*,\\ \widetilde{T}^* &= \ln Z^* + \ln w^*.\end{aligned}$$

Using a superscript "a" to denote the autarkic balanced-path value of a variable, we can say that in the absence of any government restrictions, permanent migration, brain drain, or temporary migration is attractive if, respectively,

$$\widetilde{P}(Z^a, w^a) < \widetilde{P}^*, \tag{68.1}$$

$$\widetilde{B}(Z^a, w^a) < \widetilde{B}^*, \tag{68.2}$$

$$\widetilde{T}(Z^a, w^a) < \widetilde{T}^*. \tag{68.3}$$

Two remarks can be made. First, there are cases in which more than one types of migration are simultaneously attractive to the workers in the source country. In these cases, one type of migration may dominate the other types, meaning that if workers can choose the timing and duration of migration, the dominant type is preferred. Second, since emigration affects variables in the source country, it is possible that initially one type of migration dominates the other types, but as the source country grows, the dominant type changes.

The equilibria of the present system under different types of international migration can be represented by values of Z and w that satisfy the expressions in (68.1)–(68.3) with the inequalities replaced by equalities. Different possible equilibrium combinations of values of Z and w in the source country are illustrated in figure 6. Thus schedules $\widetilde{P}\widetilde{P}'$, $\widetilde{B}\widetilde{B}'$, and $\widetilde{T}\widetilde{T}'$ represent, respectively, the equations $\widetilde{P}(Z,w) = \widetilde{P}^*$, $\widetilde{B}(Z,w) = \widetilde{B}^*$, and $\widetilde{T}(Z,w) = \widetilde{T}^*$. The functions are increasing toward the right and upward. We assume that any two schedules intersect only once.

Let us examine some of the properties of these schedules. First, we note that point H, which represents the balanced path of the host country, $(Z^*, \overline{w}^*)$, satisfies all these three functions, meaning that the three schedules must pass through point H. Second, schedule $\widetilde{B}\widetilde{B}'$ is a vertical line at $w = \overline{w}^*$. Third, schedules $\widetilde{P}\widetilde{P}'$ and $\widetilde{T}\widetilde{T}'$ are negatively sloped. Fourth, schedule $\widetilde{P}\widetilde{P}'$ is steeper than schedule $\widetilde{T}\widetilde{T}'$ at point $(Z^*, \overline{w}^*)$.[23] Fifth, a point above and to the right of a schedule represents the values of Z and w with which the corresponding type of international migration will not occur because a worker can get a higher welfare and/or a more efficient education by staying in the source country than by migrating. Similarly, a point below and to the left of a schedule means that the corresponding type of migration is preferred to no migration.

The three schedules and the horizontal line through point H divide the space in figure 6 into seven regions labelled I to VII. Whether emigration and what type of emigration will occur depends on where the autarkic point of the source country is. We now analyze each of these regions.

Region I – An autarkic point of the source country in this region represents values of Z and w that yield a welfare level to a worker in the source country higher than what she can get by moving to the host

23. The slopes of schedules $\widetilde{P}\widetilde{P}'$ and $\widetilde{T}\widetilde{T}'$ are respectively equal to

$$\left.\frac{\mathrm{d}Z}{\mathrm{d}w}\right|_{\widetilde{P}\widetilde{P}'} = -\frac{Z(1+\rho)}{w},$$

$$\left.\frac{\mathrm{d}Z}{\mathrm{d}w}\right|_{\widetilde{T}\widetilde{T}'} = -\frac{Z}{w}.$$

country. Thus, this region represents no international migration.

Region II – In this region, temporary migration but not other types of migration exists.

Region III – In this region, both temporary migration and permanent migration are attractive to the workers in the source country. This means that unskilled workers have incentives to get education in the host country. However, because the domestic wage rate is higher than that in the host country, workers from the source country will prefer to return after graduation. In other words, temporary migration dominates permanent migration.

Region IV – In this region, all three types of migration are attractive to the workers in the source country. However, because $Z < Z^*$, i.e., education is less efficient in the source country than in the host country. Workers prefer to receive education in the host country. Because $\overline{w}^* > w$, workers will choose to stay in the host country after graduation. In other words, in this region permanent migration dominates the other two types of migration.

Region V – In this region, again all three types of migration are attractive. However, because $Z > Z^*$, workers in the source country will choose to have education at home. Thus brain drain is preferred to the other two types of migration.

Region VI – In this region, only brain drain and permanent migration are preferable, but because education is more efficient in the source country than in the host country, brain drain dominates permanent migration.

Region VII – In this region, only brain drain will be chosen over no migration.

The diagram brings out several features of endogenous international labor migration which has rarely been investigated in the literature. First, it is possible that when labor emigration is allowed, more than one type of labor migration is attractive to the domestic workers. Second, usually different types of labor migration can be ranked by the workers. If workers are free to choose the type of labor migration, they may choose one over the other. Third, as the economy adjusts, the dominant type of labor migration may change.

A full analysis of possible switches from one type of labor migration to another requires a rigorous dynamic analysis, which is beyond the scope of this chapter. Instead, we focus on balanced paths. The following results can easily be obtained from figure 6 and the analysis in the previous sections.

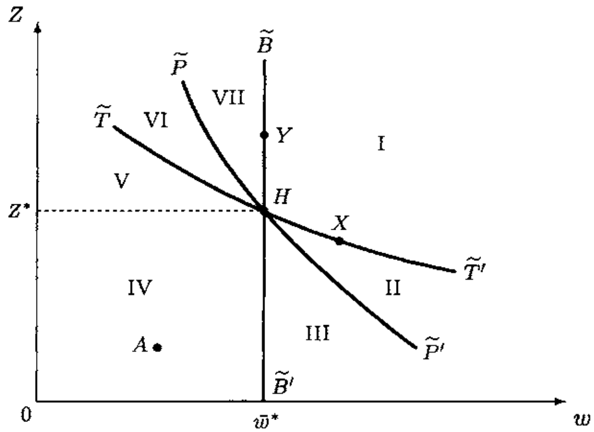


Fig. 6. Choosing between permanent migration, brain drain, and temporary migration

Proposition 12. *Suppose that with both countries initially in their balanced paths, labor is allowed to migrate from a small, source country to the host country.*

1. *There are cases in which permanent migration exists, and in some other cases, brain drain or temporary migration may exist, depending on the autarkic wage rates and education efficiency in the source and host countries.*

2. *One type of migration may switch to another type as more and more labor moves out.*

3. *No matter what type of migration exists initially, the new balanced path, if exists, is characterized by either brain drain or temporary migration.*

4. *If initially permanent migration exists, as in the case in which the autarkic point is given by point A in figure 6, it switches to temporary migration sooner or later if the wage rate goes up sufficiently faster than the education effectiveness in the source country. In this case, a new balanced path may exist at a point like X in figure 6. If instead the education effectiveness rises sufficiently faster than the wage rate in the source country, permanent migration switches to brain drain later, and a new balanced path may be represented by a point such as Y.*

5. *If initially temporary migration exists, and if in the source country the wage rate rises monotonically as labor moves out, then the new balanced path will be characterized by temporary migration.*

6. *If initially brain drain exists, and if the education effectiveness rises monotonically, then the new balanced path will be characterized by brain drain.*

7. *If brain drain exists initially and the balanced path in the long run is characterized by temporary migration, or if temporary migration exists initially and the balanced path in the long run is characterized by brain drain, then there exist some periods in which permanent migration exists.*

8. Concluding Remarks

We have constructed a simple model in which an economy grows because of accumulation of both human capital and physical capital. A balanced path is derived in which both types of capital grow at the same rate. We then extended the model to two-countries and analyzed international labor migration. For each of the three types of labor migration, we analyzed how each may affect the source country's growth and income distribution between workers and capitalists. We also explained how workers in the source country may choose between different types of migration, and showed the possibility that the dominant type of migration may change as the source country grows over time.

The perpetual growth of the source country in the present model is due to education and unbounded accumulation of human capital. Therefore the effects of labor migration on growth work through education. We argued that in the present model, both permanent migration and brain drain may have adverse effects on growth. These adverse growth effects can be explained in terms of the externalities that exist in education. Temporary migration, which does not lead to a decrease in the population of the source country, tends to have positive growth effects. These positive effects, which are not well recognized in the literature, come from the human capital brought back by the migrants, who have received education in the host country. Such transfer of human capital is analogous to the transfer of technology accompanied by the investment of foreign firms.

Appendix

Proof of Lemma 1. Rearrange condition (17) to give (the time subscript being dropped for simplicity),

$$1 - \tau = \frac{h(\tau, e)}{\rho h_\tau}. \tag{69}$$

By assumption 2, when τ approaches zero, $h(\tau, e)$ approaches unity and $h_\tau > 1/\rho$, implying that the left-hand side (LHS) of (69) is greater than the right-hand side (RHS). When τ approaches unity, the LHS of (69) is less than the RHS. The LHS of the condition is a strictly decreasing function of τ, while the RHS, because $h_\tau > 0$ and $h_{\tau\tau} < 0$ by assumption 2, is a strictly increasing function of τ. By continuity of the functions, there exists one and only one value of $\tau \in (0, 1)$ that satisfies condition (69) when given ρ and e. □

Proof of Proposition 1. By lemma 1 and condition (18), the optimal education time in an autarkic balanced path is given by $\overline{\tau}^a = \tau(\overline{N}) \in (0, 1)$. Once $\overline{\tau}^a$ is known, the growth rates of human capital stock and the labor force are determined. Let $\overline{h}^a \equiv h(\overline{\tau}^a, \alpha\overline{N})$. By lemma 1, $\overline{h}^a > 1$. So $\overline{h}^a - 1$ is the (positive) growth rate of human capital. Define the rental rate that satisfied (22) as

$$\overline{r}^a = \frac{\overline{h}^a}{\beta} + \delta - 1. \tag{70}$$

Because $\beta < 1$, the right-hand side of (70) is positive. Therefore by assumption 1, the rental rate defined in (70) is finite, positive, and unique. The wage rate is obtained from the factor price frontier, and is unique and positive. The factor prices give a unique capital-labor ratio, which is denoted by $\overline{k}^a$. □

Proof of Proposition 3. Condition (30) implies that labor movement exists when allowed. In terms of figure 2, and because $P_\tau < 0$, schedule PP is below schedule EE when $N = \overline{N}$. Suppose that significant labor emigration exists so that the labor force in the source country is small. Assumption 1 implies that the physical capital stock in the source country is finite, and that the corresponding wage rate is infinite. In other words, the rise in the local wage rate will prevent all the unskilled workers in the source country from moving out. So there exists $N \in (0, \overline{N})$ that satisfies (33). This is represented by a point of intersection between schedules PP and EE in figure 2. The education time can be obtained from (34). Other variables can be obtained accordingly. If conditions

C1–C2 are also satisfied, then a comparison of conditions (19) and (38) shows that schedule PP is steeper than schedule EE at their point of intersection in figure 2. Thus the balanced path is unique. □

Proof of Proposition 6. Condition (44) implies that brain drain exists when allowed. Assumption 1 implies that the wage rate in the source country will rise sufficiently high to prevent all workers from moving out. So there exists $N \in (0, \overline{N})$ that satisfies (48).

We now turn to uniqueness. It has been shown that when given condition C1, both schedules hh and EE are negatively sloped, but schedule hh is steeper at the point of intersection. This means schedules $h^b h^b$ and EE can intersect at most once. Thus, a balanced path with brain drain exists and is unique. □

Proof of Proposition 7. The proposition can be proved by using figure 4. Point A shows the autarkic point with the initial population. The shrinkage of population due to brain drain means that the new equilibrium point, B, is to the left of and higher than point A, which in turn implies that the growth rate of the economy under brain drain is lower than that under autarky. The rest of the proof is similar to that of Proposition 4 (given in the text) and thus is omitted. □

Proof of Proposition 8. We first consider Schur condition (i).

$$\begin{aligned} a_0^2 - a_2^2 &= (\widetilde{A}\widetilde{F} + \widetilde{B})^2 - (\widetilde{E} - \widetilde{D}\widetilde{F})^2 \\ &= \left(\overline{k}^b\overline{h}^b - \overline{h}^b\alpha\overline{w}^b(2-\overline{\tau}^b-\alpha)[(1-\alpha)\overline{w}^{b\prime}]^{-1}\right)^2 \\ &\quad -\Big(\overline{k}^b\overline{h}^{b\prime}\overline{N}^b(2-\overline{\tau}^b-\alpha) - \overline{k}^b\overline{h}^b(1-\overline{\tau}^b-\overline{N}^b\overline{\tau}^{b\prime}-\alpha) \\ &\quad -\alpha\overline{w}^b(2-\overline{\tau}^b-\alpha)[\overline{h}^b + \beta\overline{k}^b\overline{\tau}^{b\prime}][(1-\alpha)\overline{w}^{b\prime}]^{-1}\Big)^2, \end{aligned}$$

which is positive if the conditions in the proposition are satisfied. We next consider condition (ii). By using the definitions of a_0, a_1, and a_2, we have

$$\begin{aligned} &(a_0 + a_2 - a_1)(a_0 + a_1 + a_2) \\ &= (2\widetilde{A}\widetilde{F} + \widetilde{B} - \widetilde{C} - 2\widetilde{D}\widetilde{F} + \widetilde{E})(\widetilde{B} + \widetilde{C} + \widetilde{E}) \\ &= \overline{k}^b\overline{h}^{b\prime}\overline{N}^b(2-\overline{\tau}^b-\alpha)\Big\{2\overline{k}^b\overline{h}^b(\tau + \overline{N}^b\overline{\tau}^{b\prime} + \alpha) + \overline{k}^b\overline{h}^{b\prime}\overline{N}^b(2-\overline{\tau}-\alpha) \\ &\quad -4\alpha\overline{h}\overline{w}^b(2-\overline{\tau}-\alpha)[(1-\alpha)\overline{w}^{b\prime}]^{-1} \\ &\quad -2\alpha\beta\overline{k}^b\overline{w}^b\overline{\tau}^{b\prime}(2-\overline{\tau}-\alpha)[(1-\overline{\tau})\overline{w}^{b\prime}]^{-1}\Big\}, \end{aligned}$$

which is positive if $\overline{w}^{b\prime} \to \infty$ (condition C3). Thus, the Schur condition for stability with respect to N_t^s is satisfied. To find out whether N_t^s

oscillates, note that, by direct substitution, $a_2/a_0 < 0$ and $a_1/a_0 < 0$, meaning that the two roots in (56) are of opposite signs but the dominating root is positive. This proves the proposition. □

Proof of Proposition 9. Condition (57′) implies that temporary migration exists when allowed. In the presence of temporary migration, by assumption 1, the wage rate for the unskilled workers will rise sufficiently high to prevent all workers from moving out. Thus there exists $N \in (0, \overline{N})$ that satisfies (61). In terms of figure 5, schedule TT is below schedule EE when $N^u = \overline{N}$, and is above schedule EE when N^u approaches zero. Continuity of the schedules means that they cut each other at least once. Given condition C1, schedule EE is negatively sloped, and given condition C2, schedule TT is positively sloped. Thus, the balanced path is unique. □

Acknowledgement. Thanks are due to Murray Kemp, Philip Levy, Koji Shimomura, and Chong Kee Yip for helpful comments. This chapter has also been presented at the University of Hong Kong, Hong Kong Baptist University, Academia Sinica, Chung Hua Institution for Economic Research, Kobe University, Kansai University, Hosei University and Hitotsubashi University. The comments from the participants at these seminars are also gratefully acknowledged.

References

Azariadis, Costas, and Allan Drazen (1990). "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105: 501–526.

Blomqvist, Ake G. (1986). "International Migration of Educated Manpower and Social Rates of Returns to Education in LDCs," *International Economic Review* 27: 165–174.

Bond, Eric W., Ping Wang, and Chong K. Yip (1996). "A General Two-Sector Model of Endogenous Growth with Human and Physical Capital: Balanced Growth and Transitional Dynamics," *Journal of Economic Theory* 68: 149–173.

Caballé, Jordi, and Manuel S. Santos (1993). "On Endogenous Growth with Physical and Human Capital," *Journal of Political Economy* 101: 1042–1067.

Chiang, C. Alpha (1974). *Fundamental Methods of Mathematical Economics*, Second Edition. New York: McGraw-Hill.

Findlay, Ronald, and Carlos A. Rodriguez (1981). "A Model of Economic Growth with Investment in Human Capital." In Ali Kahn and Ismail Sirageldin (eds.) *Equity, Human Capital, and Development*, Vol. 2, pages 57–72. Greenwich, Conn.: JAI Press Inc.

Galor, Oded (1986). "Time Preference and International Labor Migration," *Journal of Economic Theory* 38: 1–20.

— and Oded Stark (1990). "Migrants' Savings, The Probability of Return Migration and Migrants' Performance," *International Economic Review* 31: 463–467.

— and Oded Stark (1994). "Migration, Human Capital Formation, and Long-Run Output." In Horst Siebert (ed.), *Migration: A Challenge for Europe.* Tübingen: J.C.B. Mohr, 59–68.

Grubel, Herbert B., and Anthony D. Scott (1966). "The International Flow of Human Capital" *American Economic Review* 56: 268–274.

Hu, Sheng Cheng (1976). "Education and Economic Growth," *Review of Economic Studies* 43: 509–518.

Johnson, Harry G. (1967). "Some Economic Aspects of Brain Drain," *Pakistan Development Review* 7: 379–411.

Long, Ngo Van, and Kar-yiu Wong (1997). "Endogenous Growth and International Trade: A Survey." (this volume).

Lucas, Robert E., Jr. (1988). "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22: 3–42.

Manning, R. (1975). "Optimal Aggregative Development of a Skilled Workforce," *Quarterly Journal of Economics* 89: 504–511.

— (1976). "Issues in Optimal Educational Policy in the Context of Balanced Growth," *Journal of Economic Theory* 13: 380–395.

Miyagiwa, Kaz (1991). "Scale Economies in Education and the Brain Drain Problem," *International Economic Review* 32: 743–759.

Ohyama, Michihiro (1991). "Human Capital and Endogenous Economic Growth," *Keio Economic Studies* 28: 1–14.

Razin, Assaf (1972a). "Optimum Investment in Human Capital," *Review of Economic Studies* 39: 455–460.

— (1972b). "Investment in Human Capital and Economic Growth," *Metroeconomica* 24: 101–116.

Rodriguez, Carlos Alfredo (1976). "Brain Drain and Economic Growth." In Jadgish N. Bhagwati (ed.) *The Brain Drain and Taxation.* Amsterdam: North Holland, 171–195.

Shea, K.-L., and A.E. Woodfield (1996). "Optimal Immigration, Education and Growth in the Long Run," *Journal of International Economics* 40: 495–506.

Uzawa, Hirofumi (1965). "Optimal Technical Change in an Aggregative Model of Economic Growth," *International Economic Review* 6: 18–31.

Wong, Kar-yiu (1995). *International Trade in Goods and Factor Mobility.* Cambridge, Mass.: MIT Press.

CHAPTER 11

The Human Capital Dimension to Foreign Direct Investment: Training, Adverse Selection, and Firm Location

Theo S. Eicher and Pantelis Kalaitzidakis

1. Introduction

The development literature emphasizes technology transfers as a central aspect of take-off and convergence of growth rates. Arguably the most important channel of technology transfer is foreign direct investment (FDI). While theoretical models of FDI and firm location focus largely on technology and physical capital, recent empirical evidence underscores that the success of technology transfer via FDI depends crucially on the size of the developing country's human capital stock, see Borensztein, DeGregorio, and Lee (1995). In addition, Hummels and Stern (1994) documented that the lion share of FDI occurs among nations with similar technology and human capital levels.

This chapter examines the role of multinational corporations (MNCs) in facilitating international technological diffusion, and the role of human capital in determining firm location. In focusing on human capital, we introduce a new dimension to trade and FDI: informational asymmetries. We combine an efficiency wage approach to labor markets with a model of trade and FDI by embedding an adverse selection model into a two sector general equilibrium framework that extends to the open economy. This allows us to analyze the human capital dimension to trade and FDI: the choice of firm location when investment in firm specific training is affected by adverse selection problems.

Labor market information asymmetries have not yet been analyzed in the international trade and location literature. This is surprising, since a key aspect of firm location is *ownership advantage*, e.g., a firm specific production process, blueprint, or technology, see Dunning (1977, 1981). When workers are heterogeneous in their abilities to learn about this ownership advantage and MNCs are unable to judge individual skills perfectly, MNCs cannot make full use of their ownership advantage. Any location model is thus incomplete without the specification of a distinct information set that determines how employers form expectations about worker productivity.

Our model features only one factor, heterogeneous labor. The level of human capital is important to firm location because domestic workers must learn about firm specific technologies, and firms must provide these skills through training. Investment in our model is thus investment in people and skills. We assume with Kalaitzidakis (1996) that the training efficiency depends on worker *quality*, which contains both observable and unobservable components. Since MNCs cannot ascertain workers' training efficiency with certainty ex ante, firms' hiring and location decisions are subject to adverse selection problems. The model consists of two sectors, agriculture and manufacturing. The agricultural sector establishes a reservation wage through self-employment, while the manufacturing sector pays an efficiency wage to counter the adverse selection problem. Agents differ in their productive abilities within countries, and in both their abilities and their observable human capital across countries.

The model yields a number of insights that are new to trade and foreign direct investment literature, but reminiscent of the implications of efficiency wage models. Adverse selection generates a pattern of static comparative advantage that is akin to both the Heckscher-Ohlin and the Ricardian model. As in the Ricardian model, countries with more sophisticated technologies feature higher wages and export the technology intensive good. Reminiscent of the Heckscher-Ohlin model, human capital abundant countries adopt more sophisticated technologies and possess a comparative advantage in the learning intensive good. Unlike in the Heckscher-Ohlin model, a human capital, abundant country that shares identical technologies with all other countries features a higher manufacturing wage because the expected quality of its applicant pool is higher. We also find that a country that shares the same levels of human capital and technology with all other countries, but possesses a larger labor endowment, has a comparative advantage in agriculture because the size of the population increases the adverse selection problem firms face.

In contrast to the Heckscher-Ohlin model, trade is associated with falling (rising) wages in the developed (developing) country. The wage convergence signals a novel effect introduced by the addition of adverse selection to trade: *informational efficiency gains from trade.* These gains arise because both countries relocate production to the sector where workers have a comparative advantage in terms of production *and* informational efficiency. The country with the comparative advantage in training expands production of the training intensive good, pays lower efficiency wages, but also enjoys a higher quality applicant pool. That is, trade diminishes the effect of the informational asymmetries in the developed country's manufacturing sector. Informational gains also accelerate the adjustment and increase the output of the agricultural good in the LDC.

The analysis of location reveals that the incentive to open subsidiaries in foreign countries diminishes with the technological and human capital differences between countries, as discussed by Borensztein, DeGregorio, and Lee (1995). The larger the difference in human capital levels across countries, the greater the effect of the informational asymmetry on the training efficiency for the MNC. Also, the more pronounced the technology gap between countries, the higher the training costs for MNCs. If a firm opens a production plant in a foreign country, we find that informational asymmetries naturally give rise to the type of multiple wage equilibria observed by Feenstra and Hanson (1995) and Aitken, Harrison, and Lipsey (1995). That is, MNCs pay higher wages than domestic firms in equilibrium because MNCs use higher levels of technology than domestic firms and therefore seek to attract higher quality workers to minimize training costs. Interestingly enough, the MNC pays workers in the foreign country less than in the home country, since information and training cost are higher in the developing country, where the cohort is of observationally lower quality. This is what we term the human capital dimension to foreign direct investment. Finally we examine the dynamics of the model and find that our equilibrium is locally saddle point stable. However, sustained growth comes to a halt, despite endogenous technological change, because the training cost per worker eventually outpaces productivity increases.

Labor market information asymmetries have not been extensively analyzed in the international trade literature, which commonly assumed perfect information in labor markets. Dixit (1989) modeled adverse selection in an open economy, but informational asymmetries exist only between entrepreneurs and policy makers. Moral hazard was introduced into trade models by Copeland (1989) and Bulow and Summer (1986), who examined commercial policy; and by Brecher (1992) and Brecher and Choudhri (1994), who examined the welfare effects of commercial policy given unemployment. Informational asymmetries have been introduced into various other areas of the open economy. Markusen (1995) provided an exhaustive survey of the analysis of moral hazard problems associated with licensing agreements that multinationals face. Ethier (1986) examined incomplete contracts. Dixit (1989) examined commercial policy in a model where the probability of success in a risky production sector is private information, and in a model where the migration decision is a function of the small open economy's stochastic terms of trade, see Dixit (1994).

The adverse selection problems that firms face when hiring heterogeneous workers under imperfect information, was first explicitly modeled by Weiss (1980). Weiss showed that such market imperfections require employers to identify the mass of workers that accrue the minimum cost per efficiency unit of labor. Nalebuff and Stiglitz (1982) subsequently presented a model where increases in wages increase the expected ability of the applicant pool because lower quality applicants see their probability of being hired diminish. We will utilize the insights of these adverse selection models to model how firms select applicants and efficiency wages below. The empirical evidence on the informational asymmetries in labor markets is scarce. There is little evidence for moral hazard, but compelling evidence for adverse selection problems, as documented by Foster and Rosenzweig (1993). Their study concluded that higher productivity workers participate less in time wage markets when the return to piece rate (self-employment) work increases. Foster and Rosenzweig also showed that there is considerable ignorance among employers about the individual difference in workers' abilities in developing countries.

2. A General Equilibrium Adverse Selection Model

2.1 Agriculture

We make two assumptions that pay tribute to the traditional notion of agricultural sectors. First, the sector is one of self-employment. Workers opt to work in agriculture only when they do not find employment in manufacturing, or when the value of their marginal product in agriculture exceeds the wage they would receive in manufacturing. Second, we choose a linear production function not only to simplify matters, but also to reflect the traditional notion that the marginal product equals the average product in the agricultural sectors. The total output of the agricultural good, X, is given by

$$X = \sum_{i=0}^{L^X} G[\theta(i), H], \quad i \in [0, L], \quad G_H > 0, \quad G_\theta > 0, \tag{1}$$

where subscripts indicate partial derivatives. L are the total units of labor in the economy, which divide themselves into agricultural and manufacturing employment, L^X and L^Y, respectively. Productivity, $G[\theta(i), H]$, depends on the *quality* of the individual worker, which consists of two components: *observable human capital*, H, and *unobservable ability*, θ. H represents the average level of human capital (e.g., years of schooling) that is *observable* in a country. Once we introduce trade we will assume that H varies exogenously across countries and that H_j, $j = 0, 1, 2, \ldots$, can be ordered across countries so that H_0 represents the country with the lowest level of average general human capital. We suppress country indicator superscript unless needed.

Since each self-employed worker knows her ability, θ, there exist no informational asymmetries in the agricultural sector. Hence, the return to labor in agriculture is known with certainty to each individual. We label the value of the marginal product in agriculture the *reservation wage*, ω^i, of worker i with human capital H at price π, or

$$\omega^i = \pi\, G[\theta(i), H], \tag{2}$$

where π represents the relative price of the agricultural good. Given that higher quality workers have higher productivity, the following derivatives are straightforward: $\partial\omega^i/\partial H > 0$, $\partial\omega^i/\partial\pi > 0$, and $\partial\omega^i/\partial\theta > 0$.

2.2 Manufacturing

For simplicity, we assume that each country possesses just one firm and one representative technology, A. The firm's production function for manufacturing output, Y, is given by

$$Y = F\left[T[A]\, L^Y\right], \;\; F'[\cdot] > 0, \;\; F''[\cdot] < 0, \;\; T'[\cdot] > 0, \;\; T''[\cdot] > 0. \tag{3}$$

The productivity of labor, $T[\cdot]$, depends on firm specific technology A. To be able to work with A, labor must acquire firm specific skills. Since skills are firm specific, the employer must pay the training cost, see Becker (1975). We assume with Kalaitzidakis (1996) that firms incur training costs that are a function of worker quality. Specifically, we express the cost of training worker i as

$$\bar{C}^i[\theta(i), H, A] = T[A]\, C\left[q[\theta(i), H]\right], \;\; C'[\cdot] < 0, \;\; C''[\cdot] > 0, \tag{4}$$

where, for simplicity, training costs depend linearly on the amount of training required for each specific technology, $T[A]$. The training efficiency, $q[\cdot]$, of worker i (i.e., how easily a worker can learn new skills) is determined by the worker's quality, $\theta(i)$ and H. Using (2) we can write training efficiency as $q[\omega^i, \pi, H]$, with $q_\omega > 0$ and $q_\pi < 0$.

The manufacturing sector is, however, marred by informational asymmetries. Firms hire workers whose training efficiency depends on their quality, which is not known with certainty to the employer. Manufacturers may observe general human capital, H, across countries (i.e., from UNESCO educational attainment statistics), but not the exact ability of each applicant, θ (i.e., how fast an individual worker learns about the firm specific technology). These informational asymmetries create an adverse selection problem and firms realize that the quality of the applicant pool deteriorates at any given level of observed human capital as the manufacturing wage declines.

While firms cannot observe the quality of worker i ex ante, they are capable of generating beliefs about the applicant pool's expected quality on the basis of workers' reservation wages. Since the reservation wage is increasing in θ, firms may use the manufacturing wage, w, to influence the expected quality of their applicant pool. Hence, following Weiss (1990), firms base their hiring decisions on the *expected quality*, Q,

$$Q[w, H, \pi] = \frac{\int_0^w q[w^i, \pi, H]\, D[w^i] dw^i}{\int_0^w D[w^i] dw^i}, \tag{5}$$

where $D[\omega^i]$ gives the mass of workers with reservation wage ω^i. Equation (5) states that at a given relative price π, firms can expect an applicant with observable human capital H to possess quality Q, at a given wage offer, w. From $q_\omega > 0$ and $q_\pi < 0$, it follows that $Q_w > 0$ and $Q_\pi < 0$. In addition, we assume that, at a given wage, the expected quality increases in the level of observable human capital, or $Q_H > 0$. We also assume that $Q_{\pi\pi} < 0$, $Q_{ww} < 0$, and $Q_{w\pi} = 0$. We can now rewrite the firm's training cost per worker as

$$\bar{C} = T[A]\, C\left[Q[w, H, \pi]\right]. \tag{4'}$$

The manufacturing firm's problem then consists of maximizing profits, ρ, over employment and wages

$$\max_{w, L^Y} \rho = F\left[T[A]\, L^Y\right] - \left(w + T[A]\, C\left[Q[w, H, \pi]\right]\right) L^Y. \tag{6}$$

The first order conditions can then be derived as

$$T[A]\, C'\left[Q[w, H, \pi]\right] Q_w[w, H, \pi] = -1, \tag{7}$$

$$F'\left[T[A]\, L^Y\right] = \frac{w}{T[A]} + C\left[Q[w, H, \pi]\right]. \tag{8}$$

Equation (8) solves for the optimal number of workers employed at any given wage. It simply states that the marginal product must equal the marginal cost to firms, where the cost depends on both the wage and the expected training costs. It will be convenient to define the productivity adjusted cost as the *Average Efficiency Cost* (AEC) of the firm with technology A as

$$AEC \equiv \frac{w + T[A]\, C\, [Q[w, H, \pi]]}{T[A]} \tag{9}$$

with $\partial AEC/\partial w|_{w^*} = 0$, and $\partial^2 AEC/\partial w^2 = C''[Q]Q_w^2 + C'[Q]Q_{ww} > 0$. That is, the efficiency wage minimizes the AEC of firms; and the AEC is convex in the wage. The minimization of the AEC or, equivalently, the profit maximizing wage condition, (7), determines the efficiency wage, $w^* = w^*[A, H, \pi]$, as a function of the exogenous parameters: technology, observable human capital, and the (partial equilibrium) relative price. Equation (7) replicates the typical efficiency wage condition that, at equilibrium, a unitary decrease in the wage cost must generate an equal increase in the training cost. Also, (7)–(8) reproduce the usual efficiency wage result that, as long as the labor constraint is not binding at w^*, the wage determines the amount of labor employed, instead of vice versa. That is, firms choose productivity and training efficiency of their workers optimally and independently of the amount of labor supplied at any given wage.[1] Equations (7) and (9) yield

$$\frac{\partial w^*}{\partial \pi} = -\frac{C''[Q]Q_\pi Q_w}{\partial^2 AEC/\partial w^2} > 0, \qquad \frac{\partial AEC^*}{\partial \pi} = C'[Q]Q_\pi > 0, \tag{10.1}$$

$$\frac{\partial w^*}{\partial H} = -\frac{C''[Q]Q_H Q_w + C'[Q]Q_{wH}}{\partial^2 AEC/\partial w^2} > 0, \quad \frac{\partial AEC^*}{\partial H} = C'[Q]Q_H < 0, \tag{10.2}$$

$$\frac{\partial w^*}{\partial A} = \frac{T'[A]}{T[A]^2(\partial^2 AEC/\partial w^2)} > 0, \qquad \frac{\partial AEC^*}{\partial A} = -\frac{wT'[A]}{T[A]^2} < 0. \tag{10.3}$$

It is instructive to report both the changes in wages and AECs because the two need not move in the same direction. While wages are an important part of AEC, so are the training costs, which depend on the quality of the applicant pool. If wages rise due to an increase in the relative price, then the AEC increases because firms see the quality of their applicant pool deteriorate as the value of the marginal product in agriculture rises. If, however, wages increase because firms face higher levels of observable human capital or technology, the AEC declines. In response to an increase in technology, firms raise the efficiency wage to offset increased training costs with higher quality workers. Since higher quality workers possess a comparative advantage in training and learning about more sophisticated technologies, the AEC declines.

An increase in human capital, at any given level of technology, implies two effects akin to a change in the relative price and technology.

1. For the purposes of this chapter we assume that the manufacturing sector's labor supply exceeds labor demand at any given wage. The analysis of excess labor demand in adverse selection models is standard and can be reviewed in Weiss (1990).

While higher human capital workers also have higher reservation wages, firms see the average quality of the applicant pool increase. Since high human capital workers also possess a comparative advantage in learning about new technologies, the net effect of increasing the level of human capital is upward pressure on the efficiency wage, but lower AEC.

In competitive models, when wages are directly tied to the marginal product of workers, or in models of perfect information, wages and output (employment) are usually inversely related. This need not be so in the case when firms pay efficiency wages that are influenced by worker quality or technological skill requirements. Equations (3) and (8) yield straight forward relations between manufacturing output and the three key variables:

$$\frac{\partial Y^*}{\partial \pi} = T[A]\,F'[\cdot]\frac{\partial L^{Y^*}}{\partial \pi} = \frac{F'[\cdot]}{F''[\cdot]}\frac{\partial AEC^*}{\partial \pi} < 0, \tag{11.1}$$

$$\frac{\partial Y^*}{\partial H} = T[A]\,F'[\cdot]\frac{\partial L^{Y^*}}{\partial H} = \frac{F'[\cdot]}{F''[\cdot]}\frac{\partial AEC^*}{\partial H} > 0, \tag{11.2}$$

$$\frac{\partial Y^*}{\partial A} = F'[\cdot]\left(T'[A]\,L^Y + T[A]\,\frac{\partial L^{Y^*}}{\partial A}\right) = \frac{F'[\cdot]}{F''[\cdot]}\frac{\partial AEC^*}{\partial A} > 0. \tag{11.3}$$

Manufacturing output falls as the relative price increases, since the increase in the value of the marginal product in agriculture lowers the quality of the applicant pool in manufacturing. Firms are forced to increase the efficiency wage, just to hold training costs constant. The increase in the average and total cost decreases profits which induces a contraction in output through employment. The partial equilibrium effects of increases in human capital and technology are positive on output in the manufacturing sector. Output in the manufacturing sector increases in both cases because workers are more productive. However, in the case of increased technology firms also face higher training costs. As firms are faced with increased training costs, they raise the efficiency wage to attract more able workers with a comparative advantage in training.

2.3 Demand

Most efficiency wage models, with the exception of Phelps (1994), abstract from an explicit demand side and are thus susceptible to nagging doubts that general equilibrium considerations might overturn the partial equilibrium results. To establish a meaningful notion of comparative advantage, we turn to the demand side to construct a two sector general equilibrium adverse selection model. Our introduction of two qualitatively different sectors not only represents the informational asymmetries across sectors and countries, but also allows for a complete and standard demand side that permits for a meaningful discussion of relative prices and trade.

Agents maximize utility, U, which is a function of their consumption of the agricultural and manufacturing good, X and Y,

$$U = a \ln X + \ln Y, \tag{12}$$

subject to their individual budget constraints that are determined by their income derived from their ownership in firms, plus their efficiency wage income or their income from self- employment in agriculture. Utility maximization yields the standard relation between relative demand and relative price

$$a\frac{Y}{X} = \pi. \tag{13}$$

3. Static Comparative Advantage

The condition that supply must equal demand in the closed economy, or

$$a\frac{F\left[T[A]\,L^Y\right]}{\sum_{i=0}^{L-L^Y} G[\theta(i), H]} = \pi \tag{14}$$

renders the equilibrium relative price in the closed economy, $\pi^* = \pi^*[A, H, L, a]$, a function of technology, observable human capital, the population size, preferences, and the ability distribution of a country. We refrain from assuming any specific distribution for θ and assert that these distributions are identical across countries. The comparative statics that involve the size of the labor force, L, are then based on the assumption of mean and spread preserving increases in the population and its abilities.[2] Differentiation of (14) yields the following insights into the static comparative advantage:

$$\frac{\partial \pi^*}{\partial L} = -\frac{a\frac{Y}{X^2}\frac{\partial X^*}{\partial L}}{1 - a\frac{\partial (Y/X)}{\partial \pi}} < 0, \tag{15.1}$$

$$\frac{\partial \pi^*}{\partial H} = \frac{\frac{a}{X^2}\left(X\frac{\partial Y^*}{\partial H} - Y\frac{\partial X^*}{\partial H}\right)}{1 - a\frac{\partial (Y/X)}{\partial \pi}} > 0, \tag{15.2}$$

$$\frac{\partial \pi^*}{\partial A} = \frac{\frac{a}{X^2}\left(X\frac{\partial Y^*}{\partial A} - Y\frac{\partial X^*}{\partial A}\right)}{1 - a\frac{\partial (Y/X)}{\partial \pi}} > 0, \tag{15.3}$$

2. It is easily proven that mean preserving increases in the spread of distributions increase the AEC.

$$\frac{\partial \pi^*}{\partial a} = \frac{\frac{Y}{X}}{1 - a\frac{\partial (Y/X)}{\partial \pi}} > 0. \tag{15.4}$$

The denominator is positive in all cases, as it is simply one minus the slope of the relative supply curve. An increase in π has two separate effects. First, it decreases employment in manufacturing, as explained in (11.1); second, it depresses the expected quality of the applicant pool, which elevates the AEC. Despite firms' attempts to raise the efficiency wage to diminish the deterioration in the quality of their applicant pool, quality declines.[3] Output of X (Y) rises (falls) unambiguously.

Once the sign of the denominator is established, the responses of the relative price due to changes in population and preferences are simple to sign. Ceteris paribus, a mean and spread preserving increase in the population depresses the relative price of the agricultural good, but both sectors expand. Initially, all new workers would start in the agricultural sector, since the labor demand in manufacturing is solely determined by the efficiency wage condition. As the value of the marginal product in agriculture declines, firms can offer lower efficiency wages and attract the same expected quality applicant pool, which encourages manufacturing employment. However, since the efficiency wage is independent of the amount of labor in the economy, or even the amount of labor supplied at any given wage offered, the agricultural sector's increase in supply dominates and the relative price falls. Hence countries with larger populations, even if they have the same level of observable human capital, exhibit lower efficiency wages and have a comparative advantage in the agricultural good, simply due to the adverse selection problem that firms face.

Greater preferences for the agricultural good raise the value of the marginal product in agriculture, which increases the training cost as the quality of workers forthcoming at any given wage declines. Output of the agricultural good rises while that of the manufacturing good declines. Increases in the level of technology and human capital increase the relative supply of the manufacturing good. In both cases the relative supply effect dominates the downward price pressure from the demand side. A higher level of technology implies a lower productivity adjusted wage. This allows firms to employ higher quality workers by increasing the efficiency wage. The quality of workers in manufacturing increases and with it the level of output, while output of the agricultural good decreases unambiguously. A higher level of human capital has the same effect. At any given wage, firms can attract higher quality workers, which lowers

3. $\frac{dQ}{d\pi} = Q_\pi \left(\frac{C'[Q]Q_{ww}}{C''[Q]Q_w^2 + C'[Q]Q_{ww}} \right).$

their AEC. Firms will raise the efficiency wage to compensate for the increased training efficiency and hire more workers. While fewer workers remain in the agricultural sector, and while the manufacturing sector is hiring workers of higher quality than before, the increase in H also increases the agricultural productivity, $G[\theta(i), H]$, which is taken to be dominated by the increase in output in the manufacturing sector.[4] In summary, a country with lower levels of technology or human capital, and with greater labor endowment or preferences for the agricultural good will have a comparative advantage in the agricultural sector.

Having determined the effect of a change in the equilibrium price on the key variables, we can now derive the general equilibrium effects of higher technology and human capital on the efficiency wage.

$$\frac{dw^*}{dA} = \frac{\partial w^*}{\partial \pi}\frac{\partial \pi^*}{\partial A} + \frac{\partial w^*}{\partial A} > 0, \tag{16.1}$$

$$\frac{dw^*}{dH} = \frac{\partial w^*}{\partial \pi}\frac{\partial \pi^*}{\partial H} + \frac{\partial w^*}{\partial H} > 0. \tag{16.2}$$

Increases in human capital and technology cause positive direct and indirect effects on the wage. Higher levels of technology or observable human capital increase the relative price of the agricultural good (15.2)–(15.3), which exerts upward pressure on the efficiency wage, because it decreases the quality of workers forthcoming at the current efficiency wage. In addition, higher levels of technology or observable human capital also induce firms to raise wages directly to lower their AEC, since higher quality workers have a comparative advantage in learning about new technologies (10.2)–(10.3).

In summary, the introduction of informational asymmetries, adverse selection and efficiency wages generates a pattern of static comparative advantage that is akin to both the Heckscher-Ohlin and the Ricardian model. As in the Ricardian model, the country with the more sophisticated technology features the higher wages and exports the good that is technology intensive. Reminiscent of the Heckscher-Ohlin model, the observable human capital abundant country possesses a comparative advantage in the learning intensive good. Hence this country adopts more technology and exports the technology intensive good. Unlike the Heckscher-Ohlin model, however, if two countries share the same technology, the human capital abundant country features the higher return to human capital in autarchy. Firms in the human capital abundant

4. To avoid perverse price responses we only consider the case we find intuitively most compelling, i.e., we restrict ourselves to distributions of $G[\cdot]$ that render the elasticity of output with respect to H smaller in agriculture than in manufacturing when the amount of labor and the relative quality in agriculture decline.

country can afford to pay higher wages because workers possess training efficiency. In fact, the higher quality of the applicant pool renders the AEC comparatively lower in the human capital abundant country. In general we find that

$$\frac{dAEC^*}{dA} = \frac{\partial AEC^*}{\partial \pi}\frac{\partial \pi^*}{\partial A} + \frac{\partial AEC^*}{\partial A} = C'[Q]Q_\pi \frac{\partial \pi^*}{\partial A} - \frac{wT'[A]}{T[A]^2} \gtrless 0, \quad (17.1)$$

$$\frac{dAEC^*}{dH} = \frac{\partial AEC^*}{\partial \pi}\frac{\partial \pi^*}{\partial H} + \frac{\partial AEC^*}{\partial H} = C'[Q]Q_\pi \frac{\partial \pi^*}{\partial H} + C'[Q]Q_H \gtrless 0. \quad (17.2)$$

The AEC is convex in both the level of human capital and in the level of technology.[5] The AEC is convex in the level of technology since eventually the decline in the productivity adjusted wage is dominated by the increase in the training cost. We already know that if we assume identical prices and technologies, the minimum AEC will be lower in the country with the higher human capital (10.2). If prices are allowed to vary, however, eventually the increase in relative price of the agricultural good is sufficiently strong to decrease the quality of the applicant pool to such an extent that the cost of attracting higher ability workers outweighs the benefit of paying lower productivity adjusted wage.

Similarly, at any given efficiency wage the quality of the applicant pool rises when the level of human capital increases. The associated decline in the training cost lowers the AEC only until the decrease in the quality of the applicant pool, due to an increase in the relative price of the agricultural good, dominates. The change in the AEC raises the question if it would be profitable for a firm to train workers or adopt new technologies forever, since AEC eventually rises.

4. Exogenous Technological Change and Endogenous Adoption

To examine the robustness of the comparative static results, we introduce dynamic elements into the model, namely technological change and endogenous adoption. In learning about the process of technology adoption and how this affects profits, we provide a foundation for the examination of trade and firm location. To build intuition, we begin by examining the equilibrium for the case where technological change evolves exogenously at rate γ, or $dA/dt = \gamma A$. In section 6 we then characterize the dynamics when technological change is endogenous.

This section's assumption of one "world technology", A_t, that evolves exogenously is identical to the assumption in Mankiw, Romer,

5. Convexity of the AEC curve with respect to A requires a strong effect of technology on the relative price, $\partial^2\pi/\partial A^2 > 0$, and a declining productivity adjusted wage in technology, $d(w/T[A])/dA < 0$.

and Weil (1992), which produced good fits in cross country growth regressions. From (3), (6), and (17) we can now find the effects of a change in technology on equilibrium profits and output:

$$\frac{d\rho_t}{dA_t} = -L_t^Y \frac{dAEC_t^*}{dA_t}, \tag{18}$$

$$\frac{dY_t}{dA_t} = \frac{F'[\cdot]}{F''[\cdot]} \frac{dAEC_t^*}{dA_t}. \tag{19}$$

The dynamic analysis of production and adoption implies a minimum AEC locus that depends only on the exogenous parameters of the model, ($\overline{AEC}^*[A; H, a, L]$ in figure 1). Equations (18) and (19) imply that firms adopt new technologies only up to the point where the $\overline{AEC}^*$ is at its minimum. If firms continued to adopt technology beyond the minimum $\overline{AEC}^*$, output and profits would contract. Hence, the dynamic analysis assures us that we can rule out all cases where firms would ever be on the upward sloping part of the $\overline{AEC}^*$.

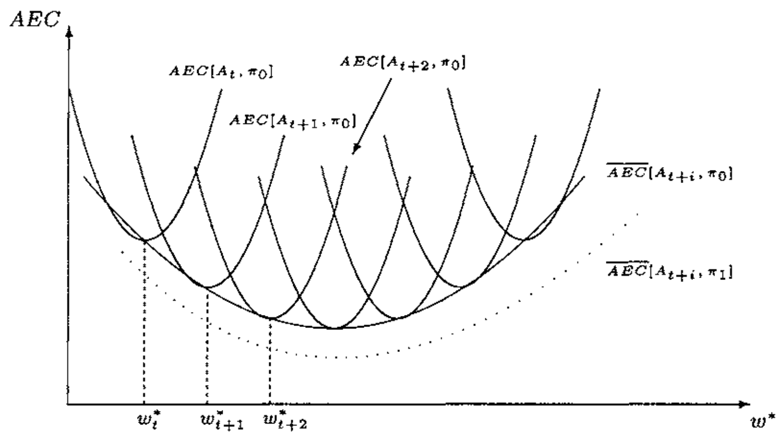


Fig. 1. Average efficiency cost (AEC) and wages (w) for different levels of technology (A)

The fact that firms cease to adopt new technologies beyond some critical level implies zero long run growth, despite the fact that ever new technologies are available to firms, free of charge. The general equilibrium adverse selection model is thus entirely void of scale effects, in terms of growth rates and *levels*. Recent empirical research has emphasized the apparent absence of scale effects in the data.[6] In this model, the complete absence of scale effects is a function of the constancy of human capital, which implies rising training costs as long as technological adoption continues. The crucial importance of the complementarity of human capital and technology has previously been stressed by Young (1993) and Eicher (1996).

Figure 1 also allows for an examination of how different levels of human capital affect the rate of technological adoption. The dotted line shows the $\overline{AEC}^*$ for a country with a higher level of human capital than the country with the solid line. From (17.2) and (16.2) we know that such a country would possess a lower $\overline{AEC}^*$ and a higher efficiency wage, which implies that countries with a relatively higher human capital, and with a comparative advantage in learning about new technologies, adopt a relatively more sophisticated technology.

We can also combine the dynamic effect of technological change with our previous insights into how changes in price alters the $\min \overline{AEC}^*$, or the maximum level of technology adopted by firms. If the price of the manufacturing good rises (for example, because a country with a comparative advantage in the manufacturing good opens to international trade), the $\overline{AEC}^*$ shifts down at any given level of technology and wages decrease (10.1). However, we can show that

$$\frac{d^2AEC^*}{dA\,d\pi} = C''[Q]Q_\pi \frac{dQ}{dA} + C'[Q]Q_{\pi\pi}\frac{\partial \pi^*}{\partial A} > 0, \tag{20}$$

which implies that as revenues (temporarily) outweigh training costs, firms deem further adoption of technology profitable. Hence, trade induces new, but temporary, incentives for technological adoption and growth in the advanced country. Conversely, the LDC would find even fewer incentives to adopt new technologies, because revenues fall in the manufacturing sector.

5. Trade and Firm Location

5.1 Informational Efficiency Gains from Trade

Since long run growth is zero and firms never adopt technologies that raise their AEC, even if the use of that technology was free, we return to the static analysis of trade and firm location, without loss of generality. We simplify matters further by restricting ourselves to the small

6. Strong empirical evidence for non-scale growth has been presented by Easterly et. al. (1993) in a large cross country data set, and in careful analysis by Jones (1995) for OECD countries. For a general theoretical discussion of non-scale models see Eicher and Turnovsky (1996).

open economy analysis below. To derive crucial intuition about the trade pattern, we assume initially that countries differ in human capital, but share identical technologies. Then we examine the implications for firm location when countries differ also in their levels of technology.

Let us commence by designating the country with the higher (lower) level of human capital, H_1 (H_0), as DC (LDC). Both countries have identical preferences and population sizes. With identical technologies, the DC has an absolute advantage in both sectors due to its higher level of observable human capital. As shown above, the DC possesses a comparative advantage in the manufacturing sector (15.2), since workers with higher human capital possess greater training efficiency and generate lower AECs. This implies a greater relative supply of the manufacturing good and relatively higher efficiency wages, compared to the LDC (16.2).

Opening to trade decreases (increases) the relative price of the agricultural good in the DC (LDC), which induces a downward (upward) shift of the AEC curve (10.1). The decline (rise) in the relative price also depresses (raises) the efficiency wage and shifts the average cost curve left (right) (10.1). The relation between AEC and the efficiency wage is the essence of the model and it is graphed in figure 2. Figure 2 reports the AECs under autarchy (solid lines) and free trade (dotted lines). The two country analysis reveals cross country wage convergence due to international trade. In the Heckscher-Ohlin model, trade also induces wage convergence but through low (high) and rising (falling) manufacturing wages in the human capital abundant (short) country. Empirically we observe, however, relatively higher average wages in skill abundant countries.[7]

While the effects of opening to trade might be standard, its *mechanism of adjustment* is novel and interesting. Wage convergence signals a new effect generated by the introduction of adverse selection to the theory of international trade: the movements in the efficiency adjusted cost curves reflect what we term *informational efficiency gains from trade.* The informational efficiency gains from trade arise because both countries relocate production to the sector where workers have a comparative advantage in terms of production *and* informational efficiency. The country with the comparative advantage in training expands production of the training intensive good, pays lower efficiency wages, but also enjoys a higher quality applicant pool. That is, trade diminishes the effect of the informational asymmetries in the DC's manufacturing sector.

7. For empirical evidence on cross country wage convergence see, among many others, Eicher (1995) and Davis (1992).

Under autarchy, the high relative supply of the agricultural good in the DC depresses the quality of the applicant pool at any given wage, and a large fraction of high quality workers are drawn into agriculture. This effect increases the informational cost to the DC's manufacturing firm and contributes to the high efficiency wage offer. Opening to trade creates profit incentives to expand output in manufacturing and lowers the reservation wage in the agricultural sector. The latter effect increases the quality of the applicant pool for manufacturing firms. This lowers the DC's manufacturers' *AEC*s and lowers the efficiency wage, because the quality of the applicant pool increases. Hence the term informational efficiency gain from trade.

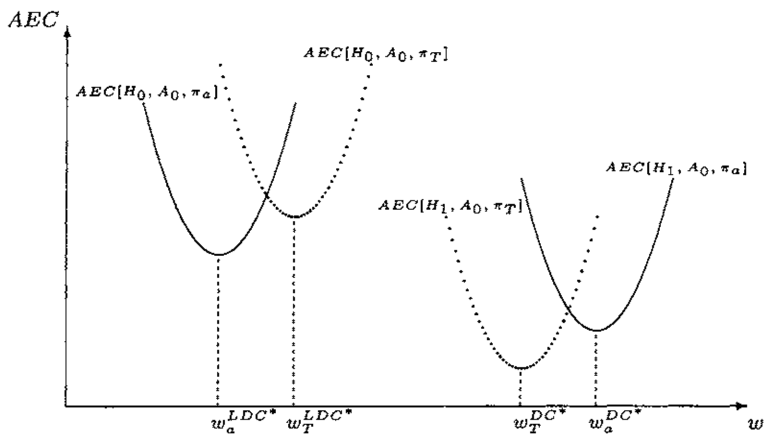


Fig. 2. Trade induced changes in *AEC* for countries that differ only in human capital endowments (*H*)

While the empirical evidence for efficiency wages is scant and often inconclusive, there exists support for the wage/efficiency adjustment mechanism outlined above. Tests of efficiency wage theories have produced no conclusive support for shirking or monitoring models, but Foster and Rosenzweig (1993) found the evidence for the existence of efficiency wages due to adverse selection. Their study reported that variations in reservation wages positively influence efficiency wages paid. Krueger (1988) previously established that higher wages increase the quality of the applicant pool.

Informational gains also accelerate the adjustment and increase the output of the agricultural good in the LDC. As long as the LDC is incompletely specialized, its manufacturing firm must pay a higher efficiency wage under free trade. Because of its excessively low autarchy price, firms in the LDC were able to attract excessively high quality workers to the manufacturing sector at relatively low efficiency wages. As the value of the marginal product in agriculture rises under trade, the quality of the applicant pool in manufacturing declines, which provides an added incentive to contract the manufacturing sector. Both countries experience a Pareto improvement because of the standard static consumption, production and the additional informational efficiency gains from trade.

5.2 Firm Location

Thus far we have embedded an adverse selection model into a general equilibrium framework and added the insights of the informational efficiency gains from trade and wage convergence due to efficiency gains. The analysis of section 5.1 does not lend itself to the analysis of firm location since firms share identical technologies. As mentioned in the introduction, a prerequisite to the analysis of firm location is that there exists (i) a factor that is internal to the firm (in this case firm-specific training), and (ii) a factor that provides a unique ownership advantage to the firm (in this case technology).

To examine the decision to locate, we must introduce countries that differ not only in observable human capital, but also in technology. This is a natural assumption in light of our discussion of endogenous adoption, where we have shown that the countries with higher levels of human capital also adopt higher levels of technology in the long run. Let us then redefine DC (LDC) as the country which has relatively higher (lower) levels of technology, A_1 (A_0), in addition to relatively higher (lower) levels of observable human capital, H_1 (H_0). Both countries continue to have identical preferences and population sizes.

The introduction of technological differences in addition to human capital differences, intensifies the DC's comparative advantage in the manufacturing sector. The DC now features an even greater relative supply of the manufacturing good, and its manufacturing sector pays an even greater efficiency wage, compared to the LDC (15.2)–(15.3) and (16). Figure 3 shows that opening to trade, the associated decline in the relative price of the agricultural good in the DC leads again to familiar shifts in *AEC*s and wages, as discussed in figure 2. In the LDC, the increase in the value of the marginal product in agriculture increases the reservation wage and makes it even more difficult to attract workers for manufacturing. The LDC's manufacturing sector is hurt not only by the price effect but also by the informational aspect that the quality of the applicant pool declines, as discussed above. Again we observe wage convergence across countries.

Having introduced technological differences between the countries, we can draw an additional average efficiency curve (bold dotted) that represents the AEC of a firm which produces the manufacturing good with technology A_1, and human capital H_0. We know from (10.2)–(10.3) that

$$AEC^*[H_0, A_0, \pi] > AEC^*[H_0, A_1, \pi] > AEC^*[H_1, A_1, \pi],$$
$$w^*[H_0, A_0, \pi] < w^*[H_0, A_1, \pi] < w^*[H_1, A_1, \pi].$$

That is, if the DC's manufacturing firm were to locate part of its production to the LDC, the relocation would require that firm to train LDC workers to work with DC technology. Because the LDC cohort is of observationally lower quality, the DC's AEC (wage) in the LDC is higher (lower) than in the DC, (16) and (10). This is the human capital dimension to direct foreign investment. Also, since the DC's plant in the LDC trains workers to produce with higher technology, A_1, than the LDC firm uses, A_0, the DC's subsidiary in the LDC pays a higher wage, attracts higher quality workers and provides more training than the LDC's manufacturing firm. Hence firm location generates a multiple-wage equilibrium, as seen in figure 3.

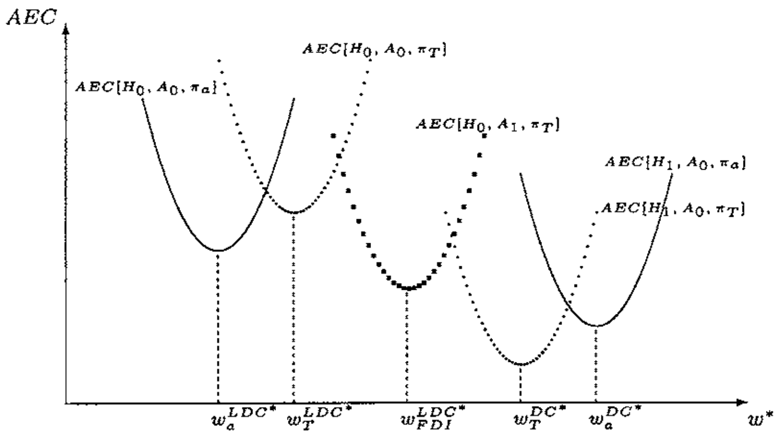


Fig. 3. Trade induced changes in AEC for countries that differ in human capital and in technology

Notice, however, that, with two firms, a strategic interaction enters into the wage setting mechanism, because firms offering low wages improve the quality of distribution of workers applying to higher wage firms. This problem is well known and has previously been addressed by Weiss (1990). To proxy the labor market interaction between domestic and foreign firms, it is useful to assume that expected quality is also a negative function of the wage offered by a competitor firm, w^c, or $Q[w, H, \pi, w^c]$ with $\partial Q[\cdot]/\partial w^c < 0$. This implies that if a competitor enters with a higher wage, $w^c > w$, the competitor firm skims the cream of the crop workers from the distribution and lowers the expected quality of workers attracted by the firm offering wage w.

Let us analyze the wage setting mechanism for the case where the MNC enters with a higher technology than the domestic firm. Formally, we require a sequence of actions where first, each firm announces a wage offer and the number of job openings. We assume that workers have rational beliefs and cannot apply to more than one firm. Second, workers decide where to apply after examining the wage offers and the probability of getting hired. The higher wage and technology of the MNC forces the local firm to offer higher wage than before to avoid quality deterioration of its applicant pool. To formalize this thought, consider the first order conditions of the *AEC* minimization problem for the domestic firm and the MNC, respectively:

$$T[A_0]\ C'\left[Q[w\ , H, \pi, w^c]\right]\ Q_w\ [w\ , H, \pi, w^c] = -1, \tag{7'}$$

$$T[A_1] C^{c'}\left[Q[w^c, H, \pi, w\ ]\right]\ Q^c_{w^c}[w^c, H, \pi, w\ ] = -1, \tag{7''}$$

where superscript c denotes the MNC, and H represents the domestic country's level of observable human capital that both firms utilize. Given the properties of the cost function, $T[A_0] < T[A_1]$ implies $w^c > w$. Note that the greater the technology gap, the larger the differential in the wage offers. Equation (7′) is the reaction function of the local firm. It determines the wage it offers at each level of the MNC's wage, taking the latter as given. The same holds for (7″). The Nash equilibrium of the model is then given by the solution of the system of (7′)–(7″):

$$w\ = w\ \left[A_0, H, \pi, w^c[A_1, H, \pi, w\ ]\right], \tag{21}$$

$$w^c = w^c\left[A_1, H, \pi, w\ [A_0, H, \pi, w^c]\right], \tag{21'}$$

but from (7′)–(7″) we know that

$$\left.\frac{dw}{dw^c}\right|_{(21)} = -\frac{C_{ww^c}}{C_{ww}} > 0, \qquad \left.\frac{dw^c}{dw}\right|_{(21')} = -\frac{C^c_{w^c w}}{C^c_{w^c w^c}} > 0.$$

Hence the equilibrium can be easily characterized, since both reaction functions are upward sloping in the w, w^c space, with each intercept being the wage a firm would offer if no competitor was in the market.

To ascertain if opening a subsidiary in the LDC is profitable, we must simply examine the first order condition of the MNC and examine if opening a plant in the LDC would provide positive profits. It is obvious from the profit condition that, depending on the DC's level of technology and the LDC's level of human capital, a production location in the LDC might not be profitable for a DC manufacturing firm. That is because the high training cost for the MNC would not be covered by the revenues. Hence, relocation becomes less and less likely the farther apart the levels of both human capital and technology are, because training costs increase as technology levels rise and/or observable human capital levels decline. This phenomenon explains not only why the lion share of FDI is among relatively similar countries, see Markusen and Venables (1995), but also the finding of Borensztein et. al. (1995) that the level of human capital is important to the success of FDI.

The LDC's firm now finds an additional impediment to production of the manufactured good. First, the price effect due to international trade raises its wage and depresses the quality of the applicant pool. Second, as the DC sets up a subsidiary and pays a higher efficiency wage, the DC skims off the high quality workers, which depresses the quality of the applicant pool for the LDC firm, yet again, as $\partial Q/\partial w^c > 0$. Both effects work to diminish the incentives for the LDC firm to produce and increase the likelihood that it will be driven out of business because it cannot generate positive profits.

6. Endogenous Technological Change under Asymmetric Information

6.1 Endogenous Technology

We start by modifying the production function of the manufacturing firm to reduce the complexity of endogenous invention process. We assume that output is linear in technology, or

$$Y_t = A_t\, F[L_t^Y]. \tag{3'}$$

Once we allow for endogenous technological change, firms hire not only production workers, but also research workers, R_t, to produce new technology according to the standard technology production function

$$\dot{A} = \phi[R_t]A_t, \tag{22}$$

where $\phi[\cdot]$ is assumed to satisfy the Inada conditions, with the exception of $\phi'[0]$, which is assumed to equal the constant γ. The slight modification of the Inada condition is necessary to insure stability of the system.

Since firms conduct research in and produce with firm-specific technology, firms must now train not only production but also research workers. In hiring for both types of positions, firms face an applicant pool with uncertain quality. Again, we will find that firms maximize profits by offering an efficiency wage to mitigate the informational asymmetry. The profit function of the firm can then be written as

$$\rho_t = Y_t - (w_t + T[A_t]\, C[w_t])\, (R_t + L_t^Y), \tag{23}$$

where $C[w_t]$ is a short for $C[Q[w_t, H, \pi]]$, to simplify the notation until we discuss comparative statics.

The manufacturing sector in our economy now solves the following maximization problem:

$$\begin{aligned} \max_{L_t^Y, R_t, w_t, A_t} \quad & \int_0^\infty \left(Y_t - (w_t + T[A_t]\, C[w_t])(R_t + L_t^Y)\right) e^{-\delta t} dt \\ \text{s.t.} \quad & \dot{A} = \phi[R_t] A_t, \end{aligned}$$

where δ represents the rate of time preference. Maximizing the Hamiltonian yields the following first order conditions:

$$F'[L^Y] = \frac{w + T[A]\, C[w]}{A} \equiv AEC[A], \tag{24}$$

$$\lambda\delta - \dot{\lambda} = F[L^Y] + \lambda\phi[R] - (L^Y + R)T'[A]\, C[w], \tag{25}$$

$$C'[w] = -\frac{1}{T[A]}, \tag{26}$$

$$\lambda\phi'[R] = \frac{w + T[A]\, C[w]}{A}. \tag{27}$$

We also add the transversality condition that

$$\lim_{t \to \infty} \lambda_t A_t e^{-\delta t} = 0. \tag{28}$$

Equations (24)–(25) indicate that the productivity adjusted marginal products of research and production workers must equal their productivity adjusted cost (AEC). In the case of research workers, the AEC is weighted by the shadow value of technology, λ. Equation (26) is the familiar efficiency wage condition that determines the wage offered by manufacturing firms on the basis of training cost, independent of the labor supply.

Any long run equilibrium requires that the growth rate of research employment goes to zero (otherwise the AEC would reach infinity). This implies from (25) that the growth rate of the average efficiency cost must be zero in equilibrium. It follows immediately that the growth rate of the shadow value of technology, and the growth rate of labor in production must be zero, too. As research employment declines, and its productivity increases, the shadow value of technology declines. The only equilibrium value of λ that satisfies (25) is then AEC/γ. From (24)–(25) can the be utilized to establish a relationship between the rates of change of employment in the two sectors:

$$\dot{\lambda}\phi'[R] + \lambda\phi''[R]\dot{R} = \frac{dAEC}{dA}\dot{A}, \tag{29}$$

$$\dot{L}^Y = \frac{dAEC/dA}{F''[\cdot]}\dot{A}, \tag{30}$$

which allow us to summarize the steady state as $\dot{\lambda} = \dot{R} = \dot{A} = \dot{L}^Y = dAEC/dA = 0$.

6.2 Dynamics

Substituting (24)–(26) into (27) and the accumulation constraint, we can summarize the differential equations that determine the dynamics of the model,

$$\dot{A} = A\,\phi\left[\phi'^{(-1)}\left[AEC[A]/\lambda\right]\right], \tag{22'}$$

$$\begin{aligned}\dot{\lambda} = {} & \lambda\delta + \left(\phi'^{(-1)}\left[AEC[A]/\lambda\right] + F'^{(-1)}\left[AEC[A]\right]\right) \\ & \cdot T'[A]\,C\left[C'^{(-1)}[-1/T[A]]\right] - F\left[F'^{(-1)}\left[AEC[A]\right]\right] \\ & -\lambda\phi\left[\phi'^{(-1)}\left[AEC[A]/\lambda\right]\right].\end{aligned} \tag{27'}$$

Equations (22′) and (27′) can be used to draw the phase diagram in the λ, A space. The slopes of the $\dot{\lambda} = 0$ and $\dot{A} = 0$ lines around the equilibrium can readily be obtained from (22′) and (27′)

$$\left.\frac{d\lambda}{dA}\right|_{\dot{A}=0} = \frac{dAEC}{dA}\frac{1}{\gamma}, \tag{31}$$

$$\left.\frac{d\lambda}{dA}\right|_{\dot{\lambda}=0} = -\frac{L^Y\left(\frac{(T'[A]\,C'[\cdot])^2}{T[A]\,C''[\cdot]} + C[\cdot]T''[A]\right)}{\delta} < 0. \tag{32}$$

Equation (31) represents the convex $\dot{A} = 0$ line that slopes downward before its minimum, and intercepts the line $\dot{\lambda} = 0$ at $\lambda = AEC[A^*]/\gamma$.

Further research would increase training costs in excess of the marginal benefit to the firm. Note also, that firms cannot be forced to adopt a technology that provides negative profits, and that it can "jump" to use older technology even if several generation of newer technologies are available. The $\dot{\lambda} = 0$ line is given in (32), and is downward sloping due to the convexity of the AEC in A, e.g. $d(T'[A]\,C[w])/dA > 0$.

The phase trajectories that map out the dynamic moments of the system indicate in figure 4 that any loci off the saddle path violate the transversality condition. For any $\dot{\lambda} = 0$ technology is increasing, and for any A to the west of the $\dot{\lambda} = 0$ demarcation line $\dot{\lambda} = 0$, so that the shadow value of technology declines as technology accumulates. The equilibrium is a saddle point with a downward sloping stable and an upward sloping unstable branch.

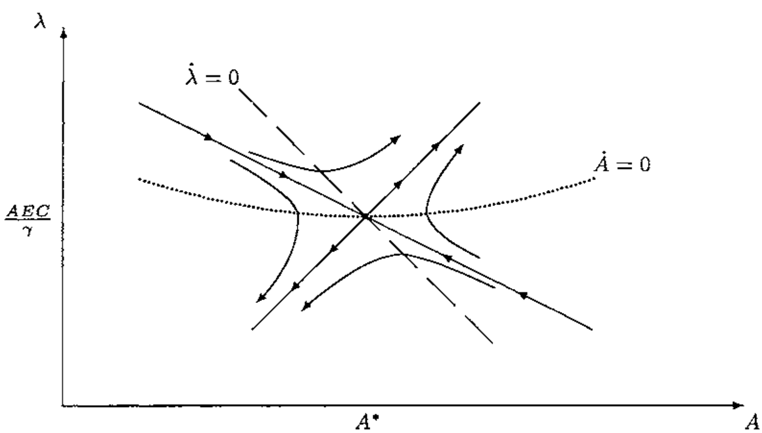


Fig. 4. Dynamics of endogenous R&D and technology adoption

The analysis of the local phase diagram can be confirmed by a complete local stability analysis. The linearization around the steady state yields

$$\begin{bmatrix} \dot{A} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{A\gamma^2}{\lambda\phi''[0]} \\ L^Y\Big(-\frac{(C'[w]T'[A])^2}{T[A]\,C''[w]} + T''[A]\,C[w]\Big) & \delta \end{bmatrix} \begin{bmatrix} A - A^* \\ \lambda - \lambda^* \end{bmatrix}. \quad (33)$$

The Jacobian reveals immediately that the determinant is negative for reasonably small values of the rate of time preference, which confirms our analysis of the phase diagram that the equilibrium is locally saddle point stable.

6.3 Dynamic Adjustment to a World Price Shock

Above we have briefly discussed the possibility of receiving a technology spillover, and under what circumstances a country would be willing to adopt such a windfall. As seen above, the economy will not adopt a new technology beyond A^* because training costs, just for production workers, would exceed the revenues from sales of Y, given their level of ability and human capital at a given world price. A change in the world price changes the incentives to produce technology permanently.

Our representative example will be a decline in the relative price of the agricultural good. In that case the small open economy finds it relatively more profitable to expand its manufacturing production and its technology production. Here it is helpful to recall the intuition we built in the static model. There it was shown that a decline in the relative price of manufacturing provides a higher quality of workers at the same efficiency wage. For any given level of technology, training costs shift down and profits in manufacturing rise again due to the fall in π.

The new equilibrium is characterized by a lower AEC, which must be due to a higher level of technology. Hence the $\dot{\lambda} = 0$ line can be shown to shift East. The transition is described in figure 5. At the old level of technology, the shadow value of another unit of technology is now positive. The economy jumps onto the new transition path, E_1, hires workers into the R&D sector and sees the shadow value decline as it moves to the new equilibrium A_1^* at E_2.

Finally a word on the dynamic effect of firm location. As discussed in the formal location analysis, the entry of a higher wage and technology multinational forces the domestic firm to offer a higher wage. From our dynamic analysis we know that this can be achieved only in two ways. First, if the country has not yet met its steady state, it will reduce its rate of technology accumulation and arrive at a lower steady state, one characterized by a higher AEC. If, on the other hand, the country has already reached its steady state, it will not only have to raise its wage, but also lower the technology employed because it faces a reduction in the quality of the applicant pool. Here the analysis is analogous to the effects on the local firm if the relative price of the manufacturing good declines. Hence under asymmetric information the multinational results in a dumbing down of the production process in the domestic firm, and an increase in the average wage paid in manufacturing, which leaves the country better off in utility terms but certainly not in terms of its level of GDP.

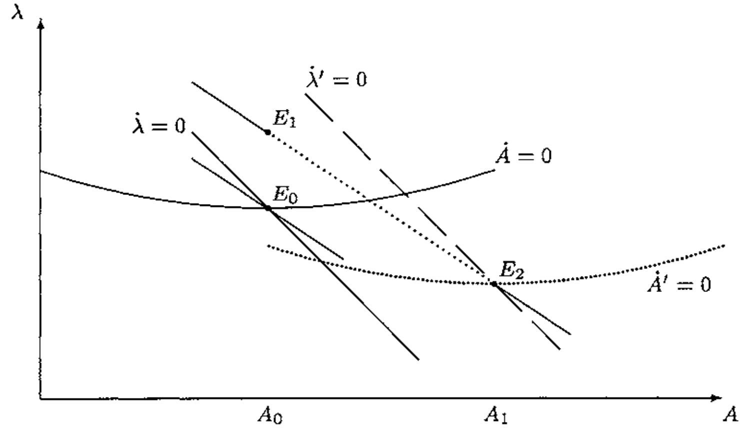


Fig. 5. Dynamic changes in R&D and technology adoption due to international trade

7. Summary and Concluding Remarks

This chapter explores the human dimension to FDI: informational asymmetries as MNCs must train workers to work with firm specific technology. Figure 3 summarizes the important conclusions of this chapter. First, it exhibits the informational gains from trade, as the country with the comparative advantage in training enjoys a higher quality applicant pool and lower efficiency wages. Second we find that the farther the cost curves are apart, the less likely FDI will be. As the cost curves indicate both the difference in the levels of technologies and human capital between the countries, we know that similar countries are more likely to receive FDI. Countries that do not provide a minimum level of human capital cannot attract technologically superior FDI, because MNCs find that the average cost of training is too high.

Most importantly efficiency wages can explain why FDI does not raise the wage level as a whole for the country, but only for workers employed in the MNC. Informational asymmetries force firms to pay wages that control the quality of the applicant pool, rather than clear the labor market in manufacturing. This chapter also shows that the MNC pays a higher wage than the domestic firm, because the MNC introduces a superior technology, and incurs higher training costs. This provides incentives to raise the wage in order to increase the quality of its applicant pool. Despite working with the same technology, workers in the LDC receive a lower wage, than in the MNC's home country, because the MNC faces lower information costs and higher quality workers in the home country. This generates multiple wage equilibria.

We find in the dynamic analysis that the model is entirely void of scale effects. That is, growth ceases in this model, even if ever more sophisticated technology were available, because the cost of adoption would eventually outpace revenues. To introduce a full fledged general equilibrium adverse selection model, we had to make some important simplifications. If both technology and human capital were endogenous, sustained growth would clearly be possible. However, we are certain that it would not overturn the qualitative nature of our location results. We have seen above that international trade mitigates the effect of informational asymmetries. Hence it is not surprising to find that there is a long history in the literature on informational asymmetries that explores the room for policy to achieve welfare improvements. Weiss (1990) and especially Copeland (1989), and Bulow and Summers (1989) addressed the wealth of welfare issues inherent in models of informational asymmetries. A full fledged commercial policy analysis is left for future research.

Acknowledgement. We thank Eric W. Bond, Bjarne S. Jensen, Kar-yiu Wong, and the discussant Jan I. Haaland for helpful comments.

References

Aitken, Brian, Ann Harrison, and Robert E. Lipsey (1995). "Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela, and the United States." NBER Working Paper No. 5102.

Becker, Gary (1975). *Human Capital.* University of Chicago Press.

Borensztein, Eduardo, Jose De Gregorio, and Jong-Wha Lee (1995). "How Does Foreign Direct Investment Affect Economic Growth?" NBER Working Paper No. 5057.

Brecher, Richard A. (1992). "An Efficiency-Wage Model With Explicit Monitoring and Welfare in an Open Economy." *Journal of International Economics* 32: 179–191.

— and Ehsan U. Choudhri (1994). "Pareto Gains From Trade, Reconsidered: Compensating for Jobs Lost." *Journal of International Economics* 36, 223–238.

Bulow, Jeremy I., and Lawrence H. Summers (1986). "A Theory of Dual Labor Markets With Applications to Industrial Policy, Discrimination and Keynesian Unemployment." *Journal of Labor Economics* 4: 376–414.

Copeland, Brian R. (1989). "Efficiency Wages in a Ricardian Model of International Trade." *Journal of International Economics* 27: 221–244.

Davis, Stephen J. (1992). "Cross Country Patterns of Change in Relative Wages." In O. Blanchard and S. Fisher (eds.), NBER Macroeconomic Annual.

Dixit, Avinash (1989). "Trade and Insurance with Adverse Selection." *Review of Economic Studies* 56: 235–248.

— (1994). "Risk-sharing, Adjustment, and Trade." *Journal of International Economics* 36: 263–285.

Dunning, John H. (1977). "Trade Location of Economic Activity and MNE: A Search for an Electic Approach." In B. Ohlin, P.O. Hesselborn, and P.M. Wijkman (eds.), *The International Allocation of Economic Activity.* London, Macmillan: 395–418.

— (1981). *International Production and the Multinational Enterprise.* London: George Allen and Unwin.

Easterly, William, Michael Kremer, Lant Pritchett, and Lawrence H. Summers (1993). "Good Policy of Good Luck? Country Growth Performance and Temporary Shocks." *Journal of Monetary Economics* 32: 459–483.

Eicher Theo S. (1995). "Trade Development and Converging Growth Rates: Dynamic Gains From Trade Reconsidered." University of Washington Working Paper.

— (1996). "Endogenous Human Capital and Technological Change." *Review of Economic Studies* 63: 127–145.

— and Stephen J. Turnovsky. (1996). "Toward a General Model of Economic Growth." University of Washington Working Paper.

Ethier, Wilfred J. (1986). "The Multinational Firm." *Quarterly Journal of Economics* 101: 805–833.

Feenstra, Robert C., and Gordon H. Hanson (1995). "Foreign Direct Investment and Relative Wages." NBER Working Paper No. 5122.

Foster, Andrew, and Mark Rosenzweig (1993). "Information Flows and Distribution in Labor Markets in Rural Areas in Developing Countries." Proceedings of the World Bank Annual Conference on Development Economics 1992: 173–203.

Hummels, David L., and Robert M. Stern (1994). "Evolving Patterns of North American Merchandise Trade and Foreign Direct Investment, 1960–1990." *World Economy* 17: 5–29.

Jones, Charles (1995). "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics* 110: 495–527.

Kalaitzidakis, Pantelis G. (1996). "On-The-Job Training Under Firm Specific Innovation and Worker Heterogeneity." Working Paper No. 96–2, University of Cyprus.

Krueger, Alan B. (1988). "The Determinants of Queues for Federal Jobs." *Industrial and Labor Relations Review* 41: 567–581.

Mankiw, N. Gregory, David Romer, and David N. Weil (1992). "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107: 407–437.

Markusen James R. (1995). "The Boundaries of Multinational Enterprises and the Theory of International Trade." *Journal of Economic Perspectives* 9: 169–189.

— and Anthony J. Venables (1995). "Multinational Firms and the New Trade Theory." NBER Working Paper No. 5036.

Nalebuff, Barry, and Joseph Stiglitz (1982). "Prices and Quality." Princeton University Econometric Research Program Paper No. 297.

Phelps, Edmund (1994). Structural Slumps. Cambridge MA: Harvard University Press.

Weiss, Andrew (1980). "Job Queues and Layoff in Labor Markets with Flexible Wages." *Journal of Political Economy* 88: 526–538.

— (1990). *Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersion.* Princeton University Press.

Young, Alwyn (1993). "Invention and Bounded Learning By Doing." *Journal of Political Economy* 101: 443–472.

Studies in International Trade Policy includes works dealing with the theory, empirical analysis, political, economic, legal relations, and evaluations of international trade policies and institutions.

General Editor: Robert M. Stern

John H. Jackson and Edwin Vermulst, Editors.
Antidumping Law and Practice: A Comparative Study

John Whalley, Editor. *Developing Countries and the Global Trading System.* Volumes 1 and 2

John Whalley, Coordinator. *The Uruguay Round and Beyond: The Final Report from the Ford Foundation Project on Developing Countries and the Global Trading System*

John S. Odell and Thomas D. Willett, Editors.
International Trade Policies: Gains from Exchange between Economics and Political Science

Jagdish Bhagwati and Hugh T. Patrick, Editors.
Aggressive Unilateralism: America's 301 Trade Policy and the World Trading System

Alan V. Deardorff and Robert M. Stern. *Computational Analysis of Global Trading Arrangements*

Ulrich Kohli. *Technology, Duality, and Foreign Trade: The GNP Function Approach to Modeling Imports and Exports*

Stephen V. Marks and Keith E. Maskus, Editors.
The Economics and Politics of World Sugar Policies

J. Michael Finger, Editor. *Antidumping: How It Works and Who Gets Hurt*

Horst Herberg and Ngo Van Long, Editors.
Trade, Welfare, and Economic Policies: Essays in Honor of Murray C. Kemp

Robert M. Stern, Editor. *The Multilateral Trading System: Analysis and Options for Change*

David Schwartzman. *The Japanese Television Cartel: A Study Based on* Matsushita v. Zenith

Barry Eichengreen.
Reconstructing Europe's Trade and Payments: The European Payments Union

Alan V. Deardorff and Robert M. Stern, Editors.
Analytical Perspectives and Negotiating Issues in the Global Trading System

Edwin Vermulst, Paul Waer, and Jacques Bourgeois, Editors.
Rules of Origin in International Trade: A Comparative Study

Alan V. Deardorff and Robert M. Stern, Editors.
The Stolper-Samuelson Theorem: A Golden Jubilee

Kent Albert Jones. *Export Restraint and the New Protectionism: The Political Economy of Discriminatory Trade Restrictions*

Alan V. Deardorff, James A. Levinsohn, and Robert M. Stern, Editors.
New Directions in Trade Theory

Robert Baldwin, Tain-Jy Chen, and Douglas Nelson. *Political Economy of U.S.–Taiwan Trade*

Bernard M. Hoekman and Petros C. Mavroidis, Editors.
Law and Policy in Public Purchasing: The WTO Agreement on Government Procurement

Danny M. Leipziger, Editor. *Lessons from East Asia*

Tamin Bayoumi. *Financial Integration and Real Activity*