NBER WORKING PAPER SERIES

A GLOBAL VIEW OF ECONOMIC GROWTH

Jaume Ventura

Working Paper 11296
http://www.nber.org/papers/w11296

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2005

I dedicate this research to the memory of Rudi Dornbusch, the best mentor, colleague and friend a young economist could have hoped for, and always capable of making others see things differently. This is a draft of a chapter in the forthcoming *Handbook of Economic Growth* edited by Philippe Aghion and Steven Durlauf. I am thankful to Fernando Broner, Gino Gancia and Francesc Ortega for their useful comments and to Matilde Bombardini, Philip Sauré and Rubén Segura-Cayuela for providing excellent research assistance. I am also grateful to the Fundación Ramón Areces for its generous financial support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

A Global View of Economic Growth
Jaume Ventura
NBER Working Paper No. 11296
April 2005
JEL No. F10, F15, F40, F43, O11, O40, O41

ABSTRACT

This paper integrates in a unified and tractable framework some of the key insights of the field of international trade and economic growth. It examines a sequence of theoretical models that share a common description of technology and preferences but differ on their assumptions about trade frictions. By comparing the predictions of these models against each other, it is possible to identify a variety of channels through which trade affects the evolution of world income and its geographical distribution. By comparing the predictions of these models against the data, it is also possible to construct coherent explanations of income differences and long-run trends in economic growth.

Jaume Ventura
CREI
Universitat Pompeu Fabra
Ramon Trias Fargas, 25-27
08005-Barcelona SPAIN
and NBER
jaume.ventura@upf.edu

Table of Contents

"All theory depends on assumptions that are not quite true. That is what makes it theory. The art of successful theorizing is to make the inevitable assumptions in such a way that the final results are not very sensitive. A "crucial" assumption is one on which the conclusions do depend sensitively, and it is important that crucial assumptions be reasonably realistic. When the results of a theory seem to flow specifically from a special crucial assumption, then if the assumption is dubious, the results are suspect."

Robert M. Solow [1956, p. 65]

0. Introduction

The world economy has experienced positive growth for an extended period of time. Figure 1 plots average world per capita income from 1500 to today, using data from Maddison's classic study of long run trends in the world economy. The most salient feature of the growth process is its nonlinear nature. For most of the past five hundred years, the world economy settled in a path of stagnation with little growth. But sometime around the early nineteenth century the world economy entered a path of sustained and even accelerating growth. While per capita income grew only by eighteen percent from 1500 to 1820, it has then grown by more than seven hundred and fifty percent from 1820 to today. And this growth has been far from steady. It averaged 0.53 percent from 1820 to 1870, and more than doubled to 1.30 from 1870 to 1913. Growth declined to 0.91 percent during the turbulent period that goes from 1913 to 1950, and then exploded to an unprecedented 2.93 percent from 1950 to 1973. Since then growth has markedly declined to 1.33 percent, even though this period still constitutes the second best growth performance in known human history.

This economic growth has not been distributed equally across the different regions of the world economy. Figure 2 shows per capita income growth for the different regions of the world economy in various time periods. Differences in regional growth experiences are quite remarkable.[1] Growth took off in Western Europe and its offshoots in the early nineteenth century and never stopped again. But other regions took longer to participate in the growth of the world economy. Perhaps the most dramatic case is that of Asia, which basically did not grow until 1950 just to become then the fastest growing region in the world. Another extreme case is that of Africa, which still today is unable to enjoy growth rates that would be considered modest in other regions. Another salient feature of the growth process is therefore its uneven geographical distribution: in each period there are some regions that have been able to grow and prosper, while others have been left behind.

World economic growth has been accompanied by more than proportional growth in world trade. Figure 3 shows the evolution of world trade as a share of world production since 1870. The picture is quite clear: from 1870 to 1998 growth in world trade has quadrupled growth in world income. There also appears to be a strong positive correlation between growth in per capita income and growth in trade. Figure 4 plots the growth rates of these two variables against each other using pooled data from various regions and periods. The simple correlation between these variables is 0.64, and the regression results indicate that regions and periods with X percent higher than average trade growth tend to have per capita income growth which is $0.3 \cdot X$ higher than average. It almost goes without saying that this statistical association between income and trade does not imply causation in any direction. But it strongly suggests that these variables are somehow related, and that there might be substantial payoffs to working with theories that jointly determine them.[2]

[1] To get a sense of the magnitudes involved, remember that an annual growth rate of G leads per capita income to multiply itself by a factor $F \approx \exp\{G \cdot T\}$ in T years. For instance, in the last quarter of the twentieth century Asia has been able to increase its per capita income by a factor of 2.5, while Latin America has only managed to increase its per capita income by a factor of 1.2 and Africa has stagnated. Even a cursory look at the data shows that this disparity in growth performances constitutes the norm rather than the exception.

[2] For empirical work on the (causal) effect of trade on income levels and income growth see Sachs and Warner [1995], Frankel and Romer [1999], Ades and Glaeser [1999], Alesina, Spolaore, and Wacziarg [2000 and their chapter in this handbook], Rodriguez and Rodrik [2000], Alcalá and Ciccone [2003 and 2004], and Dollar and Kraay [2003].

Despite this apparent relationship between income and trade, a substantial part of growth theory is built on the assumption that countries live in autarky and that there is no trade among them.[3] This is obviously a dubious assumption. But is it also a "crucial" one? And if so, what alternative assumptions would be reasonably realistic? At an abstract level, these are the questions that I attempt to answer here. A recurring theme throughout this chapter is that the growth experiences of the different world regions are intimately linked and cannot be analyzed in isolation. We therefore need a global view of economic growth that looks at the different regions of the world as parts of a single whole. Formally, this means that we should develop and systematically study world equilibrium models. These models and their predictions constitute the specific focus of this chapter.[4]

Rather than providing an all-encompassing survey of the field, my goal in writing this chapter has been to develop a unified and yet tractable framework to discuss key insights of the fields of international trade and economic growth. In particular, I examine a sequence of world equilibrium models that share a common description of technology and preferences but differ on their assumptions about trade frictions. By comparing the predictions of these models against each other, it is possible to identify a variety of channels through which trade affects the evolution of world income and its geographical distribution. By comparing their predictions against the data, it is also possible to construct coherent explanations of income differences and long run trends in economic growth. When viewed as a group, these models show that much is known about the relationship between income and trade. Despite this, I still feel we are only exploring the tip of the iceberg. The research program sketched here is ambitious, fun and it could eventually lead to a much deeper understanding of the forces that drive modern capitalist economies.

[3] A brief examination of the different chapters of this handbook should quickly convince anyone doubting this statement.
[4] Without doubt, the seminal book by Grossman and Helpman [1991] is the single most influential contribution to the development and study of world equilibrium models of the growth process. It heavily influenced a whole generation of PhD students, like myself, that were searching for dissertation topics when the book first appeared. But there are, of course, many other important contributions. The bibliography at the end of this chapter is an (admittedly imperfect) attempt to list all published papers that use world equilibrium models to study the growth process. I apologize to the authors of any relevant paper that has been overlooked.

The rest of this chapter contains four sections. The first one describes growth in the integrated economy. This is an imaginary world where trade costs are negligible and geography does not matter. Section two introduces two trade frictions: the immobility of production factors and the absence of international financial markets. Section three adds a third trade friction: costs of transporting goods. The fourth and final section briefly concludes by taking stock what we have learned and pointing out potential avenues for further research.

1. The integrated economy

Imagine a world without borders, a world in which all goods and factors can be transported across different regions at negligible cost. Some industries spread their production process across many regions searching for the ideal environment for each specific phase of production. Other industries choose instead to concentrate production in a single region to exploit increasing returns to scale. Regardless of an industry's particular circumstances, its location choice maximizes productivity and is not affected by the local availability of production factors and/or final customers. If a region does not have the necessary production factors, these can be imported from abroad. If a region does not have enough customers, the goods produced can be exported abroad. In this world, global market forces arbitrage away regional differences in goods and factor prices and all the gains from trade are reaped. This imaginary world is the integrated economy, and is the subject of this section.

The integrated economy provides a natural benchmark for the study of economic growth in an interdependent world. Moreover, its simplicity and elegance encapsulates the essence of what growth theory is all about: deriving strong results using minimalist models. In the spirit of the so-called "new growth theory", I shall use a model that jointly determines the stock of capital and the level of technology.

Admittedly, the model is somewhat lopsided. On the one hand, it contains a fairly sophisticated formulation of technology that includes various popular models as special cases. On the other hand, it uses a brutal simplification of the standard overlapping-generations model as a description of preferences. Despite this, I do not apologize for the imbalance. A robust theme in growth theory is that the interesting part of the story is nearly always on the technology side, and rarely on the side of preferences.

This section develops the basic framework that I use throughout the chapter. Sub-section 1.1 describes the integrated economy, while sub-section 1.2 derives its main predictions for world growth. Sub-section 1.3 goes back to a period in which all the regions of the world lived in autarky, and compares the growth process of this world with the integrated economy. This is just the first of various attacks to the question of globalization and its effects on the world economy.

1.1 A workhorse model

Consider a world economy inhabited by two overlapping generations: young and old. The young work and, if productive, they earn a wage. The old retire and live off their savings. All generations have size one. There are many final goods used for consumption and investment, indexed by $i \in I$. When this does not lead to confusion, I shall use I to refer both to the set of final goods and also to the number of final goods. As we shall see later, the production of these final goods requires a continuum of intermediate inputs. There are two factors of production: labor and capital. For simplicity, I assume capital depreciates fully within one generation.[5] The world economy contains many regions. But geography has no economic consequences since goods and factors can be transported from one region to another at any time at negligible cost.

[5] The main role of this assumption is to ensure that investment is always strictly positive. This simplifies the presentation without substantially affecting the main results.

The citizens of this world differ in their preferences and access to education. S_t members of the generation born in date t are patient and maximize the expected utility of old age consumption, while the rest are impatient and maximize the expected utility of consumption when young. The utility function has consumption as its single argument, and it is homothetic, strictly concave and identical for all individuals. H_t members of the generation born in date t can access education and become productive, while the rest have no access to education and remain unproductive.[6] I refer to S_t and H_t as "savings" and "human capital", and I allow them to vary stochastically over time within the unit interval. Assuming that savings and human capital are uncorrelated within each generation, we obtain:

(1) $$K_{t+1} = S_t \cdot w_t \cdot H_t$$

(2) $$C_t = (1 - S_t) \cdot w_t \cdot H_t + r_t \cdot K_t$$

where K_t and C_t are the average or aggregate capital stock and consumption; and w_t and r_t are the wage and rental rate of capital. Equation (1) states that the capital stock equals the savings of the young, which consist of the wage of those that are patient and productive. The assumption that capital depreciates fully in one generation implies that the capital stock is equal to investment. Equation (2) says that consumption equals the wage of the impatient and productive young plus the return to the savings of the old.[7]

Consumption and investment can be thought of as composites or aggregates of the different final goods. A very convenient assumption is that both composites

---

[6] The assumption that labor productivity is either one or zero is extreme, but inessential. We could also think of H_t as the average labor productivity of the world economy. The assumption that human capital is not industry specific is widespread, but not entirely innocent. See Basu and Weil [1998] and Brezis, Krugman and Tsiddon [1993] for interesting implications of relaxing this assumption.

[7] This representation of savings and consumption is nothing but a stripped-down version of Modigliani's life-cycle theory of savings. It abstracts from other motives for savings such as leaving bequests. These could be easily re-introduced in the theory through suitable and well-known modifications of the preferences of individuals. I shall not do this to keep the analysis as simple as possible. I conjecture that the bulk of the basic intuitions and results presented here would not be meaningfully affected by these extensions.

take the same Cobb-Douglas form with spending shares that vary across industries, i.e. σ_i with $\sum_{i \in I} \sigma_i = 1$. Since there is a common ideal price index for consumption and investment, it makes sense to use it as the numeraire and this implies that aggregate spending is given by $E_t \equiv C_t + K_{t+1}$. To sum up, we have that:

(3) $$E_{it} = \sigma_i \cdot E_t \qquad \text{for all } i \in I$$

(4) $$1 = \prod_{i \in I} \left(\frac{P_{it}}{\sigma_i} \right)^{\sigma_i}$$

where E_{it} and P_{it} are the total spending on and the price of the final good of industry i. Equation (3) states that spending shares are constant, while Equation (4) sets the common price of consumption and investment equal to one.

Production of final goods uses labor, capital and a continuum of different varieties of intermediate inputs, indexed by $m \in [0, M_{it}]$ for all $i \in I$. As usual, I interpret the measure of input varieties, M_{it} for all $i \in I$, as the degree of specialization or the technology of the industry. This measure will be determined endogenously as part of the equilibrium. The technology of industry i can be summarized by these total cost functions:

(5) $$B_{it} = \left[\frac{1}{Z_{it}} \cdot \left(\frac{w_t}{1-\alpha_i} \right)^{1-\alpha_i} \cdot \left(\frac{r_t}{\alpha_i} \right)^{\alpha_i} \right]^{1-\beta_i} \cdot \left[\int_0^{M_{it}} p_{it}(m)^{1-\varepsilon_i} \cdot dm \right]^{\frac{\beta_i}{1-\varepsilon_i}} \cdot Q_{it} \qquad \text{for all } i \in I$$

(6) $$b_{it}(m) = \frac{1+q_{it}(m)}{Z_{it}} \cdot \left(\frac{w_t}{1-\alpha_i} \right)^{1-\alpha_i} \cdot \left(\frac{r_t}{\alpha_i} \right)^{\alpha_i} \qquad \text{for all } m \in [0, M_{it}] \text{ and } i \in I$$

where $0 \leq \beta_i \leq 1$, $\varepsilon_i > 1$ and $0 \leq \alpha_i \leq 1$; Q_{it} is total production of final good i; $q_{it}(m)$ and $p_{it}(m)$ are the quantity and price of the m^{th} input variety of industry i; and the variables Z_{it}

are meant to capture the influence on industry productivity of geography, institutions and other factors that are exogenous to the analysis.[8] I loosely refer to the Z_{it}s as "industry productivities" and assume they vary stochastically over time within a support that is strictly positive and bounded above. Equation (5) states that the technology to produce the final good of industry i is a Cobb-Douglas function on human and physical capital, and intermediate inputs. The latter are aggregated with a standard CES function. Equation (6) states that the production of intermediates is also a Cobb-Douglas function on human and physical capital, and that there are fixed and variable costs.[9] I interpret the fixed costs as including both the costs of building a specialized production plant and the costs of inventing or developing a new variety of intermediate. An important simplifying assumption is that input varieties become obsolete in one generation and, as a result, all generations must incur these fixed costs.[10]

Since there are constant returns in the production of final goods, it is natural to assume that final good producers operate under perfect competition. Therefore, prices and intermediate input demands are given as follows:

(7) $$P_{it} = \frac{\partial B_{it}}{\partial Q_{it}} \qquad \text{for all } i \in I$$

(8) $$q_{it}(m) = \frac{\partial B_{it}}{\partial p_{it}(m)} \qquad \text{for all } m \in [0, M_{it}] \text{ and } i \in I$$

Equation (7) states that price equals marginal cost, while Equation (8) uses Shephard's lemma to describe the demand for intermediate inputs. Equations (5) and (8) imply that an increase in the price of a given input variety lowers its market share. But Equation (3) shows that the lost market share goes entirely to other input

[8] Although popular, this is a quite simplistic view of the effects of geography and institutions. See Levchenko [2004] for an interesting discussion of alternative ways of modeling the effects of institutions.
[9] As usual, the fixed cost is paid if and only if there is strictly positive production.
[10] This assumption is crucial for tractability, since it eliminates a potentially large set of state variables, i.e. M_{it} for all $i \in I$.

varieties of the same industry and does not affect the industry's overall market share.

Since the production of intermediate inputs exhibits increasing returns that are internal to the firm, input producers cannot operate under perfect competition. I assume instead they operate under monopolistic competition with free entry. This has the following implications:

(9) $$p_{it}(m) = \frac{e_{it}(m)}{e_{it}(m)-1} \cdot \frac{\partial b_{it}(m)}{\partial q_{it}(m)} \qquad \text{for all } m \in [0, M_{it}] \text{ and } i \in I$$

(10) $$p_{it}(m) \cdot q_{it}(m) = b_{it}(m) \qquad \text{for all } m \in [0, M_{it}] \text{ and } i \in I$$

where $e_{it}(m)$ is the price-elasticity of input demand: $e_{it}(m) = -\frac{p_{it}(m)}{q_{it}(m)} \cdot \frac{\partial q_{it}(m)}{\partial p_{it}(m)}$ with the derivative in this definition being applied to Equation (8). Equation (9) states that monopolistic firms charge a markup over marginal cost that is decreasing on the demand elasticity faced by the firm. As usual, the CES formulation implies that this demand elasticity is equal to the elasticity of substitution among inputs, i.e. $e_{it}(m)=\varepsilon_i$. Equation (10) states that profits must be zero and this is, of course, a direct implication of assuming free entry.

Finally, we must impose appropriate resource constraints or market-clearing conditions:

(11) $$P_{it} \cdot Q_{it} = E_{it} \qquad \text{for all } i \in I$$

(12) $$H_t = \sum_{i \in I} H_{it} \qquad \text{with } H_{it} = \frac{\partial B_{it}}{\partial w_t} + \int_0^{M_{it}} \frac{\partial b_{it}(m)}{\partial w_t} \cdot dm$$

(13) $$K_t = \sum_{i \in I} K_{it} \qquad \text{with } K_{it} = \frac{\partial B_{it}}{\partial r_t} + \int_0^{M_{it}} \frac{\partial b_{it}(m)}{\partial r_t} \cdot dm$$

where H_{it} and K_{it} are the labor and capital demanded by industry i. Since the integrated economy is a closed economy, Equation (11) forces the aggregate supply of each good to match its demand, while Equations (12)-(13) state that the aggregate supply of labor and capital must equal their demands. The latter are the sum of their industry demands, and these are calculated using Shephard's lemma.

This completes the description of the model. For any admissible initial capital stock and sequences for S_t, H_t, and Z_{it}, an equilibrium of the integrated economy consists of sequences of prices and quantities such that Equations (1)-(13) hold in all dates and states of nature. The assumptions made ensure that this equilibrium always exists and is unique. I shall show this by construction in the next section.

The reader might be wondering why I have not formally introduced financial markets. I have allowed individuals to construct their own capital and use it as a vehicle to carry on their savings into retirement (a world of family-owned firms?). But I have not allowed them to trade securities in organized financial markets. The reason is simply to save notation. The assumptions made ensure that asset trade does not matter in this world economy. [11] To see this, assume there exist sophisticated financial markets where all individuals can trade a wide array of state-contingent securities. Naturally, the old would not be able to trade these securities since they will not be back to settle claims one period later. But the young would not trade with each other either. Impatient young would not be willing to trade securities since they do not have income in their old age and are happy to consume all their income during their youth. Patient young are the only ones willing and able to trade these securities. But they all have identical preferences and face the same distribution of returns to capital, and therefore they find no motive to trade with each other. Thus, we can safely assume the integrated economy contains sophisticated

[11] This statement is not entirely correct. It applies to assets whose price reflects only fundamentals, but without additional assumptions it does not apply to securities whose price contains a bubble. I shall disregard the possibility of asset bubbles in this chapter, although this is far from an innocuous assumption. See Ventura [2002] for an example where asset bubbles have an important effect on the growth of the world economy and its geographical distribution.

financial markets that allow individuals to enter contracts that specify exchanges of various quantities of the different goods to be delivered at various dates and/or states of nature. It just happens that these financial markets do not make any difference for consumption and welfare.

1.2 Diminishing returns, market size and economic growth

To study the forces that determine economic growth in the integrated economy, it is useful to start with a familiar expression:

$$\frac{K_{t+1}}{K_t} = s_t \cdot \frac{Q_t}{K_t} \qquad (14)$$

where Q_t is the integrated economy's output or production, i.e. $Q_t \equiv \sum_{i \in I} P_{it} \cdot Q_{it}$; and s_t is the economy's (gross) savings rate, i.e. $s_t \equiv \frac{K_{t+1}}{Q_t}$. Equation (14) states that the (gross) growth rate of the capital stock is equal to the savings rate times the output-capital ratio or average product of capital. If this product stays above one asymptotically, the world economy exhibits sustained or long run growth. Otherwise, economic growth eventually ceases and the world economy stagnates. We shall study then the determinants of savings and the average product of capital.

To compute the savings rate, remember that industry i receives a share σ_i of aggregate spending of which a fraction $1-\alpha_i$ goes to labor. Adding across industries, it follows that aggregate labor income is $w_t \cdot H_t = (1-\alpha) \cdot Q_t$, where α is the aggregate or average share of capital, i.e. $\alpha \equiv \sum_{i \in I} \sigma_i \cdot \alpha_i$. Since only the patient young save, the savings rate consists of the fraction of labor income in the hands of patient consumers:

(15) $s_t = (1-\alpha)\cdot S_t$

Since the savings rate is bounded above, sustained economic growth requires that the average product of capital remain above one as the economy grows. But what determines the aggregate output-capital ratio? I shall answer this question in a few steps, so as to develop intuition.

The first step consists of finding the output-capital ratio of a given industry as a function of its technology and factor proportions:[12]

(16) $$\frac{Q_{it}}{K_{it}} = \left(\frac{\varepsilon_i}{\varepsilon_i - 1}\right)^{-\beta_i} \cdot M_{it}^{\frac{\beta_i}{\varepsilon_i - 1}} \cdot Z_{it} \cdot \left(\frac{K_{it}}{H_{it}}\right)^{\alpha_i - 1} \qquad \text{for all } i \in I$$

Equation (16) shows the effects of changes in factor proportions on the industry's output-capital ratio, *holding constant technology*. Since there are diminishing returns to physical and human capital in production, we find the standard result that increases in the physical to human capital ratio reduce the output-capital ratio. But technology is endogenously determined in this model, and it depends on the size of the industry:[13]

(17) $$M_{it} = \frac{\beta_i}{\varepsilon_i} \cdot Z_{it} \cdot H_{it}^{1-\alpha_i} \cdot K_{it}^{\alpha_i} \qquad \text{for all } i \in I$$

Equation (17) shows that increases in factor usage or industry size raise the incentives to specialize and therefore improve technology. The larger is the size of the market, the easier it is to recoup the fixed costs of producing a new input variety

[12] From Equations (7) and (11) find that $P_{it}\cdot Q_{it}=B_{it}$, and use this to eliminate B_{it} from Equation (5). Then, solve Equation (9) with Equation (6), substitute into Equation (5) and eliminate factor prices by noting that the industry factor shares, i.e. $w_t\cdot H_{it}/P_{it}\cdot Q_{it}$ and $r_t\cdot K_{it}/P_{it}\cdot Q_{it}$ are given by $1-\alpha_i$ and α_i, respectively.

[13] Symmetry of intermediates and perfect competition in the final goods industry implies that $M_{it}\cdot p_{it}\cdot q_{it}=\beta_i\cdot P_{it}\cdot Q_{it}$; where p_{it} and q_{it} are the common price and quantity of all varieties of intermediates of industry i. Then, use Equations (6), (9) and (10) to eliminate p_{it} and q_{it} from this expression. Finally, eliminate factor prices once again by noting that the industry factor shares are $1-\alpha_i$ and α_i.

and therefore the higher is the number of input varieties that can be sustained in equilibrium. We can now put these two pieces together and write the output-capital ratio as follows:

$$(18)\qquad \frac{Q_{it}}{K_{it}} = A_{it} \cdot H_{it}^{\mu_i \cdot (1-\alpha_i)} \cdot K_{it}^{\mu_i \cdot \alpha_i - 1} \qquad \text{for all } i \in I$$

where μ_i is a measure of the importance of market size effects, i.e. $\mu_i = 1 + \frac{\beta_i}{\varepsilon_i - 1}$;

and A_{it} is a measure of industry productivity, i.e. $A_{it} = \left(\frac{\varepsilon_i}{\varepsilon_i - 1}\right)^{-\beta_i} \cdot \left(\frac{\beta_i}{\varepsilon_i}\right)^{\frac{\beta_i}{\varepsilon_i - 1}} \cdot Z_{it}^{\mu_i}$. I shall refer to both Z_{it} and A_{it} as "industry productivities" when this is not a cause for confusion. Equation (18) summarizes the aggregate industry technology and shows direct and indirect effects of factor usage on the industry's output-capital ratio. Increases in human capital raise the output-capital ratio, as the direct positive effect of making physical capital scarce is reinforced by the indirect effect of increasing input variety. Increases in physical capital have an ambiguous effect on the output-capital ratio, as the direct negative effect of making physical capital abundant and the positive indirect effect of increasing input variety work in opposite directions. If diminishing returns are strong and market size effects are weak ($\mu_i \cdot \alpha_i < 1$) increases in physical capital reduce the industry's output-capital ratio. If instead diminishing returns are weak and market size effects are strong ($\mu_i \cdot \alpha_i \geq 1$) increases in physical capital raise the industry's output-capital ratio.

The next step is to aggregate these effects across industries. To do this, note first that factor allocations and aggregate output are determined as follows:[14]

[14] Equations (19) and (20) are direct implications of the constant factor and spending shares. One way to think about Equation (21) is as the definition of the Cobb-Douglas aggregate that defines consumption and investment and therefore underlies Equations (3) and (4). Another way of thinking about Equation (21) is as an implication of Equations (3), (4) and (11).

(19) $$H_{it} = \sigma_i \cdot \frac{1-\alpha_i}{1-\alpha} \cdot H_t \quad \text{for all } i \in I$$

(20) $$K_{it} = \sigma_i \cdot \frac{\alpha_i}{\alpha} \cdot K_t \quad \text{for all } i \in I$$

(21) $$Q_t = \prod_{i \in I} Q_{it}^{\sigma_i}$$

Equations (19) and (20) show that the equilibrium allocations of human and physical capital to industry i depend on the corresponding factor share and the size of the industry. Equation (21) says that output is a Cobb-Douglas aggregate of industry outputs. This is, of course, the production function associated with the cost function in Equation (4). It is now immediate to substitute Equations (18), (19) and (20) into Equation (21) to find the aggregate output-capital ratio of the world economy:

(22) $$\frac{Q_t}{K_t} = A_t \cdot H_t^{\mu \cdot (1-\alpha) - \upsilon} \cdot K_t^{\mu \cdot \alpha + \upsilon - 1}$$

where μ is the average value of μ_i, i.e. $\mu \equiv \sum_{i \in I} \sigma_i \cdot \mu_i$; υ is the covariance between μ_i and α_i, i.e. $\upsilon \equiv \sum_{i \in I} \sigma_i \cdot (\mu_i - \mu) \cdot (\alpha_i - \alpha)$; and A_t is an aggregate measure of productivity, i.e. $A_t \equiv \prod_{i \in I} \left[\sigma_i^{\mu_i} \cdot \left(\frac{1-\alpha_i}{1-\alpha} \right)^{\mu_i \cdot (1-\alpha_i)} \cdot \left(\frac{\alpha_i}{\alpha} \right)^{\mu_i \cdot \alpha_i} \cdot A_{it} \right]^{\sigma_i}$. Equation (22) is the aggregate production function and will play an important role in what follows. It shows that the industry intuitions on the effects of changes in factor usage carry on to the aggregate effects of changes in factor supplies. While increases in human capital unambiguously raise the output-capital ratio, increases in physical capital have ambiguous effects.[15] If the "representative" industry has strong diminishing returns and weak market-size effects ($\mu \cdot \alpha + \upsilon < 1$) physical capital accumulation

[15] Note that $\mu \cdot (1-\alpha) - \upsilon \geq 0$.

reduces the aggregate output-capital ratio. If instead the "representative" industry has weak diminishing returns and strong market-size effects ($\mu \cdot \alpha + \upsilon \geq 1$) physical capital accumulation raises the output-capital ratio.

We are ready now to characterize the process of economic growth in the integrated economy. Substituting Equation (22) into Equation (14), we obtain the following law of motion for the capital stock:

$$(23) \qquad K_{t+1} = s_t \cdot A_t \cdot H_t^{\mu \cdot (1-\alpha)-\upsilon} \cdot K_t^{\mu \cdot \alpha + \upsilon}$$

Equation (23) shows that the integrated economy behaves as if it were a Solow model with a Cobb-Douglas production function that exhibits increasing returns to scale, i.e. the sum of the share coefficients is $\mu \geq 1$. Figures 5 and 6 illustrate the dynamics of the stock of physical capital with the help of two simple examples. The first example is the "deterministic" world where savings, human capital and productivity are constant over time, i.e. $\{s_t,H_t,A_t\}=\{s,H,A\}$ for all t. The second example is the "stochastic" world where savings, human capital and productivity fluctuate between a "bad" state with $\{s_t,H_t,A_t\}=\{s_B,H_B,A_B\}$ and a "good" state with $\{s_t,H_t,A_t\}=\{s_G,H_G,A_G\}$; with $s_G \cdot A_G \cdot H_G^{\mu \cdot (1-\alpha)-\upsilon} > s_B \cdot A_B \cdot H_B^{\mu \cdot (1-\alpha)-\upsilon}$. The central point of these examples is to show that economic growth solves a tension between diminishing returns and market size effects.

Figure 5 shows the case in which diminishing returns are strong and market-size effects are weak, i.e. $\mu \cdot \alpha + \upsilon < 1$. The top panel depicts the evolution of the "deterministic" world. There is a unique steady state and the stock of physical capital converges monotonically towards it from any initial position. The steady state is stable because increases (decreases) in the stock of physical capital lower (raise) the output-capital ratio and lead to a lower (higher) growth rate. The bottom panel shows that the "stochastic" world exhibits similar dynamics, with the stock of physical capital monotonically converging to a steady state interval, rather than a steady state value. Once the stock of physical capital is trapped within this interval, its growth rate fluctuates between positive and negative values and averages zero in the long run. These examples illustrate why sustained growth is not possible if diminishing returns are strong and market size effects are weak.

Figure 6 shows the case in which diminishing returns are weak and market-size effects are strong, i.e. $\mu \cdot \alpha + \upsilon \geq 1$. The top panel shows the "deterministic" world again. There is unique steady state that is unstable. If the stock of physical capital starts above the steady state, it grows without bound at an accelerating rate. If it starts below, the stock of physical capital contracts over time also at an accelerating rate. The steady state is now unstable because increases (decreases) in the stock of physical capital raise (lower) the output-capital ratio and lead to a higher (lower) growth rate. The bottom panel shows that the "stochastic" world also exhibits similar dynamics. One difference however is that there is no steady state. Instead, there is a threshold interval. If the stock of physical capital is above (below) this interval, it grows (contracts) at an accelerating rate. If the stock of physical capital starts within the threshold interval, it fluctuates within it until it eventually exits. This happens with probability one, and only luck determines when this exit occurs and whether the world economy exits above and enters an expansionary path or, alternatively, it exits below and enters a contractionary path. Therefore, sustained growth is possible (but not necessary) if diminishing returns are weak and market size effects are strong.

This model suggests a simple account of the history of the world economy since the 1500s. It is based on the "stochastic" world of Figure 6 and it goes as follows: for centuries, the size of the world economy was too small to generate sustained growth. Located within the threshold interval, the world economy was subject to periodic expansions and contractions with virtually zero average growth. This is consistent with Maddison's calculation that the world economy grew only about eighteen percent from 1500 to 1820. But this was an unstable situation in the very long run. The Industrial Revolution marks the moment in which, after a series of favorable shocks, the world economy reached enough size to exit the threshold interval and started traveling on the path of accelerating growth reported in Figure 1. As a result of this successful exit, the world economy grew more than seven hundred and fifty percent from 1820 to 1998.

Although suggestive, this account is far too sketchy and incomplete to be taken seriously. Moreover, I find highly improbable that the last five hundred years of the world economy can be understood in terms of a model that postulates negligible costs of transporting goods and factors and constant world population. Surely the demographic revolution and the process of globalization have both played central roles in shaping the growth process during this period. This chapter is not the place for a discussion of the growth effects of the demographic revolution.[16] But it is definitely the place to study the growth effects of globalization, and we turn to this topic next.

1.3 The effects of economic integration

Assume the world economy initially consisted of many regions or locations separated by geographical obstacles that made the costs of transporting goods and factors among them prohibitive. As a result, these regions were forced to live in autarky. I index these regions by $c \in C$, and let them differ on their savings, human capital, industry productivities and initial capital stock, i.e. on $S_{c,t}$, $H_{c,t}$, $Z_{c,it}$ and $K_{c,0.}$ When this does not lead to confusion, I shall use C to refer to both the set of regions and also to the number of regions. Throughout, I denote world aggregates by omitting the region sub-index. Typically, world aggregates refer to the sum of all corresponding regional variables. For instance, world aggregate savings, human and

---

[16] In this model, a sustained increase in population would generate sustained growth even if $\alpha \cdot \mu + \upsilon < 1$. The reason is that, holding constant both factor endowments and productivity, population growth increases the size of the market and this raises income. I have ruled out this possibility by simply assuming that the world population is constant. Given the purpose of this chapter, I think this is not a "crucial" assumption. But it might be so in other contexts. See Jones' chapter in this volume for a thorough and clear discussion of scale effects in growth models.

physical capital are $S_t = \sum_{c \in C} S_{c,t}$, $H_t = \sum_{c \in C} H_{c,t}$ and $K_t = \sum_{c \in C} K_{c,t}$. But there will be some exceptions. For instance, the relationship between $Z_{c,it}$ and the corresponding world aggregate Z_{it} is a bit more intricate and will be explained shortly.

Although it is not really necessary to take a stand on the geographical distribution of population, I assume throughout that it is equally distributed across regions. This simplifies somewhat the presentation since absolute and per capita regional comparisons coincide. For instance, if $S_{c,t}>S_{c',t}$ then c also has higher savings per person than c'. Note also that, as the number of regions becomes arbitrarily large, the size of each of them becomes arbitrarily small and the effects of shocks to their characteristics on world aggregates become arbitrarily small. This limiting case is usually referred to as the small economy assumption.

The model of globalization considered here is embarrassingly simple: at date t=0, all the geographical obstacles to trade suddenly disappear forever and the costs of transporting goods and factors fall from prohibitive to negligible. What are the effects of such a dramatic reduction in transport costs on world economic growth and its geographical distribution? To answer this question, we must characterize the growth process in the autarkic world economy and in the integrated world economy and compare them. Although this way of modeling globalization and its effects is almost a caricature, it turns out to be quite useful to develop intuitions that survive as we move to more sophisticated and realistic models.

In the world of autarky, each region constituted a smaller version of the integrated economy. Therefore, the world economy at t<0 can be described by:[17]

(24) $$Y_{c,t} = A_{c,t} \cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_{c,t}^{\mu\cdot\alpha+\upsilon} \qquad \text{for all } c \in C$$

(25) $$K_{c,t+1} = s_{c,t} \cdot A_{c,t} \cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_{c,t}^{\mu\cdot\alpha+\upsilon} \qquad \text{for all } c \in C$$

[17] Equation (25) is an analogue to Equation (23), while Equation (24) follows from the region counterparts to Equation (22) and the fact that $Y_{c,t}=Q_{c,t}=C_{c,t}+K_{c,t}$ in autarky.

where $Y_{c,t}$ is the income of the region and, in autarky, it coincides with its production and spending, i.e. $Y_{c,t}=Q_{c,t}=E_{c,t}$; and $A_{c,t}$ is the corresponding measure of regional productivity, i.e. $A_{c,t} \equiv \prod_{i \in I} \left[\sigma_i^{\mu_i} \cdot \left(\frac{1-\alpha_i}{1-\alpha} \right)^{\mu_i \cdot (1-\alpha_i)} \cdot \left(\frac{\alpha_i}{\alpha} \right)^{\mu_i \cdot \alpha_i} \cdot A_{c,it} \right]^{\sigma_i}$ with

$A_{c,it} = \left(\frac{\varepsilon_i}{\varepsilon_i - 1} \right)^{-\beta_i} \cdot \left(\frac{\beta_i}{\varepsilon_i} \right)^{\frac{\beta_i}{\varepsilon_i - 1}} \cdot Z_{c,it}^{\mu_i}$. Equations (24) and (25) have been discussed at length already and need no further comment.

In the integrated economy it is not possible in general to determine the production or spending located in a given region. Since goods and factors can move at negligible cost, any geographical distribution of production and factors that ensures all production takes place in the regions with the highest industry productivity is a possible equilibrium. Despite this indeterminacy, prices and aggregate quantities are uniquely determined as shown in section 1.2. This means that it is possible to track the stock of physical capital owned by the original inhabitants of region c and their descendants as well as their income:[18]

$$(26) \qquad Y_{c,t} = \left[(1-\alpha) \cdot \frac{H_{c,t}}{H_t} + \alpha \cdot \frac{K_{c,t}}{K_t} \right] \cdot A_t \cdot H_t^{\mu \cdot (1-\alpha) - \upsilon} \cdot K_t^{\alpha \cdot \mu + \upsilon} \qquad \text{for all } c \in C$$

$$(27) \qquad K_{c,t+1} = \frac{S_{c,t} \cdot H_{c,t}}{S_t \cdot H_t} \cdot s_t \cdot A_t \cdot H_t^{\mu \cdot (1-\alpha) - \upsilon} \cdot K_t^{\alpha \cdot \mu + \upsilon} \qquad \text{for all } c \in C$$

for all $c \in C$ and $t \geq 0$; and A_t is a measure of world productivity. Remember that we have now specified a set of industry productivities for each region, $Z_{c,it}$. But we only specified one set of industry productivities for the integrated economy in section 1.1.

[18] Equation (26) follows from adding the income from human and physical capital of the inhabitants of the region, and noting that aggregate or world shares of human and physical capital are constant and equal to 1-α and α, respectively. Equation (27) follows from Equations (1) and (23), and the observation that wages are the same for all productive workers of the world. Without loss of generality, I keep assuming that there is no trade in securities.

The reason was that industries never locate in a region that offers less than the highest possible productivity. As a result, in the integrated world economy the only industry productivities that matter are the highest ones, i.e. $Z_{it} = \max_{c \in C}\{Z_{c,it}\}$. This implies that $A_t \geq A_{c,t}$ for all $c \in C$, and we can interpret aggregate productivity not as average productivity, but instead as the highest possible productivity or the world productivity frontier. With this in mind, Equation (27) traces the holdings of capital of the original inhabitants of region c and their descendants, while Equation (26) describes their income.

We are ready now to examine the growth effects of economic integration. Consider first the static or impact effects on the incomes of regions. A bit of straightforward algebra shows that:[19]

$$(28) \qquad \ln\left(\frac{Y^{I}_{c,0}}{Y^{A}_{c,0}}\right) = \underbrace{\ln\left(\frac{A_0}{A_{c,0}}\right)}_{\substack{higher \\ productivity}} + \underbrace{\ln\left(\frac{(1-\alpha)\cdot\frac{H_{c,0}}{H_0} + \alpha\cdot\frac{K_{c,0}}{K_0}}{\left(\frac{H_{c,0}}{H_0}\right)^{1-\alpha}\cdot\left(\frac{K_{c,0}}{K_0}\right)^{\alpha}}\right)}_{\substack{improved\ \ factor \\ allocation}} + \underbrace{\ln\left(\frac{H_0^{1-\alpha}\cdot K_0^{\alpha}}{H_{c,0}^{1-\alpha}\cdot K_{c,0}^{\alpha}}\right)^{\mu-1}}_{\substack{increased \\ market\ \ size}} \geq 0$$

where $Y^{I}_{c,0}$ is the actual income of the inhabitants of region c at date t=0, and $Y^{A}_{c,0}$ is the income they would have had at date t=0 if globalization had not taken place. Since each of the terms in Equation (28) is non-negative, the first result we obtain is that the overall impact or static gains from economic integration are non-negative as well.

These gains can be decomposed into three sources corresponding to each of the terms of Equation (28). The first one shows the growth of income that results

[19] To derive this expression I have assumed a zero cross-industry correlation between α_i and μ_i, i.e. υ=0. This parameter restriction is useful because it allows us to unambiguously disentangle the "increased-market-size" and "improved-factor-allocation" effects.

from moving industries from low to high productivity locations. This term would vanish if region c had the highest productivity in all industries. The second term shows the growth of income that results from relocating factors away from those regions and/or industries in which they were abundant in autarky into those in which they were scarce. This term would vanish if region c had world average factor proportions. The third term shows the growth in income that is due to an increase in market size that allows industries to support a higher degree of specialization. This term would vanish if the size of region c were arbitrarily large with respect to the rest of the world. An implication of Equation (28) is that the static gains from economic integration are greater for regions with low productivity, extreme factor proportions and modest amounts of physical and human capital.

If coupled with an appropriate transfer scheme, globalization leads to a Pareto improvement in the world economy. Equation (28) shows that, with the same production factors, the integrated economy generates more output than the world of autarky. It is therefore possible to implement a transfer scheme that keeps constant the income of all current and future young and gives more income to all current and future old. Under this transfer scheme, investment and the stock of physical capital would be unaffected by economic integration. But the production and consumption of all generations born at date t=0 or later would increase. Of course, there exist many alternative transfer schemes that ensure that globalization benefits all. Moreover, since each region gains from trade there exist Pareto-improving transfer schemes that can be implemented without the need for inter-regional transfers. That is, ensuring that globalization generates a Pareto improvement does not require compensation from one region to another.

How “large” the transfer scheme must be to ensure that economic integration leads to a Pareto improvement? The answer is “not much” if most of the gains from economic integration come from higher productivity and increased market size. The reason is that in this case all factors share in the gains from integration. The required transfer scheme could be “substantial” if the gains from integration come mostly from improved factor allocation. This is because within each region the owners of the abundant factor obtain more than proportional gains from integration while the owners of the region's scarce factor might have losses. In this case, implementing a Pareto improvement requires a transfer from the former to the latter.

Without a transfer scheme, it is relatively straightforward to trace the dynamic effects of economic integration. Assume for simplicity that the world contains many symmetric regions so that before integration all of them had the same law of motion. The top panel of Figure 7 shows the effects of economic integration in the "deterministic" world when diminishing returns are strong and market size effects are weak. Economic integration raises the steady state stock of physical capital and sets up a period of high growth that eventually ends. It is straightforward to see that the effects would be similar in the "stochastic" world, with economic integration permanently raising the steady state interval. Using the jargon of growth theory, if $\mu \cdot \alpha + \upsilon < 1$ economic integration has level effects on income. The bottom panel of Figure 7 shows the opposite case in which diminishing returns are weak and market size effects are strong. In this case, economic integration shifts down the steady state value, increasing the growth rate permanently. Once again, it is straightforward to see that the effects would be similar in the "stochastic" world, with trade shifting the threshold interval to the left. Using again the jargon of growth theory, if $\mu \cdot \alpha + \upsilon \geq 1$ integration has growth effects on income.

It is tempting now to revisit our earlier account of the history of the world economy since the 1500s, and propose an alternative version which is also based on the "stochastic" world with $\mu \cdot \alpha + \upsilon > 1$. It goes as follows: for centuries, the world economy consisted of a collection of autarkic regions that were too small to sustain economic growth. Located within the threshold interval, these regions were subject to periodic expansions and contractions with virtually zero average growth. Once again, this is consistent with Maddison's calculation that the world economy grew only about eighteen percent from 1500 to 1820. The Industrial Revolution occurs when a series of reductions in trade costs between some British regions raised their combined size above the threshold interval and set them on the path of accelerating growth. As time went on, more and more regions joined the initial core and the Industrial Revolution spread throughout Britain and moved into France, Germany and beyond. It is therefore a reduction of trade costs and the progressive extension of markets that made possible sustained growth and allowed the world economy to grow more than seven hundred and fifty percent from 1820 to 1998. This might also explain why this growth in world income was accompanied by an even higher growth in world trade.[20]

This view of the development process is also broadly consistent with the general observations about inequality between center and periphery discussed in the introduction. Regions that join the integrated economy (the "center") become rich and take off into steady growth. Regions that do not join the integrated economy (the "periphery") are left behind, technologically backward and capital poor. As more and more regions enter the integrated economy, those that are left behind become relatively poorer and world inequality increases. Eventually all regions will enter the integrated economy and world inequality will decline. Therefore, this model generates an inverted U-shape or Kuznets curve, with world inequality rising in the first stages of world development and declining later. Pritchett [1997], Bourguignon and Morrison [2002] and others have shown that world inequality has increased from 1820 to now. It remains to be seen if this inequality will decline in the future.

This stylized model also illustrates some of the conflicts that globalization might create. It follows from Equation (28) that the gains from trade are large for regions whose factor proportions are far from the world average. Ceteris paribus, this means that regions in the center would like that new entrants into the integrated economy to move the world average factor proportions away from them. In fact, unless productivity and market size effects are substantial, the entry of a large region creates losses to other regions with similar factor proportions. This implies, for instance, that the Chinese process of economic integration should be seen with

[20] The word "might" reflects the earlier observation that regional production and therefore trade is indeterminate in the integrated economy.

some concern in countries with similar factor proportions such as Mexico and Indonesia, but with hope in the European Union or the United States.

This view of globalization and growth leads to a powerful prescription for economic development: open up and integrate into the world economy. I believe this is a fundamentally sound policy prescription, and history is largely consistent with it. But there are a number of important qualifications that this stylized model cannot capture. Integrating into the world economy is not an "all-or-nothing" type of affair in which regions move overnight from autarky to complete integration. The process of economic integration is slow and full of treacherous steps. Obtaining general prescriptions for development in a world of imperfect integration has proved to be a much more challenging task. I shall come back to this important point later, but we must first introduce trade frictions into the story.

2. Specialization, trade and diminishing returns

Let us revise our model of globalization. As in section 1.3, assume that at date t=0 the costs of transporting goods across regions suddenly fall from prohibitive to negligible. Unlike section 1.3, assume now that the costs of transporting factors across regions remain prohibitive after date t=0. An implication of this setup is that globalization equalizes goods prices across regions, but it does not necessarily equalize factor prices. This particular view of globalization has a longstanding tradition in trade theory and the goal of this section is to analyze it.

Assuming that human capital is immobile internationally is somewhat dubious, as there are some well-known examples of large contingents of people working overseas. But most of the results discussed here would go through with only minor changes under the weaker and reasonably realistic assumption that international flows of people are quantity constrained, although not necessarily at zero.[21] Assuming that physical capital is immobile is appropriate for buildings and structures and, probably, not too unreasonable for the most important types of machinery and equipment. Moreover, assuming that existing physical capital cannot be transported does not preclude physical capital to effectively “move” across regions over time, as it declines in some regions through depreciation and increases in others through investment.[22]

If physical capital is immobile, pieces of capital located in different regions might offer different return distributions. This opens up a role for financial markets. Although the old and the impatient young still have no incentive to trade securities, the patient young now have a motive. Those that are located in regions where physical capital offers an attractive distribution of returns want to sell securities and use the proceeds to finance additional purchases of domestic physical capital. Those patient young that are located in regions where physical capital offers an unattractive distribution of returns want to buy securities and reduce their holdings of domestic physical capital. And, regardless of their location, the patient young want to buy and sell securities in order to share regional risks. Thus, the immobility of physical capital creates a potentially important role for international financial markets: the geographical reallocation of investments and production risks.

Despite this, I will not let international financial markets play this role. This failure of financial markets could be due to technological motives or informational problems of various sorts. But I prefer instead to think of it as being caused by lack of incentives to enforce international contracts. In the integrated economy,

[21] Of course, this becomes a weak or empty excuse if quantity constraints respond to economic incentives in a systematic way. See Lundborg and Segerstrom [2002] and Ortega [2004] for models in which this happens.

[22] Remember that we have assumed that physical capital depreciates in one generation. Therefore, assuming physical capital is immobile only means that it is not possible at date t to move around the stock of physical capital created and deployed at date t-1, and that is being used for production at date t. But it is certainly possible to choose where to deploy the new stock of physical capital created at date t that will be used for production at date t+1. The effects of physical capital immobility would be more severe quantitatively with a slower rate of depreciation. Note also that immobility matters only because physical capital is irreversible or putty-clay. In fact, it would be logically inconsistent to assume that physical capital is immobile if it could be converted back into mobile goods.

individuals could enter into contracts that specify exchanges of various quantities of the different goods to be delivered at various dates and/or states of nature. It is standard convention to refer to the signing of contracts that involve only contemporaneous deliveries as “goods” trade, while the signing of contracts that involve future (and perhaps state contingent) deliveries is usually referred to as “asset” trade. Both types of trade require sufficiently low costs of transporting goods. But asset trade also requires that the signing parties credibly commit to fulfill their future contractual obligations. The domestic court system punishes those that violate contracts, thus creating the credibility or trust that serves as the foundation for domestic financial markets. But there is no international court system that endows sovereigns with the same sort of credibility, and this hampers international financial markets. I assume next this problem is so severe that it precludes all asset trade.

Unlike the integrated economy, in the world analyzed in this section each region’s total production, spending and capital stock are always determined. Since trade balances and current accounts are zero, the income of each region equals the value of both its production and spending, i.e. $Y_{c,t}=Q_{c,t}=E_{c,t}$. Since the only vehicle for savings available to the young is physical capital, analogues to Equations (1)-(2) apply to each region. We can therefore write regional incomes and the laws of motion of regional capital stocks as follows:

(29) $$Y_{c,t} = w_{c,t} \cdot H_{c,t} + r_{c,t} \cdot K_{c,t} \qquad \text{for all } c \in C$$

(30) $$K_{c,t+1} = S_{c,t} \cdot w_{c,t} \cdot H_{c,t} \qquad \text{for all } c \in C$$

These Equations apply to all the models of this section, including the world of autarky before globalization. Therefore, a complete analysis of the world income distribution and its evolution requires us to determine the cross-section of factor prices, i.e. $w_{c,t}$ and $r_{c,t}$ as a function of the state of the world economy. The latter consists of the savings, factor endowments and industry productivities of all regions of the world, i.e. $S_{c,t}$, $H_{c,t}$, $K_{c,t}$ and $Z_{c,it}$ for all $i\in I$ and all $c\in C$; plus the date, since trade in goods is only possible if $t\geq 0$.

The rest of this section is organized as follows. Sub-section 2.1 studies further the world of autarky, while the rest of the section studies the world after globalization. In sub-section 2.2, we explore a world in which frictions to factor mobility and asset trade are not binding after globalization. Sub-section 2.3 provides a formal description of the model. Sub-sections 2.4 and 2.5 examine worlds where frictions to factor mobility and asset trade remain binding after globalization.

2.1 Economic growth in autarky

The analysis of the effects of globalization starts in the world of autarky. As explained in section 1.3, before globalization each region is a smaller and less efficient version of the integrated economy and factor prices can be written as:[23]

$$(31)\qquad w_{c,t} = (1-\alpha)\cdot A_{c,t}\cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon-1}\cdot K_{c,t}^{\mu\cdot\alpha+\upsilon}\qquad \text{for all } c\in C$$

$$(32)\qquad r_{c,t} = \alpha\cdot A_{c,t}\cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon}\cdot K_{c,t}^{\mu\cdot\alpha+\upsilon-1}\qquad \text{for all } c\in C$$

Equations (31)-(32) describe the cross-section of factor prices. Holding constant factor endowments, regions with higher than average industry productivities have higher than average factor prices. Holding constant industry productivities, the relationship between factor prices and factor endowments depends on two familiar forces: diminishing returns and market size. For a given set of industry technologies, an increase in one factor makes this factor relatively more abundant, lowering its price and raising the price of the other factor. But an increase in one factor also raises income and demand in all industries, improving industry technologies and

[23] These Equations follow from Equation (24) and the observation that the shares of human capital and physical capital are 1-α and α.

raising the prices of both factors. Equations (31)-(32) put these two effects together. Hence, regions with higher-than-average human capital have higher-than-average rental rates for all parameter values, and also higher-than-average wages if $\mu\cdot(1-\alpha)-\upsilon>1$. Similarly, regions with higher-than-average physical capital have higher-than-average wages for all parameter values, and also higher-than-average rental rates if $\mu\cdot\alpha+\upsilon>1$.

It follows from Equations (29)-(32) that, before globalization, we can write regional incomes and capital stocks as follows:[24]

(33) $$Y_{c,t} = A_{c,t} \cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_{c,t}^{\mu\cdot\alpha+\upsilon} \qquad \text{for all } c\in C$$

(34) $$K_{c,t+1} = s_{c,t} \cdot A_{c,t} \cdot H_{c,t}^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_{c,t}^{\mu\cdot\alpha+\upsilon} \qquad \text{for all } c\in C$$

Equation (33) shows the income of regions, and it can be used to determine the relative contribution of factor endowments and productivity to income differences. For instance, assume income is λ times higher than average in a given region. It could be that in this region human capital is $\lambda^{1/(\mu\cdot(1-\alpha)-\upsilon)}$ higher than average or that physical capital is $\lambda^{1/(\mu\cdot\alpha+\upsilon)}$ higher than average. It could also be that the region's productivity in industry i is $\lambda^{1/\sigma_i\cdot\mu_i}$ times higher than average.[25] Naturally, it could also be any combination of these factors.

Equation (34) is the law of motion of the capital stocks and can be used to analyze the dynamic response to a region-specific shock to savings, human capital and/or industry productivity. Positive (and permanent) shocks to any of these variables raise the region's capital stock and income. As Equation (34) shows, these shocks have growth effects if $\alpha\cdot\mu+\upsilon\geq1$, but only have level effects if $\alpha\cdot\mu+\upsilon<1$. Regardless of the case, the effects of these shocks never spill over to other regions.

[24] These Equations are identical to Equations (24)-(25) and have been reproduced here only for convenience.
[25] Here industry productivity means $Z_{c,it}$, and not $A_{c,it}$.

Assume the joint distribution of savings, human capital and industry productivities is stationary. Then, Equations (33)-(34) imply a strong connection between the cross-sectional and time-series properties of the growth process.. If diminishing returns are strong and market size effects are weak, i.e. if $\mu \cdot \alpha + \upsilon < 1$, world average income ($Y_t$) and its regional distribution ($Y_{c,t}/Y_t$) are both stationary. If instead diminishing returns are weak and market size effects are strong, i.e. $\mu \cdot \alpha + \upsilon > 1$, world average income and its regional distribution are both non-stationary. This result provides a tight link between the long run properties of the growth process and the stability of the world income distribution. A weaker version of this result assumes that the world productivity frontier (A_t) is non-stationary but regional productivity gaps ($A_{c,t}/A_t$) are stationary. Under this assumption, world average income is non-stationary even if diminishing returns are strong and market size effects are weak.

It is commonplace among growth theorists to interpret cross-country data from the vantage point of the autarky model.[26] One influential example is the work of Mankiw, Romer and Weil [1992]. They combined Equations (33)-(34) to obtain an Equation relating income to savings, human capital, country productivity, and lagged income; and estimated it using data for a large cross-section of countries. They interpreted the residuals of this regression as measuring differences in country productivities and measurement error, and concluded that differences in savings and human capital explain (in a statistical sense) about 80 percent of the cross-country variation in income. Their procedure imposed the restriction $\mu=1$ (and therefore $\upsilon=0$) and yielded an estimate of α of about two thirds. Hall and Jones [1999] and Klenow and Rodríguez-Clare [1997] interpreted this high estimate of α as a signal that the regression was miss-specified. Their argument was that savings, human capital and productivity were positively correlated and the omission of productivity from the

[26] Unfortunately, the absence of direct and reliable measures of productivity precludes carrying out formal tests of the theory. The most popular empirical response to this problem has been to simply assume the theory is correct and use available data to make inferences about the determinants of the world income distribution and its evolution.

regression biased upwards the estimate of α. These authors used Equations (33)-(34) to calibrate country productivities keeping the assumption that μ=1, but instead imposing a value of α of about one third.[27] With these productivities at hand, they found that about two thirds of the variation in incomes reflects variation in productivity, and only one third can be attributed to cross-country variation in savings and human capital.

Another influential example of the use of the autarky model to interpret available data is Barro [1991] who found that, after controlling for human capital and saving rates, poor countries tend to grow faster than rich ones. This finding has been labeled "conditional convergence" since it implies that, if two countries have the same country characteristics, they converge to the same level of income.[28] If Equations (33)-(34) provide a good description of the real world, observing conditional convergence is akin to finding that $\mu \cdot \alpha + \upsilon < 1$.[29] Many have therefore interpreted the conditional convergence finding as evidence that diminishing returns are strong relative to market size effects.

These inferences about the nature of the growth process heavily rely on Equations (33)-(34), and these Equations have been derived from a theoretical model that assumes that all regions of the world live in autarky. This assumption is obviously unrealistic. Is it also crucial? And if so, what alternative assumption would be reasonably realistic? I next turn to these questions. But the script should not be surprising. Globalization (as described at the beginning of this section) has profound effects on the world income distribution and its evolution. The newfound ability of regions to specialize and trade alters, sometimes quite dramatically, the effects of

[27] This value corresponds to the share of capital in income in national accounts. This sort of calibration exercise is known as development accounting. Caselli's chapter in this volume is the definitive source on this topic.
[28] As Barro himself emphasized, this does not mean that per capita incomes tend to converge unconditionally since countries with high initial incomes also tend to have good country characteristics. There is a large number of papers that try to determine whether there is conditional convergence and measure how fast it takes place. See, for instance, Knight, Loayza and Villanueva [1993] and Caselli, Esquivel and Lefort [1996].
[29] An additional maintained assumption of this line of research is that savings, human capital and productivity are jointly stationary.

factor endowments and industry productivities on factor prices. This is most clearly illustrated in subsection 2.2, which depicts a world in which goods trade allows the world economy to replicate the prices and allocations of the integrated economy. Of course, this is not a general feature of goods trade. Subsection 2.3 prepares the ground for the analysis of worlds where economic integration is imperfect and factor prices vary across regions. This analysis is then performed in subsections 2.4 and 2.5.

2.2 Factor price equalization

A good starting point for the analysis of the world economy after globalization is to ask whether restricting factor mobility matters at all. Somewhat surprisingly, the answer is "perhaps not". As Paul Samuelson [1948, 1949] showed more than half a century ago, goods trade might be all that is needed to ensure global efficiency. When this happens, we say that the equalization of goods prices leads to the equalization of factor prices. I shall describe Samuelson's result and its implications step by step, so as to develop intuition.[30]

Consider the set of all possible partitions of the world factor endowments at date t, H_t and K_t, among the different regions of the world or, for short, the set of all possible factor distributions. This set is formally defined as follows:

$$(35) \qquad D_t \equiv \left\{ (H_{c,t}, K_{c,t}) \text{ for all } c \in C \;\middle|\; H_{c,t} \geq 0,\ K_{c,t} \geq 0 \text{ s.t. } \sum_{c \in C} H_{c,t} = H_t \text{ and } \sum_{c \in C} K_{c,t} = K_t \right\}$$

Define FPE_t as the subset of D_t for which the world economy replicates the prices and allocations of the integrated economy. To construct FPE_t, fix $d_t \in D_t$ and consider the integrated economy prices and quantities. At these prices, consumers

[30] The analysis here follows a long tradition in international trade. See Norman and Dixit [1989], Helpman and Krugman [1985] and Davis [1995].

are willing to purchase the integrated economy quantities of the different goods and also have enough income to do so. At these prices, producers located in regions with the highest industry productivities are willing to produce the integrated economy quantities of the different goods using the integrated economy quantities of factors. If these producers can find these quantities of factors in their regions, the integrated economy prices and quantities are in fact the equilibrium ones and we say that $d_t \in FPE_t$. Otherwise, the integrated economy prices and quantities cannot be the equilibrium ones and we say that $d_t \notin FPE_t$. Therefore, the set FPE_t can be formally defined as follows:

(36)

$$FPE_t \equiv \left\{ d_t \in D_t \;\middle|\; \begin{array}{l} \exists x_{c,it}(m) \geq 0,\ x^F_{c,it} \geq 0 \text{ with} \sum_{c \in C} x_{c,it}(m) = 1,\ \sum_{c \in C} x^F_{c,it} = 1 \text{ and} \\ x_{c,it} = (1-\beta_i) \cdot x^F_{c,it} + \frac{\beta_i}{M_{it}} \cdot \int_0^{M_{it}} x_{c,it}(m) \cdot dm; \text{ such that:} \\ \text{(R1) } x_{c,it} = 0 \text{ if } Z_{c,it} < \max_{c \in C}\{Z_{it}\}; \\ \text{(R2) } H_{c,t} = \sum_{i \in I} x_{c,it} \cdot H_{it} \text{ and } K_{c,t} = \sum_{i \in I} x_{c,it} \cdot K_{it}; \text{ and} \\ \text{(R3) } x_{c,it}(m) \in \{0,1\} \text{ for all } m \in [0, M_{it}] \text{ and } i \in I \end{array} \right\}$$

where M_{it}, H_{it} and K_{it} are defined in Equations (17), (19) and (20). To understand this definition, interpret $x_{c,it}$ as the share of the world production of industry i located in region c at date t; and note that this share includes the production of intermediate inputs, $x_{c,it}(m)$, and final goods, $x^F_{c,it}$. Definition (36) then says that $d_t \in FPE_t$ if it is possible to achieve full employment of human and physical capital in all regions producing only in those regions with the highest productivity (requirement R1), using the same factor proportions as in the integrated economy (requirement R2), and without incurring the fixed cost of production more than once (requirement R3). The set FPE_t is never empty since the factor distribution that applies in the integrated economy always belongs to it. In fact, the set FPE_t consists of all the factor distributions that are equilibria of the integrated economy. The larger is the size of the indeterminacy in the geographical distribution of production and factors of the integrated economy, the larger is the size of FPE_t.

The patterns of production and trade that support factor price equalization after globalization are easy to state and quite intuitive:

1. *In regions where human (physical) capital is relatively abundant, production shifts towards industries that, on average, use human (physical) capital intensively. Excess production in these industries is converted into exports that finance imports of industries that use physical (human) capital intensively.*

Example 2.1.1: Consider a world economy with H- and K-industries, such that $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume $\alpha_i=\alpha_H$ if $i \in I^H$, and $\alpha_i=\alpha_K$ if $i \in I^K$; and $\alpha_H<\alpha_K$; and $\beta_i=0$ for all $i \in I$. All regions have the same industry productivities, but A-regions have a higher ratio of human to physical capital than B-regions. Factor price equalization is possible if the differences in factor proportions between A- and B-regions are not too large relative to the differences in factor proportions between H- and K-industries. Figure 8 shows the geometry of this example. Since all regions have the same factor costs, industries use the same factor proportions in all regions. A-regions contain a more than proportional fraction of the integrated economy's H-industry, and a less than proportional fraction of the K-industry. The opposite happens in B-regions. This is how specialization and trade ensure that in this world economy factor endowments are used efficiently.

2. *In industries where a region's productivity is less than the world's highest, production falls to zero and domestic spending shifts towards imports. To finance*

the latter, production expands in industries in which the region has the highest possible productivity and the excess production is exported abroad.

Example 2.1.2: Consider a world economy with H- and K-industries, such that $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume $\alpha_i=\alpha_H$ if $i \in I^H$, and $\alpha_i=\alpha_K$ if $i \in I^K$; and $\alpha_H<\alpha_K$; and $\beta_i=0$ for all $i \in I$. Within each type there are "advanced" and "backward" industries. A-regions have the highest possible productivity in all industries, regardless of whether they are "advanced" or "backward". B-regions have the highest possible productivity only in "backward" industries. Factor price equalization is possible if the combined factor endowments of A-regions are large enough and the subset of "advanced" industries is not too large. Figure 9 shows the geometry of this example. Since all regions have the same factor costs, only producers located in regions with the highest productivity can survive international competition. A-regions produce the integrated economy quantities of "advanced" goods and a fraction of the integrated economy quantities of "backward" goods. B-regions produce the remaining quantities of "backward" goods. This is how specialization and trade ensure that in this world economy production takes place only where industry productivities are higher.

3. *Within each industry, only one region produces each input variety and exports it to all other regions. If an industry is split among various regions, there is likely to be two-way trade within the same industry*.[31]

Example 2.1.3: Consider any of the world economies of the previous examples, but assume now that $\beta_i=1$ for all $i \in I$. Assume $d_t \in FPE_t$. Since the fixed costs of

[31] I say "likely to be" because a region might produce the final good for domestic use, and import the necessary input varieties. It is usual in trade models to set $\beta_i=1$ and then drop the "likely to be" from the statement.

producing inputs contain the cost of building a specialized production plant, all input producers choose to concentrate their production in one region in order not to duplicate these costs. Therefore, each region produces a disjoint set of input varieties. This is how specialization and trade allow the world economy to exploit increasing returns to scale and therefore benefit from a larger market size.

By adopting these patterns of specialization and trade, the world economy is able to reap all the benefits of economic integration without any factor movements. Using the jargon of trade theory, goods trade is a "perfect substitute" for factor movements if $d_t \in FPE_t$. When this is the case, factor prices are given by:

$$w_{c,t} = (1-\alpha)\cdot A_t \cdot H_t^{\mu\cdot(1-\alpha)-\upsilon-1}\cdot K_t^{\mu\cdot\alpha+\upsilon} \qquad \text{for all } c\in C \tag{37}$$

$$r_{c,t} = \alpha\cdot A_t \cdot H_t^{\mu\cdot(1-\alpha)-\upsilon}\cdot K_t^{\mu\cdot\alpha+\upsilon-1} \qquad \text{for all } c\in C \tag{38}$$

The world economy is able to operate at the same level of efficiency as the integrated economy despite the immobility of factors. Equations (31)-(32) showed that, before globalization, cross-regional differences in factor proportions and industry productivities lead to differences in the way industries operate (i.e. their factor proportions and productivity) and also in the size of their markets. Regions with a high ratio of human to physical capital have high wage-rental ratios. Regions with high industry productivities and abundant human and physical capital have high factor prices. But Equations (37)-(38) show that, after globalization (and if $d_t \in FPE_t$), cross-regional differences in factor proportions and industry productivities neither change the way industries operate, nor do they affect the size of their markets. Goods trade allows regions to absorb their differences in factor endowments and industry productivities by specializing in those industries that use their abundant factors and have the highest possible productivity, without the need for having different factor prices. Goods trade also eliminates the effects of regional size on factor prices by creating global markets.

These observations have important implications for the world income distribution and, consequently, for any attempt to determine the relative contribution of factor endowments and productivity to income differences. Substituting Equations (37)-(38) into Equation (29), we find that:

$$(39) \qquad Y_t = A_t \cdot H_t^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_t^{\mu\cdot\alpha+\upsilon}$$

$$(40) \qquad \frac{Y_{c,t}}{Y_t} = (1-\alpha)\cdot\frac{H_{c,t}}{H_t} + \alpha\cdot\frac{K_{c,t}}{K_t} \qquad \text{for all } c \in C$$

A comparison between these Equations and Equation (33) shows that the relative contribution of factor endowments and productivity to income differences is fundamentally affected by globalization. Equation (33) differs from Equations (39)-(40) in three important respects: the elasticity of substitution between domestic human and physical capital is one in Equation (33) but infinite in Equations (39)-(40); domestic productivity appears in Equation (34) but not in Equations (39)-(40); and income is homogeneous of degree μ on domestic factor endowments in Equation (34) but only of degree one in Equation (39)-(40). Each of these differences echoes a different aspect of globalization, and I shall discuss them in turn.

Globalization raises the elasticity of substitution between human and physical capital from one to infinity because structural transformation (a shift towards industries that use the abundant factor) replaces factor deepening (forcing industries to use more of the locally abundant factor) as a mechanism to absorb differences in factor proportions. Assume a region has a ratio of human to physical capital λ times higher than average. Before globalization, each of its industries is forced to operate with a ratio of human to physical capital that is λ times average, and this requires a wage-rental ratio that is λ^{-1} times average. After globalization, the region simply shifts its production towards industries that are human-capital intensive, keeping the ratio of human to physical capital of its industries constant. This does not require changes in the wage-rental ratio.

Globalization eliminates differences in industry productivities as a source of income differences because structural transformation (a shift towards industries that have high productivity) also replaces productivity deepening (forcing low-productivity industries to produce) as a mechanism to absorb differences in industry productivities. Assume now that a region has average factor endowments but higher than average industry productivities. For instance, the region's productivity is λ times higher than the rest of the world in a subset of industries of combined size σ, and equal to the rest of the world in the remaining ones. Before globalization, this productivity advantage allows the region to produce $\lambda^{\sigma\cdot\mu}$ output than average with the same factors, holding constant technology. After globalization, the region takes over all world production of those industries in which its productivity is higher and scales back the rest of its industries. This allows the rest of the world to take full advantage of the region's high productivity and catch up with it in terms of income (even though not in productivity).

Globalization reduces the effects of factor endowments on relative incomes because it converts regional markets into global ones. Assume now that a region has average industry productivities, but its human and physical capitals are both λ times above average. Before globalization, the region's higher factor endowments allow it to produce more output than the average region. This effect is further reinforced because the region's larger market size allows it to have a better technology than average. Therefore, in autarky the region's income is λ^{μ} times higher than the world's average. After globalization, this additional market size effect disappears since the relevant market is the world market and this is the same for all regions. Therefore, after globalization the region's income is only λ times higher than the world average income.

Globalization also influences the dynamics of the world economy. Assume $d_t \in FPE_t$ for all t, then it follows from Equation (30) and (37)-(38) that:

(41) $K_{t+1} = s_t \cdot A_t \cdot H_t^{\mu\cdot(1-\alpha)-\upsilon} \cdot K_t^{\mu\cdot\alpha+\upsilon}$

(42) $\frac{K_{c,t+1}}{K_{t+1}} = \frac{S_{c,t} \cdot H_{c,t}}{S_t \cdot H_t}$ for all $c \in C$

A comparison between these Equations and Equation (34) shows how globalization affects the dynamic responses to region-specific shocks. After globalization, positive (and permanent) shocks to savings and human capital still raise a region's capital stock and income. But now the effects of these shocks spill over to other regions. Shocks to productivity can only affect a region's income if they push outward the world productivity frontier. And, in this case, all countries equally benefit.[32]

Another important implication of Equations (39)-(42) is that globalization breaks down the connection between the long run properties of the growth process and the stability of the world income distribution.[33] Assume again that the joint distribution of savings, human capital and productivities is stationary. Then, Equation (41) shows that it still is the relative strength of diminishing returns and market size effects that determines whether world average income is stationary or not. But Equation (42) shows that now the world distribution of capital stocks is stationary regardless of parameter values. The same applies to the world income distribution (see Equation (40)). Therefore, all regions share a common growth rate in the long run. The reason is simple: physical capital accumulation in high-savings and high-human capital regions is absorbed by increased production in industries that use physical capital intensively, and this lowers the prices of these industries and increases the prices of industries that use human capital intensively. This increases wages and savings in low-savings and low-human capital regions. In a nutshell,

[32] See Ventura [1997] and Atkeson and Kehoe [2000] for analyses of shocks to small open economies in the factor-price-equalization world.
[33] Ventura [1997] provides a dramatic example of this by constructing a world in which time-series convergence to a steady state is associated with cross-sectional divergence and vice versa.

movements in goods prices positively transmit growth across regions and ensure the stability of the world income distribution. [34]

The main feature of the factor-price-equalization world is that diminishing returns and market size effects are global and not local. This observation has important implications for growth theory. Explanations for why the world grows faster today than in the past should feature diminishing returns and market size effects in the lead role, and relegate savings and human capital to a secondary one. But explanations of why some countries grow faster than others should do exactly the opposite, giving the lead role to savings and human capital and relegating diminishing returns and market size effects to a secondary role. A distinctive feature of the integrated economy is therefore a sharp disconnect between the determinants of average or long run growth and the determinants the dispersion or the cross-section of growth rates.[35]

The factor-price-equalization world neatly illustrates the potential effects of trade on the world income distribution and its dynamics, and it shows why and how goods trade can be a perfect substitute for factor movements. But the real world has not achieved yet the degree of economic integration that this model implies. One does not need sophisticated econometrics to conclude that wages vary substantially around the world. It is less obvious but probably true as well that rental rates also vary substantially around the world. These differences in factor prices indicate that regional differences in factor endowments and/or industry productivities are so large that goods trade cannot make up for factor immobility.

---

[34] As a general proposition, it is not necessary that trade leads to the stability of the world income distribution. In fact, the study of the stability of world income distribution has received considerable attention recently. While Acemoglu and Ventura [2002] rely on specialization to generate a stable world income distribution, Deardorff [2001] presents a model in which mere differences in initial endowments create persistent difference in world income and “club convergence”. Krugman [1987] and Howitt [2000] rely on endogenous technology change to generate such effects. See Brezis, Krugman, and Tsiddon [1993] for a model of human capital accumulation that explains leapfrogging in the international income distribution.

[35] One implication of this is that Barro’s conditional convergence finding cannot be used to determine whether diminishing returns are weak or strong relative to market size effects. See Ventura [1997].

What trade always does is to create a global market in which only the most competitive producers of the world can survive. Trade forces high-cost industries to close down and offers low-cost industries the opportunity to grow. If $d_t \in FPE_t$, all regions contain enough of these low-cost industries to employ all of their factors at common or equalized factor prices. But this need not be always the case. If $d_t \notin FPE_t$, regions with low industry productivities and sizeable factor endowments are forced to offer cheap factors to compete, while regions with high industry productivities and small factor endowments are able to enjoy expensive factors. These price differences indicate that factors are not deployed where they should and the world economy does not operate efficiently. To study the origins and effects of these world inefficiencies, it is necessary first to review some formal aspects of the model after globalization.

2.3 Formal aspects of the model

As mentioned already, in the absence of asset trade analogues of Equations (1) and (2) apply now to each region of the world economy. A regional analogue to Equation (3) also applies since it is a direct implication of our Cobb-Douglas assumption for the consumption and investment composites. Since all regions share spending patterns and face the same goods prices, the price of consumption and investment is the same for all. We keep this common price as the numeraire and, as a result, Equation (4) also applies. Equations (5)-(6) describing technology apply to all regions, with the corresponding factor prices and industry productivities.

After globalization, Equations (7)-(10) describing pricing policies, input demands and the free-entry condition apply only to those regions that host the lowest-cost producers of the world. The rest cannot compete in global markets. To formalize this notion, define the following sets of industries:

$$(43)\qquad I_{c,t} \equiv \left\{ i \in I \;\middle|\; c \in \underset{c' \in C}{\arg\min} \left\{ \frac{1}{Z_{c',it}} \cdot \left(\frac{w_{c',t}}{1-\alpha_i} \right)^{1-\alpha_i} \cdot \left(\frac{r_{c',t}}{\alpha_i} \right)^{\alpha_i} \right\} \right\} \qquad \text{for all } c \in C$$

An industry belongs to $I_{c,t}$ if and only if producers located in region c are capable of competing internationally in this industry at date t.[36] Note that a region can be competitive in a given industry because it offers high productivity or a cheap combination of factor prices. The main implication of goods trade is that industries do not locate in regions where they are not competitive:

$$(44)\qquad Q_{c,it} = 0 \qquad \text{if } i \notin I_{c,t} \qquad \text{for all } i \in I \text{ and } c \in C$$

Since goods markets are integrated, Equation (11) describing market clearing in global goods markets still applies. But now Equations (12)-(13) describing market clearing in global factor markets must be replaced by analogue conditions imposing market clearing in each regional factor market:

$$(45)\qquad H_{c,t} = \sum_{i \in I} H_{c,it} \qquad \text{with } H_{c,it} = \frac{\partial B_{c,it}}{\partial w_{c,t}} + \int_0^{M_{it}} \frac{\partial b_{c,it}(m)}{\partial w_{c,t}} \cdot dm \qquad \text{for all } c \in C$$

$$(46)\qquad K_{c,t} = \sum_{i \in I} K_{c,it} \qquad \text{with } K_{c,it} = \frac{\partial B_{c,it}}{\partial r_{c,t}} + \int_0^{M_{it}} \frac{\partial b_{c,it}(m)}{\partial r_{c,t}} \cdot dm \qquad \text{for all } c \in C$$

Equations (45)-(46) state that the regional supplies of labor and capital must equal their regional demands. The latter are the sum of their industry demands, and these are calculated by applying Shephard’s lemma to Equations (5) and (6).

[36] This follows directly from the cost functions in Equations (5)-(6) and the observation that all producers in the world face the same world demand.

This completes the formal description of the model. For any admissible set of capital stocks, i.e. $K_{c,0}$ for all $c \in C$, and sequences for the vectors of savings, human capital and industry productivities, i.e. $S_{c,t}$, $H_{c,t}$, and $A_{c,it}$ for all $c \in C$ and for all $i \in I$, an equilibrium of the world economy after globalization consists of sequences of prices and quantities such that the equations listed above hold at all dates and states of nature. Although there might be multiple geographical patterns of production and trade that are consistent with world equilibrium, the assumptions made ensure that prices and world aggregates are uniquely determined.[37]

We are ready now to re-examine the effects of globalization on factor prices and the world income distribution. We have already found that, if $d_t \in FPE_t$, globalization eliminates all regional differences in factor prices and permits the world economy to operate at the same level of efficiency as the integrated economy. In this case, global market forces are strong enough to ensure that diminishing returns and market size effects have a global rather than a regional scope. This is no longer the case if $d_t \notin FPE_t$ since globalization cannot eliminate all regional differences in factor prices. These factor price differences reflect inefficiencies of various sorts in the world economy.

Efficiency requires that factor usage within an industry be the same across regions. This is a direct implication of assuming diminishing returns to each factor in production. The problem, of course, is that regional factor proportions vary. Structural transformation allows regions to accommodate all or part of their differences in factor proportions without factor deepening. If there are enough industries that use different factor proportions, factor prices are equalized across regions. If there are not enough industries that use different factor proportions, regions must lower the price of their abundant factor and raise the price of their

---

[37] Despite the indeterminacy in trade patterns, the trade theorist will immediately recognize that, if $\beta_i=1$ for all $i \in I$, the volume of trade is determined and the popular gravity equation applies to this world economy.

scarce one to attract enough firms to employ their factor endowments. In this case, industries in different regions use different factor proportions and the world economy is inefficient. Subsection 2.4 studies the properties of the growth process in this situation.

Efficiency also requires that industries locate in those regions that offer them the highest possible productivity. Structural transformation allows regions to accommodate all or part of their differences in industry productivities without productivity deepening. If all regions have enough industries with the highest productivity, factor prices are equalized across regions. If some regions do not have enough industries with the highest productivity, they are forced to produce in low productivity industries and must lower their factor prices to be able to compete internationally. Subsection 2.5 shows how this affects the properties of the growth process.

In the presence of these two types of inefficiency, diminishing returns retain a regional scope even after globalization. Regional differences in factor prices still reflect regional differences in factor abundance and industry productivities, although the mapping between these variables is much more subtle than in the world of autarky. However, even in the presence of these inefficiencies regional differences in factor prices cannot reflect regional differences in market size. For market size effects to retain a regional scope after globalization we need to introduce impediments to goods trade. And this task is left for section 3.

2.4 Limits to structural transformation (I): factor proportions

It follows from Definition (36) that factor prices are equalized if and only if it is possible to achieve full employment of human and physical capital in all regions producing only with the highest productivity (requirement R1), with the factor proportions used in the integrated economy (requirement R2), and without incurring a fixed cost more than once (requirement R3). Moving away from the factor-price-equalization world means that we must consider the violation of one or more of these requirements. Since the market for each input is “small”, I assume that regions are large enough to ensure that requirement R3 is always satisfied.[38] Therefore, in the remainder of this section I will focus on violations of requirements R1 and R2. In this subsection, we study the effects on the growth process of violations to requirement R2, keeping the assumption that requirement R1 is not binding. This assumption will be removed in sub-section 2.5.

To formalize the notion that requirement R1 is not binding, define $I^*_{c,t}$ as the set of industries in which region c has the highest possible productivity: $I^*_{c,t} \equiv \left\{ i \in I \,\middle|\, c \in \arg\max_{c' \in C} \{Z_{c',it}\} \right\}$ for all $c \in C$. To ensure that requirement R1 is not binding in the models of this section, for each of them I first construct the set of “unrestricted” world equilibria by assuming that $I^*_{c,t} = I$ for all $c \in C$. As mentioned, all these equilibria share the same prices and world aggregates, but might exhibit different geographical patterns of production. In these “unrestricted” world equilibria, some industries might not operate in all regions. Naturally, prices and world aggregates would not be affected if regions did not have the best possible technologies in some or all of the industries in which they do not produce. Therefore, we can trivially relax the assumption that $I^*_{c,t}$ contains all industries, and instead assume only that there exists an “unrestricted” equilibrium such that, for all $c \in C$, the industries not included in $I^*_{c,t}$ do not operate in the region. This defines the extent to which regional differences in industry productivities are allowed in this section. It follows that requirement R1 is never binding and comparative advantage is determined solely by regional differences in factor proportions.

[38] I shall explore the effects of violations to requirement R3 in sections 3.2 and 3.3.

In the worlds we consider in this sub-section it is not possible in general to employ all factors in all regions using the techniques of the integrated economy. Even if they concentrate all of their production in industries that use human capital intensively, regions with abundant human capital might lack enough physical capital to produce with the factor proportions that these industries would use in the integrated economy. These regions are therefore forced to use a higher proportion of human capital in their industries and this requires them to have a lower wage-rental ratio than in the integrated economy. Naturally, the exact opposite occurs in regions with abundant physical capital. This situation can be aptly described as a geographical mismatch between different factor endowments.

To study the causes and effects of this mismatch, I present two examples that help build intuitions that apply more generally. The first example is the two-industry case that is so popular in trade theory:

Example 2.4.1: Consider a world economy with H- and K-industries, $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume $\alpha_i=\alpha_H$, $\sigma_i=\sigma_H$ and $\max_{c\in C}\{Z_{c,it}\} = Z_{Ht}$ if $i\in I^H$, $\alpha_i=\alpha_K$, $\sigma_i=\sigma_K$ and $\max_{c\in C}\{Z_{c,it}\} = Z_{Kt}$ if $i\in I^K$, with $\alpha_H\leq\alpha_K$. (Note that $I^H\cdot\sigma_H + I^K\cdot\sigma_K=1$) For simplicity, assume also that $\varepsilon_i=\varepsilon$ and $\beta_i=\beta$ for all $i\in I$. The first step is to relate prices and world income to production:[39]

$$(47) \qquad P_{it} = \sigma_i \cdot \frac{\prod_{i'\in I}\left(\sum_{c\in C} Q_{c,i't}\right)^{\sigma_i}}{\sum_{c\in C} Q_{c,it}} \qquad \text{for all } i\in I$$

$$(48) \qquad Y_t = \prod_{i\in I}\left(\sum_{c\in C} Q_{c,it}\right)^{\sigma_i}$$

[39] These Equations follow from Equations (3) and (4).

Equation (47) can be thought of as the “demand” side of the model, since it shows how prices depend negatively on quantities, while Equation (48) simply describes world income. The “supply” side of the model is given by the following set of Equations:[40]

$$(49)\quad \frac{(1-\alpha_k)\cdot P_{Kt}}{w_{c,t}}\cdot\sum_{i\in I^K} Q_{c,it}+\frac{(1-\alpha_H)\cdot P_{Ht}}{w_{c,t}}\cdot\sum_{i\in I^H} Q_{c,it}=H_{c,t}\qquad \text{for all } c\in C$$

$$(50)\quad \frac{\alpha_k\cdot P_{Kt}}{r_{c,t}}\cdot\sum_{i\in I^K} Q_{c,it}+\frac{\alpha_H\cdot P_{Ht}}{r_{c,t}}\cdot\sum_{i\in I^H} Q_{c,it}=K_{c,t}\qquad \text{for all } c\in C$$

$$(51)\quad \left(\frac{w_{c,t}}{1-\alpha_H}\right)^{1-\alpha_H}\cdot\left(\frac{r_{c,t}}{\alpha_H}\right)^{\alpha_H}\geq\frac{\varepsilon-1}{\varepsilon}\cdot Z_{Ht}\cdot p_{Ht}=\left(A_{Ht}\cdot P_{Ht}\right)^{\frac{1}{\mu}}\cdot\left(\sigma_H\cdot Y_t\right)^{\frac{\mu-1}{\mu}}=\phi_H\cdot f_{Ht}\quad \text{for all } c\in C$$

$$(52)\quad \left(\frac{w_{c,t}}{1-\alpha_K}\right)^{1-\alpha_K}\cdot\left(\frac{r_{c,t}}{\alpha_K}\right)^{\alpha_K}\geq\frac{\varepsilon-1}{\varepsilon}\cdot Z_{Kt}\cdot p_{Kt}=\left(A_{Kt}\cdot P_{Kt}\right)^{\frac{1}{\mu}}\cdot\left(\sigma_K\cdot Y_t\right)^{\frac{\mu-1}{\mu}}=\phi_K\cdot f_{Kt}\quad \text{for all } c\in C$$

where $\phi_i=(1-\alpha_i)^{\alpha_i-1}\cdot\alpha_i^{-\alpha_i}$ for all $i\in I$; and f_{Ht} and f_{Kt} are measures of the lowest factor costs in the world for the H- and K-industries since in equilibrium $f_{it}=\min_{c\in C}\left\{w_{c,t}^{1-\alpha_i}\cdot r_{c,t}^{\alpha_i}\right\}$ for all $i\in I$. Equations (49)-(50) are factor market clearing conditions, while Equations (51)-(52) are just a transformation of the pricing equations of each industry (for both final goods and intermediate inputs). Naturally, these pricing equations hold with strict equality if there is positive production in the corresponding industry. Equations (49)-(52) determine the production of each type of industry and the factor prices of region c, as a function of world prices and income.[41]

[40] Equations (49)-(50) follow from Equations (45) and (46), while Equations (51)-(52) follow from Equations (7) and (9) after using Equation (17) to eliminate the number of input varieties.

[41] If one is willing to take goods prices and factor endowments parametrically and further assume that the pricing equations hold with strict equality, it is possible to derive two popular results of trade theory from Equations (49)-(52). The Stolper-Samuelson effect says that an increase in the relative price of an industry leads to a more than proportional increase in the price of the factor that is used intensively in this industry and a decline in the price of the other factor. The Rybcynski effect says that an increase in a factor endowment leads to a more than proportional increase in the production of the industry that uses this factor intensively and a decline in the production of the other industry.

Equations (47)-(52) determine prices and quantities as a function of the distribution of factor endowments. Together with the regional analogues to Equation (1), the initial condition and the dynamics of the exogenous state variables, these Equations provide a complete characterization of the world equilibrium. Next, I describe some its most salient features.

Regions with extreme factor proportions have specialized production structures, while regions with intermediate factor proportions have diversified production structures. Let C_{Kt} (C_{Ht}) be the set of regions where there is production only in K-industries (H-industries), and let C_{Mt} be the set of regions where there is production in both types of industries. In fact, it follows from Equations (49)-(52) that these sets of regions are defined as follows:

$$(53)\qquad C_{Kt} = \left\{ c \in C \;\middle|\; \frac{H_{c,t}}{K_{c,t}} \leq \frac{1-\alpha_K}{\alpha_K} \cdot \left(\frac{f_{Kt}}{f_{Ht}} \right)^{\frac{1}{\alpha_K - \alpha_H}} \right\}$$

$$(54)\qquad C_{Ht} = \left\{ c \in C \;\middle|\; \frac{H_{c,t}}{K_{c,t}} \geq \frac{1-\alpha_H}{\alpha_H} \cdot \left(\frac{f_{Kt}}{f_{Ht}} \right)^{\frac{1}{\alpha_K - \alpha_H}} \right\}$$

$$(55)\qquad C_{Mt} = \left\{ c \in C \;\middle|\; \frac{1-\alpha_K}{\alpha_K} \cdot \left(\frac{f_{Kt}}{f_{Ht}} \right)^{\frac{1}{\alpha_K - \alpha_H}} < \frac{H_{c,t}}{K_{c,t}} < \frac{1-\alpha_H}{\alpha_H} \cdot \left(\frac{f_{Kt}}{f_{Ht}} \right)^{\frac{1}{\alpha_K - \alpha_H}} \right\}$$

It follows from Equations (51)-(52) that factor prices are the same for all $c \in C_M$. If the dispersion in regional factor proportions is not too large, and the dispersion in factor intensities is not too low, $C_{Kt}=C_{Ht}=\emptyset$ and there is factor price equalization. Otherwise, this world economy exhibits a limited version of the factor price equalization result since factor prices are still equalized for all $c \in C_{Mt}$. It is common in trade theory to refer to a group of regions that share the same factor prices as a "cone of diversification". In fact, we can write the wage and the rental as a function of f_{Ht} and f_{Kt} as follows:

$$(56) \qquad w_{c,t} = \begin{cases} (1-\alpha_K)\cdot\phi_K\cdot f_{Kt}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{-\alpha_K} & \text{if } c \in C_{Kt} \\ f_{Kt}^{\frac{-\alpha_H}{\alpha_K-\alpha_H}}\cdot f_{Ht}^{\frac{\alpha_K}{\alpha_K-\alpha_H}} & \text{if } c \in C_{Mt} \\ (1-\alpha_H)\cdot\phi_H\cdot f_{Ht}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{-\alpha_H} & \text{if } c \in C_{Ht} \end{cases}$$

$$(57) \qquad r_{c,t} = \begin{cases} \alpha_K\cdot\phi_K\cdot f_{Kt}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{1-\alpha_K} & \text{if } c \in C_{Kt} \\ f_{Kt}^{\frac{1-\alpha_H}{\alpha_K-\alpha_H}}\cdot f_{Ht}^{\frac{\alpha_K-1}{\alpha_K-\alpha_H}} & \text{if } c \in C_{Mt} \\ \alpha_H\cdot\phi_H\cdot f_{Ht}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{1-\alpha_H} & \text{if } c \in C_{Ht} \end{cases}$$

The wage is continuous and weakly declining on the human to physical capital ratio, while the rental is also continuous but increasing on this same ratio. The most noteworthy feature of these relationships is that they exhibit a "flat" for the set of human to physical capital ratios that define the cone of diversification. The top panel of Figure 10 shows how the wage-rental varies with a region's ratio of human to physical capital. Regional differences in this ratio reflect factor abundance in the usual way. In regions with a high (low) ratio of human to physical capital the price of human capital is low (high) relative to physical capital. Factor prices do not reflect however regional differences in industry productivities and/or market size.

It is now straightforward to compute the world income distribution as a function of f_{Ht} and f_{Kt}:

(58) $$Y_{c,t} = \begin{cases} \phi_K \cdot f_{Kt} \cdot H_{c,t}^{1-\alpha_K} \cdot K_{c,t}^{\alpha_K} & \text{if } c \in C_{Kt} \\ f_{Kt}^{\frac{-\alpha_H}{\alpha_K-\alpha_H}} \cdot f_{Ht}^{\frac{\alpha_K}{\alpha_K-\alpha_H}} \cdot H_{c,t} + f_{Kt}^{\frac{1-\alpha_H}{\alpha_K-\alpha_H}} \cdot f_{Ht}^{\frac{\alpha_K-1}{\alpha_K-\alpha_H}} \cdot K_{c,t} & \text{if } c \in C_{Mt} \\ \phi_H \cdot f_{Ht} \cdot H_{c,t}^{1-\alpha_H} \cdot K_{c,t}^{\alpha_H} & \text{if } c \in C_{Ht} \end{cases}$$

We can use Equation (58) to re-evaluate earlier results about the relative contribution of factor endowments and industry productivities to income differences across regions. The first result is that the elasticity of substitution between human and physical capital is one outside the cone of diversification, but infinity within the cone. This elasticity reflects the relative importance of structural transformation and factor deepening as means to absorb regional differences in factor proportions. The second result is that regional differences in industry productivities continue not playing a role in determining regional income differences. This, of course, is not surprising given the assumption we have made about requirement R1 not being binding. The third and final result is that relative incomes are homogenous of degree one on factor endowments. This not surprising either since it simply confirms the absence of market size effects at the regional level.

We can also write the dynamics of the capital stock as a function of f_{Ht} and f_{Kt}:

(59) $$K_{c,t+1} = \begin{cases} S_{c,t} \cdot (1-\alpha_K) \cdot \phi_K \cdot f_{Kt} \cdot H_{c,t}^{1-\alpha_K} \cdot K_{c,t}^{\alpha_K} & \text{if } c \in C_{Kt} \\ S_{c,t} \cdot f_{Kt}^{\frac{-\alpha_H}{\alpha_K-\alpha_H}} \cdot f_{Ht}^{\frac{\alpha_K}{\alpha_K-\alpha_H}} \cdot H_{c,t} & \text{if } c \in C_{Mt} \\ S_{c,t} \cdot (1-\alpha_H) \cdot \phi_H \cdot f_{Ht} \cdot H_{c,t}^{1-\alpha_H} \cdot K_{c,t}^{\alpha_H} & \text{if } c \in C_{Ht} \end{cases}$$

The specific dynamics of this example are hard to determine, since f_{Ht} and f_{Kt} change from generation to generation. It is easy to construct examples in which the world economy moves towards factor-price equalization; examples in which the world economy moves away from factor-price equalization; or examples in which the world economy alternates between periods in which factor prices are equalized and periods in which they are not. These dynamics depend on all the parameters the model (including initial condition) and the evolution of the exogenous state variables, i.e. savings, human capital and industry productivities. Regardless of the specific dynamics, the world income distribution is stable if the joint distribution of these variables is stationary. Economic growth is positively transmitted across regions through changes in goods prices. This stabilizing role of trade is further reinforced by the fact that regions outside the cones cannot absorb capital accumulation through structural transformation and, consequently, experience diminishing returns in production.

Identifying cones of diversification is important because regional differences in factor proportions lead to structural transformation inside them, but to factor deepening outside them. In example 2.4.1, there is one of such cones and contains regions with intermediate factor proportions. Regions with extreme factor proportions do not belong to any cone. This need not be always the case, as the next example shows.

Example 2.4.2: Consider a world economy with H-, M- and K-industries, $I^H \cup I^M \cup I^K = I$, $I^H \cap I^M = \varnothing$, $I^H \cap I^K = \varnothing$ and $I^M \cap I^K = \varnothing$. Assume α_i=0 and $\max_{c \in C}\{Z_{c,it}\} = Z_{Ht}$ if $i \in I^H$; $\alpha_i = \alpha_M$ and $\max_{c \in C}\{Z_{c,it}\} = Z_{Mt}$ if $i \in I^M$; and α_i=1 and $\max_{c \in C}\{Z_{c,it}\} = Z_{Kt}$ if $i \in I^K$. For simplicity, assume also that $\varepsilon_i = \varepsilon$ and $\beta_i = \beta$ for all $i \in I$. The “demand” side of this model is still described by Equations (47)-(48), but the “supply side is now given by:

(60) $$\frac{(1-\alpha_M)\cdot P_{Mt}}{w_{c,t}} \cdot \sum_{i \in I^M} Q_{c,it} + \frac{P_{Ht}}{w_{c,t}} \cdot \sum_{i \in I^H} Q_{c,it} = H_{c,t} \qquad \text{for all } c \in C$$

(61) $$\frac{P_{Kt}}{r_{c,t}} \cdot \sum_{i \in I^K} Q_{c,it} + \frac{\alpha_M \cdot P_{Mt}}{r_{c,t}} \cdot \sum_{i \in I^M} Q_{c,it} = K_{c,t} \qquad \text{for all } c \in C$$

$$(62)\qquad w_{c,t} \geq \frac{\varepsilon-1}{\varepsilon}\cdot Z_{Ht}\cdot p_{Ht} = \left(A_{Ht}\cdot P_{Ht}\right)^{\frac{1}{\mu}}\cdot\left(\sigma_H\cdot Y_t\right)^{\frac{\mu-1}{\mu}} = f_{Ht} \text{ for all } c\in C$$

$$(63)\qquad \left(\frac{w_{c,t}}{1-\alpha_M}\right)^{1-\alpha_M}\cdot\left(\frac{r_{c,t}}{\alpha_M}\right)^{\alpha_M} \geq \frac{\varepsilon-1}{\varepsilon}\cdot Z_{Mt}\cdot p_{Mt} = \left(A_{Mt}\cdot P_{Mt}\right)^{\frac{1}{\mu}}\cdot\left(\sigma_M\cdot Y_t\right)^{\frac{\mu-1}{\mu}} = \phi_M\cdot f_{Mt} \text{ for all } c\in C$$

$$(64)\qquad r_{c,t} \geq \frac{\varepsilon-1}{\varepsilon}\cdot Z_{Kt}\cdot p_{Kt} = \left(A_{Kt}\cdot P_{Kt}\right)^{\frac{1}{\mu}}\cdot\left(\sigma_K\cdot Y_t\right)^{\frac{\mu-1}{\mu}} = f_{Kt} \quad \text{for all } c\in C$$

Unlike the previous example, we find now that regions with extreme factor proportions have diversified production structures, while regions with intermediate factor proportions have specialized production structures. These sets of regions are now given by: [42]

$$(65)\qquad C_{Kt} = \left\{ c\in C \,\middle|\, \frac{H_{c,t}}{K_{c,t}} \leq \frac{1-\alpha_M}{\alpha_M}\cdot\left(\frac{f_{Kt}}{f_{Mt}}\right)^{\frac{1}{1-\alpha_M}} \right\}$$

$$(66)\qquad C_{Ht} = \left\{ c\in C \,\middle|\, \frac{H_{c,t}}{K_{c,t}} \geq \frac{1-\alpha_M}{\alpha_M}\cdot\left(\frac{f_{Mt}}{f_{Ht}}\right)^{\frac{1}{\alpha_M}} \right\}$$

$$(67)\qquad C_{Mt} = \left\{ c\in C \,\middle|\, \frac{1-\alpha_M}{\alpha_M}\cdot\left(\frac{f_{Kt}}{f_{Mt}}\right)^{\frac{1}{1-\alpha_M}} < \frac{H_{c,t}}{K_{c,t}} < \frac{1-\alpha_M}{\alpha_M}\cdot\left(\frac{f_{Mt}}{f_{Ht}}\right)^{\frac{1}{\alpha_M}} \right\}$$

Regions in C_{Kt} (C_{Ht}) produce in the M-industries and the K-industries (H-industries), while regions in C_{Mt} produce only in M-industries. Factor prices are determined as follows:

[42] Note that the sets C_{Kt} and C_{Ht} never intersect in world equilibrium. Assume the opposite, then it follows that equilibrium input prices must satisfy $f_{Mt}<(f_{Ht})^{1-\alpha_M}\cdot(f_{Kt})^{\alpha_M}$. But if this inequality held nobody would produce in M-industries and markets for the products of these industries would not clear.

(68) $$w_{c,t}=\begin{cases} f_{Mt}^{\frac{1}{1-\alpha_M}}\cdot f_{Ht}^{\frac{-\alpha_M}{1-\alpha_M}} & \text{if } c\in C_{Kt} \\ (1-\alpha_M)\cdot\phi_M\cdot f_{Mt}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{-\alpha_M} & \text{if } c\in C_{Mt} \\ f_{Ht} & \text{if } c\in C_{Ht} \end{cases}$$

(69) $$r_{c,t}=\begin{cases} f_{Kt} & \text{if } c\in C_{Kt} \\ \alpha_M\cdot\phi_M\cdot f_{Mt}\cdot\left(\dfrac{H_{c,t}}{K_{c,t}}\right)^{1-\alpha_M} & \text{if } c\in C_{Mt} \\ f_{Mt}^{\frac{1}{\alpha_M}}\cdot f_{Ht}^{\frac{\alpha_M-1}{\alpha_M}} & \text{if } c\in C_{Ht} \end{cases}$$

Once again, the wage is continuous and weakly declining on the human to physical capital ratio, while the rental is also continuous but increasing on this same ratio. But now these relationships exhibit at most two "flats", one for each set of human to physical capital ratios that defines a cone of diversification. Regional differences in factor prices reflect again factor abundance in the usual way. This world economy contains at most two cones of diversification.[43] Regions with extreme factor proportions belong to one of them, while regions with intermediate factor proportions do not. The middle panel of Figure 10 shows how the wage-rental varies with a region's ratio of human to physical capital.

It is straightforward to compute the analogues of Equation (58)-(59) for this example and check that the mapping from factor endowments to incomes and capital accumulation is also linear within the cones and takes the Cobb-Douglas form outside of them. The picture of the growth process that comes out of this example is therefore very similar to the on in Example 2.4.1.

[43] I say "at most" because it is also possible that $\underline{R}_{Mt}=\overline{R}_{Mt}$, in which the case there would be a single cone. Cuñat and Mafezzoli [2004a] analyze a similar model under the assumption that none of the regions of the world have specialized production structures, i.e. $C_{Mt}=\varnothing$.

Examples 2.4.1 and 2.4.2 can be generalized by introducing further industries with different factor intensities. As we do so, the potential number of cones increases. But the overall picture remains the same. The world economy sorts itself out in a series of cones of diversification. The bottom panel of Figure 10 depicts a case with multiple cones of diversification.[44] Small regional differences in factor proportions lead to structural transformation within cones, but to factor deepening outside them. Large regional differences in factor proportions might span one or more cones and therefore lead to a mix of structural transformation and factor deepening. Therefore, the world of diversification cones can be seen as being somewhere in between the world of factor-price-equalization and the world of autarky.[45]

In the light of these results, we must slightly revise our earlier discussion of the effects of globalization on the source of income differences. As in the world of factor-price equalization, differences in domestic productivities cannot be a source of income differences and relative incomes are homogeneous of degree one with respect to factor endowments. But unlike the world of factor-price-equalization, the elasticity of substitution between domestic factors is no longer infinity but instead lies somewhere between one (outside cones) and infinity (within cones). As mentioned, this elasticity measures the relative importance of structural transformation and factor deepening as a means to accommodate regional differences in factor proportions. And this relative importance in turn depends on various factors, most notably how dispersed are factor intensities across industries. Two extreme examples make this point forcefully. If the dispersion in industry factor intensities is extreme, i.e. $\alpha_i \in \{0,1\}$ for all $i \in I$, then regional differences in factor proportions always lead to structural transformation and the world income distribution is given by

---

[44] Dornbusch, Fischer and Samuelson [1980] develop a similar model with a continuum of goods that vary in their factor intensity, although they do not specifically study the formation of cones.
[45] In pure Heckscher-Ohlin models, Deardorff [2001] and Cuñat and Maffezzoli [2004a] generate "club convergence". Stiglitz [1970] and Devereux and Shi [1991] are examples where cones of diversification establish due to inherently different time-preferences and incomes diverge. Oniki and Uzawa [1965] analyze conditions for diversification cones in two-sector model.

Equations (39)-(40).[46] If the dispersion in industry factor intensities is instead negligible, i.e. $\alpha_i=\alpha$ for all $i\in I$, then regional differences in factor proportions always lead to factor deepening and the world income distribution is given by:[47]

$$(70) \qquad Y_{c,t} = A_t \cdot H_{c,t}^{1-\alpha} \cdot K_{c,t}^{\alpha}$$

As in the world of autarky, the elasticity of substitution across factors is one (see Equation (33)). But unlike the world of autarky, regional differences in industry productivities and market size play no role in explaining regional income differences.

We do not need to revise however our earlier discussion of how globalization affects the dynamic responses to region-specific shocks. In this respect, the world with diversification cones offers the same insights as the world of factor-price-equalization. Region-specific shocks to savings and human capital have positive effects that spill over to other regions, while shocks to industry productivities only have effects if they push outwards the world productivity frontier. Economic growth is positively transmitted across regions through changes in goods prices and this keeps the world income distribution stable. In fact, this force towards stability is further reinforced in regions that are outside a cone by the existence of diminishing returns in production.

We conclude therefore that violations to requirement R2 do not alter much the picture came out of the factor-price-equalization world. Surely the geographical mismatch between different factor endowments implied by these violations might generate large inefficiencies that, in turn, might lead to sizeable regional differences in factor prices. Therefore, there might be important quantitative differences between a world with many diversification cones and the world of factor-price-equalization. But the qualitative properties of the growth process of these two worlds remain relatively close to each other, and far away from those of the world of autarky.

[46] This is the model used by Ventura [1997].

[47] One of many ways to find this result is as the appropriate limiting case of Examples 2.4.1 or 2.4.2.

2.5 Limits to structural transformation (II): industry productivities

Consider next worlds where requirement R1 is either binding or fails. Regions with few high-productivity industries might find that even if they concentrate all of their production in those industries, they cannot employ all of their factors and produce the same quantities as the integrated economy. These regions are therefore forced to exceed the production of the integrated economy in those industries and/or move into low-productivity industries. Whatever the case, this requires these regions to offer low factor prices to employ all of their factors. This situation can be aptly described as a geographical mismatch between industry productivities and factor endowments.

To make further progress, it is necessary to be more explicit about why and how industry productivities differ across regions. The first example considers the case in which regional differences in productivities take the popular factor-augmenting form:

Example 2.5.1: Consider a world where $Z_{c,it} = \pi_{c,Ht}^{1-\alpha_i} \cdot \pi_{c,Kt}^{\alpha_i}$ for all $i \in I$ and all $c \in C$; with $\sum_{c\in C} \pi_{c,Ht} \cdot \frac{H_{c,t}}{H_t} = 1$ and $\sum_{c\in C} \pi_{c,Kt} \cdot \frac{K_{c,t}}{K_t} = 1$. As usual, $\pi_{c,Ht}$ and $\pi_{c,Kt}$ are interpreted as labor- and capital-augmenting productivity differences, The world productivity frontier is given by $Z_{it} = \max_c \{\pi_{c,Ht}^{1-\alpha_i} \cdot \pi_{c,Kt}^{\alpha_i}\}$. In the integrated economy, industries would be located exclusively in the regions that are in this frontier. The set FPE_t is "small" and, except for a few very special or knife-edge cases, factor-price equalization is not possible and requirement R1 fails. [48]

[48] Take, for instance, the case of two regions and two industries. If one region has the highest productivity in both industries the only factor distribution that leads to factor-price equalization is the one

To understand the logic of this world, it is useful to follow the usual procedure of re-normalizing the model in terms of "efficiency" or "productivity-equivalent" factor units. That is, we can pretend that regional factor endowments are given by $\hat{H}_{c,t} = \pi_{c,Ht} \cdot H_{c,t}$ and $\hat{K}_{c,t} = \pi_{c,Kt} \cdot K_{c,t}$ for all $c \in C$; and that industry productivities are identical across regions, i.e. $\hat{Z}_{c,it} = 1$ for all $i \in I$ and all $c \in C$. Then, productivity-adjusted factor prices are given by $\hat{w}_{c,t} = \frac{w_{c,t}}{\pi_{c,Ht}}$ and $\hat{r}_{c,t} = \frac{r_{c,t}}{\pi_{c,Kt}}$. The key observation is that the re-normalized model is formally equivalent to the model of the previous section.[49] Therefore, all the results we obtained in the previous sections regarding the cross-section of factor prices also apply here to productivity-adjusted factor prices, i.e. $\hat{w}_{c,t}$ and $\hat{r}_{c,t}$; but not to factor prices as usually measured, i.e. $w_{c,t}$ and $r_{c,t}$.[50]

As the worlds of the previous section, this world economy sorts itself out in a series of cones of diversification. All regions within a cone have the same productivity-adjusted factor prices, although possibly different factor prices as usually measured. Regional differences in productivity-adjusted factor proportions lead to structural transformation within cones, and to factor deepening across them. When all regions are located within a single cone, we have the conditional factor-price-equalization result emphasized by Trefler [1993]. That is, regional differences in factor prices reflect only differences in factor-augmenting productivities and are not related to differences in productivity-adjusted factor proportions.

in which all factors are located in this region. If instead each region has the highest productivity in a different industry, the only factor distribution that leads to factor-price equalization is the one in which each region receives the exact quantity of factors that its high-productivity industry uses in the integrated economy.

[49] The re-normalized model is a bit less general than the model of the previous section since it does not display regional differences in industry productivities. We could (trivially) generalize this example to allow for regional differences in industry productivities, but keeping the assumption that requirement R1 is not binding in the re-normalized model.

[50] For instance, Equations (56)-(57) describe the productivity-adjusted factor if we further assume that the world economy contains two types of industries as in Example 2.4.1. Similarly, Equations (68)-(69) describe productivity-adjusted factor prices if we instead assume that the world economy contains three types of industries as in Example 2.4.2.

Although the presence of factor-augmenting productivity differences does not alter much the formal or mathematical structure of the model, it has important implications for the question of why some regions are richer than others. Unlike the worlds of section 2.4, we now have that productivity differences become a source of income differences across countries. For instance, if all regions belong to a single cone of diversification we have the following counterpart to Equation (40):

$$(71) \qquad \frac{Y_{c,t}}{Y_t} = (1-\alpha)\cdot\frac{\pi_{c,Ht}\cdot H_{c,t}}{H_t} + \alpha\cdot\frac{\pi_{c,Kt}\cdot K_{c,t}}{K_t} \qquad \text{for all } c \in C$$

Alternatively, if all the industries in the world have the same factor intensity we have the following counterpart of Equation (70):

$$(72) \qquad Y_{c,t} = A_{c,t}\cdot H_{c,t}^{1-\alpha}\cdot K_{c,t}^{\alpha}$$

where $A_{c,t} = \hat{A}_t \cdot \pi_{c,Ht}^{1-\alpha} \cdot \pi_{c,Kt}^{\alpha}$.[51] The inability of the world economy to match best technologies with appropriate factors moves us a step closer to the world of autarky, since regional productivities now affect regional incomes. Moreover, since now the world operates below its productivity frontier shocks to regional factor productivities have effects even if they do not push this frontier. Note however that, as in the worlds of section 2.4, the elasticity of substitution between domestic factors still lies somewhere between one (outside cones) and infinity (within cones); and relative incomes are homogeneous of degree one with factor endowments.

The rest of the picture of the growth process that comes out of this world remains close to the world of factor-price-equalization. Region-specific shocks to savings and human capital have positive effects that spill over to other regions. Economic growth is positively transmitted across regions through goods prices and

[51] This model therefore provides an alternative theoretical foundation for the work of Mankiw, Romer and Weil [1992], Hall and Jones [1999] and Klenow and Rodríguez-Clare [1997].

this keeps the world income distribution stable. If the conditional version of the factor-price-equalization theorem does not hold, regions outside the cones experience diminishing returns and this reinforces the effects of changes in product prices on the stability of the world income distribution.

Assuming that regional productivity differences take the factor-augmenting form discussed in example 2.5.1 is popular because it yields tractable models. But the factor-augmenting view of productivity differences hides some interesting effects of trade on the world income distribution and its stability. One reason is that, in the world of factor-augmenting productivity differences, comparative advantage is still determined solely by regional differences in factor proportions, albeit productivity-adjusted ones. The next example provides a dramatic illustration of how regional differences in industry productivities could determine comparative advantage, and how this brings about a new effect of trade on the world income distribution:

Example 2.5.2: Consider a world with many industries and regions. Assume that $Z_{c,it} = 1$ if $i \in I^*_{c,t}$, and $Z_{c,it} = 0$ if $i \notin I^*_{c,t}$; where $I^*_{c,t}$ for all $c \in C$ constitutes a partition of I: $\bigcup_{c \in C} I^*_{c,t} = I$ and $I^*_{c,t} \cap I^*_{c',t} = \varnothing$ for all $c \in C$ and $c' \in C$. Assume also that $I^*_{c,t} \neq \varnothing$ for all $c \in C$. That is, each region knows how to produce a disjoint subset of goods. Since only one region knows how to produce each good, the corresponding industry is located in that region. That is, $I_{c,t} = I^*_{c,t}$ for all $c \in C$, regardless of the factor distribution. In this world, comparative advantage is driven solely by regional differences in industry productivities, and differences in factor proportions play no role. In this example, requirement R1 does not fail but it is binding, except for a few very special and knife-edge cases.

A bit of straightforward algebra shows that production and factor allocations are given as follows:

(73) $$Y_{c,t} = \sum_{i \in I^*_{c,t}} \phi_i \cdot f_{it} \cdot H_{c,it}^{1-\alpha_i} \cdot K_{c,it}^{\alpha_i} \quad \text{for all } c \in C$$

(74) $$H_{c,it} = \frac{\sigma_i}{\sum_{i' \in I^*_{c,t}} \sigma_{i'}} \cdot \frac{1-\alpha_i}{\sum_{i' \in I^*_{c,t}} \sigma_{i'} \cdot (1-\alpha_{i'})} \cdot H_{c,t} \text{ if } i \in I^*_{c,t}\text{; and } H_{c,it} = 0 \text{ if } i \notin I^*_{c,t}$$

(75) $$K_{c,it} = \frac{\sigma_i}{\sum_{i' \in I^*_{c,t}} \sigma_{i'}} \cdot \frac{\alpha_i}{\sum_{i' \in I^*_{c,t}} \sigma_{i'} \cdot \alpha_{i'}} \cdot K_{c,t} \text{ if } i \in I^*_{c,t}\text{; and } K_{c,it} = 0 \text{ if } i \notin I^*_{c,t}$$

where, as usual by now, $\phi_i = (1-\alpha_i)^{\alpha_i - 1} \cdot \alpha_i^{-\alpha_i}$ and $f_{it} = \min_{c \in C} \left\{ w_{c,t}^{1-\alpha_i} \cdot r_{c,t}^{\alpha_i} \right\}$ for all $i \in I$.

Equation (73) describes the world income distribution as a function of factor allocations and goods prices, while Equations (74)-(75) provide the equilibrium factor allocations as a function of aggregate factor endowments. By substituting Equations (74)-(75) into Equation (73), we obtain the world income distribution as a function of factor endowments and input prices.[52] It is immediate to show that the elasticity of substitution between human and physical capital is between one and infinity; that regional differences in industry productivities affect regional differences in income; and that the world income distribution is homogeneous of degree one with respect to factor endowments.

These results are obtained from a relationship between incomes, factor endowments and industry productivities that holds constant input prices. Once we substitute input prices into this relationship, we find that the world income distribution is given by:

(76) $$\frac{Y_{c,t}}{Y_t} = \sum_{i \in I^*_{c,t}} \sigma_i \quad \text{for all } c \in C$$

[52] This relationship is formally analogous, for instance, to Equation (58) in Example 2.4.1.

Equation (76) states that the share of world income of each region equals that share of world spending on the industries located in the region, and it does not depend on domestic factor endowments. What is going on? Assume a region has a ratio of human to physical capital λ times higher than average. Since the region is producing a fixed set of goods, it is forced to operate with a ratio that is λ times higher than average, and this requires a wage-rental ratio that is λ^{-1} higher than average. Therefore the elasticity of substitution between human and physical capital in production is one. What is different here is that relative incomes are now homogeneous of degree zero with respect to factor endowments. Assume a region's human and physical capitals are both λ times average. Since production is homogenous of degree one with factor endowments, its production of all industries is λ times average. But since the country faces a demand for its products with price-elasticity equal to one, the prices of its products are λ^{-1} times average. As a result, the income of the region is just average, despite its factor endowments being λ times average.

So what should we conclude about the degree of homogeneity of relative incomes with respect to factor endowments? As Equations (73)-(75) and (76) show, in empirical applications it will depend on whether we are holding goods prices constant or not. If we are holding these prices constant, then relative incomes are homogeneous of degree one in factor endowments. If we are not holding goods prices constant, then the degree of homogeneity of relative incomes with respect of factor endowments lies between zero and one. In this example, this degree of homogeneity is zero because regional differences in factor endowments are absorbed by regional variation in the quantities produced of each input. In Examples 2.4.1, 2.4.2 and 2.5.1, this degree of homogeneity was one because regional differences in factor endowments were absorbed by regional variation in the number of input varieties produced. The next example, inspired by Dornbusch, Fischer and Samuelson [1977], neatly clarifies this point by showing an intermediate world where both margins are at work.

Example 2.5.3: Consider a world with two regions, C={N,S}; and a continuum of industries, I=[0,1]. Assume all industries have the same factor intensity, $\alpha_i=\alpha$ for all $i \in I$. For simplicity, let also $\varepsilon_i=\varepsilon$ and $\beta_i=\beta$ for all $i \in I$. It follows immediately that: [53]

(77) $$Y_{c,t} = \phi \cdot f_{c,t} \cdot H_{c,t}^{1-\alpha} \cdot K_{c,t}^{\alpha} \qquad \text{for all } c \in C$$

where $\phi = (1-\alpha)^{\alpha-1} \cdot \alpha^{-\alpha}$; and $f_{c,t}$ is a measure of factor costs of region c, i.e. $f_{c,t} = w_{c,t}^{1-\alpha} \cdot r_{c,t}^{\alpha}$ for all $c \in C$. To characterize the world income distribution in this world, we need to determine factor costs. Equation (77) is akin to Equation (58) or Equations (73)-(75) in the sense that it shows the world income distribution as a function of factor endowments and input prices. Not surprisingly, these relative incomes are homogeneous of degree one with respect to factor endowments. The next step is to determine input prices and substitute them into Equation (77).

Define $T_i \equiv \frac{Z_{N,it}}{Z_{S,it}}$ for all $i \in I$ as the industry productivity of North relative to South. Then, assign indices or order goods so that T_i is non-increasing in i. Note that T_i might be neither continuous nor invertible.[54] It follows from this ordering that $I_{N,t} \equiv \left\{ i \in I \;\middle|\; \frac{f_{N,t}}{f_{S,t}} \leq T_i \right\}$ and $I_{S,t} \equiv \left\{ i \in I \;\middle|\; \frac{f_{N,t}}{f_{S,t}} \geq T_i \right\}$. That is, North (or N) specializes on low-index industries while South (or S) specializes in high-index industries. The cutoff industry, i*, is determined as follows:[55]

---

[53] This follows directly from the observation that the share of human and physical capital in income are 1-α and α, respectively.

[54] This ranking can vary over time, but this does not play any role here. Without loss of generality, the reader can focus on the case in which the ranking is time-invariant.

[55] If T_i is not invertible in the region of interest, this condition determines a set of candidate values for i*.

(78) $$\frac{f_{N,t}}{f_{S,t}} = T_{i^*}$$

Let X_i be world share of spending on all industries with indices equal or lower than i, that is, $X_i \equiv \int_0^i \sigma_j \cdot dj$. Note that X_i is non-decreasing in i, and takes values zero and one for i=0 and i=1. It follows from this definition that $Y_{N,t}=X_{i^*}\cdot(Y_{N,t}+Y_{S,t})$ and, using Equation (78), this can be rewritten as follows:

(79) $$\frac{f_{N,t}}{f_{S,t}} = \frac{X_{i^*}}{1-X_{i^*}} \cdot \left(\frac{H_{S,t}}{H_{N,t}}\right)^{1-\alpha} \cdot \left(\frac{K_{S,t}}{K_{N,t}}\right)^{\alpha}$$

Equations (78)-(79) jointly determine the pattern of production and trade (i*) and relative factor costs ($f_{N,t}/f_{S,t}$) as a function of spending patterns, industry productivities and factor endowments. Finally, we can use the numeraire rule in Equation (4) to find that:

(80) $$Y_t = \sum_{c\in\{N,S\}} \phi \cdot f_{c,t} \cdot H_{c,t}^{1-\alpha} \cdot K_{c,t}^{\alpha} = (\varepsilon-1)\cdot \exp\left\{\int_0^{i^*} Z_{N,it} \cdot \sigma_i \cdot di + \int_{i^*}^{1} Z_{S,it} \cdot \sigma_i \cdot di\right\}$$

Having already found the pattern of production and trade (i*) and relative factor costs ($f_{N,t}/f_{S,t}$), Equation (80) can then be used to determine absolute factor costs.

This world is somewhat different form the ones we have seen so far in that we have only two regions. To think about the effects of factor endowments on relative incomes, I consider next a situation in which both regions have symmetric technologies and differ in that North's factor endowments are λ (>1) times larger than South's.[56] Figure 11 depicts this world. The AA and BB lines represent

[56] By symmetric technologies, I mean that if there exists an industry i such that $T_i=\tau$ then there also exists another industry i' such that $T_{i'}=1/\tau$ and $\alpha_i=\alpha_{i'}$, $\beta_i=\beta_{i'}$, $\varepsilon_i=\varepsilon_{i'}$ and $\sigma_i=\sigma_{i'}$.

Equations (78) and (79), respectively. The AA line is non-increasing because T_i is non-increasing in i, while the BB line is non-decreasing because X_i is non-decreasing in i. The existence of a unique crossing point follows since the BB line takes value zero at i=0 and slopes upward towards infinity at i=1.

The top panel of Figure 11 shows the case in which T_i is flat. This case corresponds to a world in which differences in industry productivities are minimal or irrelevant at the margin as in Examples 2.4.1, 2.4.2 and 2.5.1. This allows North to employ its larger factor endowments by producing a larger number of varieties than South. Factor costs are the same in both regions and, as a result, North's income is λ times South's. Relative incomes (after substituting in goods prices) are homogenous of degree one on factor endowments.

The middle panel of Figure 11 shows the opposite case in which T_i is vertical. This case corresponds to a world in which differences in industry productivities are extreme as in Examples 2.5.2. North is forced to employ its larger factor endowments by producing a higher quantity of each of its varieties. Factor costs in North are λ^{-1} times those of South and, as a result, North's income equals that of South. Relative incomes (after substituting in goods prices) are homogenous of degree zero on factor endowments.

The bottom panel shows the intermediate case in which T_i is neither flat nor vertical. Since the slope reflects how strong are differences in industry productivities, we are somewhere in between the two extreme examples considered up to now. North employs its larger factor endowments by producing a larger number of varieties and also a larger quantity of each of them. Factor costs in North are somewhere between λ^{-1} and one times those of South. The degree of homogeneity of relative incomes (after substituting in goods prices) on factor endowments is therefore somewhere between zero and one.

It is possible to generalize Example 2.5.3 in a variety of directions. For instance, one could allow for industry variation in factor intensities and many regions.[57] This is important in empirical applications, of course. But the central message remains. The effects of factor endowments on relative incomes depend on regional differences in industry productivities. If these differences are small, regions with larger factor endowments absorb them mostly through structural transformation: not changing much their production in existing industries and moving into new industries where the region's productivity relative to the rest of the world is similar to existing ones. If differences in industry productivities are large, regions with larger factor endowments absorb them by productivity deepening: substantially increasing their production in existing industries and/or moving into industries where the region's productivity relative to the rest of the world is substantially lower than in existing ones.

One can conclude from this discussion that differences in industry productivities create another force for diminishing returns to physical capital accumulation. As physical capital is accumulated, quantities produced increase and the terms of trade worsen. The result is a reduction in factor prices that lowers wages, savings and capital accumulation. This is a central aspect of the growth process in a world of interdependent regions generates a force towards the stability of the world income distribution.[58]

I argued at the end of section 1 that, if globalization leads to the integrated economy, there is a powerful prescription for economic development: open up and integrate into the world economy. This allows regions to benefit from higher productivity, improved factor allocation and increased market size. Not much has changed here. Naturally, if factor prices are equalized the effects are literally the same as in section 1 since the globalization leads to the integrated economy. If

[57] See Wilson [1980], Eaton and Kortum [1999, 2000], Matsuyama [2000] and Alvarez and Lucas [2004].

[58] See Acemoglu and Ventura [2002] on this point.

factor prices are not equalized, the world economy operates with a lower productivity and a worse factor allocation than the integrated economy. This also means that the size of the world economy will be smaller than that of the integrated economy. As a result, all the benefits from globalization are smaller in the worlds of sub-sections 2.4 and 2.5 than in the world of factor-price equalization. But it is still relatively straightforward to see that coupled with an appropriate transfer scheme globalization constitutes a Pareto improvement for the world economy. Moreover, since all regions gain from trade there exist Pareto-improving transfer schemes that do not require inter-regional transfers.[59] Therefore, the prescription for development remains the same: open up and integrate into the world economy.

We have traveled much already, and the global view of economic growth is starting to take shape. This view is more realistic and rich in details than the views that came out from either the world of autarky or the integrated economy. Despite this progress, we should not rest here yet. We have assumed so far that globalization eliminates all impediments to goods trade. This is obviously an unrealistic assumption. Is it also a crucial one?

3. Transport costs and market size

Despite the already large and growing importance of international trade, there are some important areas of economic activity where the degree of market integration is still relatively low. Surely the clearest case in point is the service sector.[60] As the textbook example of a haircut suggests, many services are

[59] How do we know that all regions have non-negative gains from globalization? Since regions have the choice of not trading, it is therefore possible to achieve the level of income and welfare of the world of autarky after globalization. Realizing these gains might require implementing an appropriate tax-subsidy scheme, though.

[60] In industrial economies, the service sector accounts for more than two thirds of production but only for about one fifth of exports and imports. Moreover, most trade in services is concentrated in activities related to transportation and travel even though these activities only constitute a small component of overall services production.

inherently more difficult to transport than agricultural and manufacturing products. Services also tend to be more vulnerable to various governmental barriers to trade, such as professional licensing requirements that discriminate against foreigners, domestic content requirements in public procurement, or poor protection of intellectual property rights.[61] In addition, there are important examples of weak market integration that go beyond the service sector. Trade in some agricultural and manufacturing products is also severely restricted as a result of protectionist practices in industrial countries.

The goal of this section is to study the effects on the growth process of partial segmentation in goods markets. The new model of globalization that I shall adopt here is as follows: at date t=0 the costs of transporting some (but not all) goods across regions suddenly fall from prohibitive to negligible. In particular, I partition the set of all industries into the sets of tradable and nontradable industries, i.e. T_t and N_t such that $T_t \cup N_t = I$ and $T_t \cap N_t = \varnothing$. The costs of transporting intermediate inputs and final goods fall from prohibitive to negligible at t=0 if $i \in T_t$. But even after t=0, the costs of transporting either the intermediate inputs, or the final goods, or both remain prohibitive if $i \in N_t$.[62] We keep assuming that the costs of transporting factors across regions remain prohibitive after t=0, and that international trade in assets is not possible. Naturally, the model analyzed in section 2 (and formally described in section 2.3) obtains as the special case of this model in which $T_t = I$ and $N_t = \varnothing$ for t≥0.

[61] There are also signs that this is changing rapidly. Advances in telecommunications technology, the appearance of e-commerce and the development of new and standardized software have all opened up the possibility of trading a wider range of services. Recent multilateral trade negotiations and the process of European integration have also led to the dismantling of various non-tariff barriers to service trade.

[62] The most popular alternative to this model is the "iceberg" cost model whereby all goods are subject to the same proportional transport cost. In particular, a quantity τ (>1) of a good must be shipped from source to ensure that one unit of it arrives to destination. The rest "melts" away in transit. See Matsuyama [2004] for yet another model of transport costs.

A central aspect of the analysis turns out to be whether transport costs apply only to final goods, to intermediate inputs, or to both. Section 3.1 presents the case in which transport costs apply only to final goods. This model neatly generalizes the results obtained in the previous section. Section 3.2 studies the case in which transport costs apply only to intermediate inputs. This gives rise to agglomeration effects that can have a large and somewhat unexpected impact on the world income distribution. Section 3.3 analyzes the case in which transport costs apply to both final goods and intermediate inputs. The interaction between the two types of frictions brings about a new perspective on the role of local markets.

3.1 Nontraded goods and the cost of living

Consider next a world where some final goods are not tradable, although the intermediate inputs required to produce them are always tradable. In particular, the costs of trading intermediate inputs are negligible for all $i \in I$; and the costs of transporting final goods are negligible if $i \in T_t$ but prohibitive if $i \in N_t$. Since prices of final goods can differ across regions, a novel feature of this model is that regions will have different price levels.

I must now revise the formal description of the model. While regional analogues of Equations (1) and (2) continue to apply, one must now recognize that final goods prices in nontradable industries might differ across regions. As a result, the price of consumption and investment will vary across them even if Equation (3) describing spending patterns still applies to all regions. Therefore, we must write the analogues of Equations (1) and (2) as follows:

$$(81) \qquad K_{c,t+1} = S_{c,t} \cdot \frac{w_{c,t}}{P_{c,t}} \cdot H_{c,t}$$

(82) $$C_{c,t} = (1 - S_{c,t}) \cdot \frac{w_{c,t}}{P_{c,t}} \cdot H_{c,t} + \frac{r_{c,t}}{P_{c,t}} \cdot K_{c,t}$$

where $P_{c,t}$ is the price level (or cost of living) of region c, i.e. $P_{c,t} = \prod_{i \in I} \left(\frac{P_{c,it}}{\sigma_i} \right)^{\sigma_i}$ for all $c \in C$. A natural choice of numeraire now is the ideal price index for tradable industries:

(83) $$1 = \prod_{i \in T_t} \left(\frac{P_{it}}{\sigma_i} \right)^{\sigma_i}$$

Equation (83) replaces Equation (4). The latter obtains as the special case of the former in which all goods are tradable, i.e. $T_t = I$ and $N_t = \varnothing$. An implication of this choice of numeraire is that the price level of region c is equal to the ideal price index of its nontradable industries:

(84) $$P_{c,t} = \prod_{i \in N_t} \left(\frac{P_{c,it}}{\sigma_i} \right)^{\sigma_i} \qquad \text{for all } c \in C$$

Since now price levels differ across regions it is necessary to distinguish between two concepts of income and factor prices: (1) market-based incomes and factor prices, i.e. $Y_{c,t}$, $w_{c,t}$ and $r_{c,t}$; and (2) real or PPP-adjusted incomes and factor prices, i.e. $Y_{c,t}/P_{c,t}$, $w_{c,t}/P_{c,t}$ and $r_{c,t}/P_{c,t}$. Whenever there is no risk of confusion, I shall refer to the former simply as income and factor prices, and to the latter as real income and real factor prices. As before, Equations (5)-(6) describing technology apply to all regions, with the corresponding factor prices and industry productivities.

After globalization, producers of intermediate inputs in all industries and producers of final goods in tradable industries face a global market and Equations

(7)-(10) describing pricing policies, input demands and the free-entry condition therefore apply only to those regions where the lowest-cost producers are located. But even after globalization, producers of final goods in nontradable industries remain sheltered from foreign competition, and Equations (7)-(8) apply to all regions and not only to the lowest-cost ones. Thus, Equation (44) no longer applies to the producers of final goods in nontradable industries (Equation (43) still stands as a definition, though).

Market clearing conditions are also affected by the presence of transport costs. While Equations (45)-(46) describing market clearing in regional factor markets still apply, Equation (11) describing market clearing in global markets for final goods applies only to tradable industries. In nontradable industries, Equation (11) must be replaced by analogue conditions imposing market clearing in each regional market:

(85) $$P_{c,it} \cdot Q_{c,it} = E_{c,it} \qquad \text{for all } i \in N_t \text{ and } c \in C$$

This completes the formal description of the model. For any admissible set of capital stocks, i.e. $K_{c,0}$ for all $c \in C$; sequences for the vectors of savings, human capital and industry productivities, i.e. $S_{c,t}$, $H_{c,t}$, and $A_{c,it}$ for all $c \in C$ and for all $i \in I$; and a sequence for the set N_t (or T_t); an equilibrium of the world economy after globalization consists of a sequence of prices and quantities such that the equations listed above hold at all dates and states of nature. Although there might be multiple geographical patterns of production that are consistent with world equilibrium, the assumptions made ensure that prices and world aggregates are uniquely determined.

The best way to start the analysis is by asking again whether the assumed trade restrictions matter at all. That is, to ask whether restricting factor mobility and now goods trade impede the world to achieve the level of efficiency of the integrated economy. Re-define the set FPE_t now as follows:

(86)
$$FPE_t \equiv \left\{ d_t \in D_t \,\middle|\, \begin{array}{l} \exists x_{c,it}(m) \geq 0,\ x^F_{c,it} \geq 0 \text{ with } \sum_{c \in C} x_{c,it}(m) = 1,\ \sum_{c \in C} x^F_{c,it} = 1 \text{ and} \\ x_{c,it} = (1-\beta_i) \cdot x^F_{c,it} + \frac{\beta_i}{M_{it}} \cdot \int_0^{M_{it}} x_{c,it}(m) \cdot dm; \text{ such that :} \\ \text{(R1) } x_{c,it} = 0 \text{ if } Z_{c,it} < \max_{c \in C}\{Z_{it}\}; \\ \text{(R2) } H_{c,t} = \sum_{i \in I} x_{c,it} \cdot H_{it} \text{ and } K_{c,t} = \sum_{i \in I} x_{c,it} \cdot K_{it}; \\ \text{(R3) } x_{c,it}(m) \in \{0,1\} \text{ for all } m \in [0, M_{it}] \text{ and } i \in I; \text{ and} \\ \text{(R4) } x^F_{c,it} \geq (1-\beta_i) \cdot \frac{Y^{IE}_{c,t}}{Y^{IE}_t} \text{ if } i \in N_t \end{array} \right\}$$

Comparing Definitions (36) and (86), we observe that the latter contains an additional requirement: each region should be able to produce all the final goods used for its own consumption and investment in nontradable industries. This additional restriction is a direct consequence of transport costs. The presence of this additional restriction reduces the size of FPE_t. In fact, it is now even possible that $FPE_t = \emptyset$. For instance, assume regional differences in industry productivities are such that there exists no region that has the highest possible productivity in all nontradable industries simultaneously. Then, it is not possible to replicate the integrated economy.[63]

If $d_t \in FPE_t$, factor prices are equalized across regions and the world economy operates with the same efficiency as the integrated economy despite factor immobility and goods market segmentation. In this case, the world economy

[63] That one or more regions with the highest possible productivity in all nontradable industries exist is a necessary but not sufficient condition for $FPE_t \neq \emptyset$. Since factor-price equalization requires that all factors be located in these regions, it is also necessary that at least one of these regions have the highest possible productivity for each tradable industry.

behaves exactly as the world of factor-price equalization of section 2.2.[64] If $d_t \notin FPE_t$, the world economy cannot operate at the same level efficiency as the integrated economy. As a result, either market-based factor prices, or real factor prices, or both differ across regions. But even in this case the behavior of the world economy does not depart much from what we observed in the worlds of section 2. To see this, define $H^T_{c,t}$ and $K^T_{c,t}$ as the factor endowments devoted to the production of tradable goods, i.e. all intermediate inputs and the final goods of tradable industries. Straightforward algebra shows that:[65]

$$\text{(87)} \quad H^T_{c,t} = \max\left\{0, H_{c,t} \cdot \left(1 - \sum_{i \in N_t} (1-\beta_i)\cdot(1-\alpha_i)\cdot\sigma_i\right) - K_{c,t} \cdot \left(\frac{w_{c,t}}{r_{c,t}}\right)^{-1} \cdot \sum_{i \in N_t} (1-\beta_i)\cdot(1-\alpha_i)\cdot\sigma_i\right\} \text{ for all } c \in C$$

$$\text{(88)} \quad K^T_{c,t} = \max\left\{0, K_{c,t} \cdot \left(1 - \sum_{i \in N_t} (1-\beta_i)\cdot\alpha_i\cdot\sigma_i\right) - H_{c,t} \cdot \frac{w_{c,t}}{r_{c,t}} \cdot \sum_{i \in N_t} (1-\beta_i)\cdot\alpha_i\cdot\sigma_i\right\} \text{ for all } c \in C$$

Equations (87)-(88) show the factor supplies that are left after subtracting from aggregate factor supplies the factors used in the production of final goods in nontradable industries. In the special case in which $N_t = \varnothing$, these factor supplies equal the aggregate factor supplies and are independent of factor prices. But in the general case, these factor supplies depend on factor prices in the usual way. The higher is the wage-rental, the lower is the human to physical capital ratio used for the production of final goods in nontradable industries and, as a result, the higher is the relative supply of human to physical capital that is left after production of final goods in nontradable industries.

---

[64] Even the price levels would be equalized across regions, i.e. $P_{c,t}=1$ for all $c \in C$. Note however that there is less indeterminacy regarding the patterns of production and trade, since nontradable final goods must now be produced in the same region where they are used for consumption or investment.

[65] To see this, note that the shares of human and physical capital devoted to producing the final good of the i[th] nontradable industry are $(1-\beta_i)\cdot(1-\alpha_i)$ and $(1-\beta_i)\cdot\alpha_i$. Add over industries and note that the share of spending in the i[th] industry is $\sigma_i \cdot Y_{c,t}$.

With Equations (87)-(88) at hand, it is straightforward to see that all the results in sections 2.4 and 2.5 regarding incomes and factor prices still go through in the presence of nontradable final goods. Take, for instance, Example 2.4.1. Equations (47)-(48) must be rewritten as follows:

$$(89) \qquad P_{it} = \sigma_i \cdot \frac{\prod_{i' \in T_t} \left(\sum_{c \in C} Q_{c,i't} \right)^{\frac{\sigma_i}{\sum_{i' \in T_t} \sigma_{i'}}}}{\sum_{c \in C} Q_{c,it}} \qquad \text{for all } i \in T_t$$

$$(90) \qquad Y_t = \prod_{i \in T_t} \left(\sum_{c \in C} Q_{c,it} \right)^{\frac{\sigma_i}{\sum_{i' \in T_t} \sigma_{i'}}}$$

while Equations (49)-(52) still apply provided that we write $H^T_{c,t}$ and $K^T_{c,t}$ instead of $H_{c,t}$ and $K_{c,t}$, in Equations (49)-(50). Factor prices and the pattern of trade are determined by these modified versions of Equations (47)-(52) together with Equations (87)-(88). Since factor supplies are well behaved, a brief analysis of this system reveals that all the discussion of the properties of the world income distribution and its dynamics after Equations (58)-(59) still goes through. In fact, all the results and intuitions developed in the examples of sections 2.4 and 2.5 still apply after we remove the assumption that $N_t = \varnothing$.

The major difference between the world of this sub-section and the one in section 2 is that there is a discrepancy between market-based and real incomes and factor prices. To see this, we need to compute regional price levels. Equations (5)-(7) and (83) imply that:

$$(91) \qquad P_{c,t} = \prod_{i \in N_t} \left\{ \frac{1}{\sigma_i} \cdot \left[\frac{1}{Z_{c,it}} \cdot \left(\frac{w_{c,t}}{1-\alpha_i} \right)^{1-\alpha_i} \cdot \left(\frac{r_{c,t}}{\alpha_i} \right)^{\alpha_i} \right]^{1-\beta_i} \cdot \left[\int_0^{M_{it}} p_{it}(m)^{1-\varepsilon_i} \cdot dm \right]^{\frac{\beta_i}{1-\varepsilon_i}} \right\}^{\sigma_i}$$

Since all regions face the same input prices, Equation (91) shows that, ceteris paribus, the price level is high in regions that have high factor prices and low productivity in nontradable industries. This relationship is the first piece of a theory of the price level. The second piece is a relationship between factor prices, factor endowments and industry productivities. The following examples show how to obtain this additional relationship.

Example 3.1.1: Consider a world economy with H- and K-industries, $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume $\alpha_i=\alpha_H$ and $\max_{c\in C}\{Z_{c,it}\} = Z_{Ht}$ if $i \in I^H$, $\alpha_i=\alpha_K$ and $\max_{c\in C}\{Z_{c,it}\} = Z_{Kt}$ if $i \in I^K$, with $\alpha_H \le \alpha_K$. For simplicity, assume also that $\varepsilon_i=\varepsilon$ and $\beta_i=\beta$ for all $i \in I$. As in section 2.4, we assume that requirement R1 is not binding.[66] The only difference between this world and the one in Example 2.4.1 is the presence of nontradable industries, i.e. $N_t \neq \varnothing$.

Let P_{Ht} and P_{Kt} be the prices of final goods in tradable H- and K-industries. If a region is internationally competitive in tradable H-industries, then the price of final goods of its nontradable H-industries is also P_{Ht}.[67] If a region is not competitive internationally, then the price of final goods in its nontradable H-industries exceeds P_{Ht}. In fact, it follows from Equations (5) and (51)-(52) that the price of the final goods in nontradable H-industries is $\frac{w_{c,t}^{1-\alpha_H} \cdot r_{c,t}^{\alpha_H}}{f_{Ht}} \cdot P_{Ht} \geq 1$. A parallel argument shows that the price of the final goods in nontradable K-industries is $\frac{w_{c,t}^{1-\alpha_K} \cdot r_{c,t}^{\alpha_K}}{f_{Kt}} \cdot P_{Kt} \geq 1$. It then follows from Equations (83)-(84) that the price level of region c is given by:

[66] Note that this implies that all regions have the same productivity in nontradable industries. That is, $Z_{c,it}=Z_{Ht}$ if $i \in N_t \cap I^H$ and $Z_{c,it}=Z_{Kt}$ if $i \in N_t \cap I^K$ for all $c \in C$.

[67] This follows because the technology to produce final goods is the same for all H-industries, and also because the number of input varieties of H-industries does not depend on whether the industry is tradable or nontradable.

$$(92) \qquad P_{c,t} = \left[\frac{w_{c,t}^{1-\alpha_H} \cdot r_{c,t}^{\alpha_H}}{f_{Ht}} \cdot \frac{P_{Ht}}{\sigma_H} \right]^{\sum_{i \in N_t \cap I^H} \sigma_i} \cdot \left[\frac{w_{c,t}^{1-\alpha_K} \cdot r_{c,t}^{\alpha_K}}{f_{Kt}} \cdot \frac{P_{Kt}}{\sigma_K} \right]^{\sum_{i \in N_t \cap I^K} \sigma_i} \quad \text{for all } c \in C$$

As in Example 2.4.1, regions with intermediate factor proportions have diversified production structures while regions with extreme factor proportions have specialized production structures. The sets C_{Mt} , C_{Kt} and C_{Ht} are still defined by Equations (53)-(55) provided we write $H_{c,t}^{T}$ and $K_{c,t}^{T}$ instead of $H_{c,t}$ and $K_{c,t}$. It follows from Equations (51), (52) and (92) that $P_{c,t} = \left(\frac{P_{Ht}}{\sigma_H} \right)^{\sum_{i \in N_t \cap I^H} \sigma_i} \cdot \left(\frac{P_{Kt}}{\sigma_K} \right)^{\sum_{i \in N_t \cap I^K} \sigma_i}$ if $c \in C_{Mt}$ and $P_{c,t} \geq \left(\frac{P_{Ht}}{\sigma_H} \right)^{\sum_{i \in N_t \cap I^H} \sigma_i} \cdot \left(\frac{P_{Kt}}{\sigma_K} \right)^{\sum_{i \in N_t \cap I^K} \sigma_i}$ if $c \in C_{Kt} \cup C_{Ht}$. All regions within the cone share the same price level, and this is the lowest in the world. The reason, of course, is that these regions are competitive both in H- and K-industries. Regions outside the cone have different price levels. Moreover, it is possible to show that these price levels increase the farther away the regions are from the cone. The reason is that the farther away from the cone, the less competitive a region is in one of the industry types and the more expensive it is to produce the final goods of the nontradable industries of this type.

Example 3.1.1 provides us with a simple theory of why and how the price level varies across regions. But it is difficult to reconcile this theory with the data. The later show that price levels are positively correlated with income, so that regional differences in real incomes are substantially smaller than regional differences in market-based incomes. To obtain this pattern in the world of Example 3.1.1 would require that poor regions be located inside the cone and rich regions outside of it. Although this is not impossible from a theoretical standpoint, it does not seem a promising starting point for the construction of an empirically successful theory of the price level.

A positive association between incomes and price levels could arise somewhat more naturally in the world of Example 2.4.2 once we remove the assumption that $N_t = \emptyset$. For instance, if nontradable industries tend to be more human-capital intensive than tradable industries the price level would be high in regions that belong to C_{Kt}; intermediate in regions that belong to C_{Mt}; and low in regions that belong to C_{Ht}. Assume then that most of the variation in income levels is due to differences in savings rates, so that rich regions are those that have low human to physical capital ratios. This does not seem implausible, since most nontradable industries tend to be in the service sector and this sector tends to use a higher human to physical capital ratio than other sectors.

More generally, in the worlds of sub-section 2.4 the correlation between income and price levels is positive or negative depending on how factor proportions vary with income and the factor intensities of nontradable industries relative to tradable ones. The central observation is that price levels should be high in regions that have factor proportions that are inadequate to produce nontradable goods. Building an empirically successful theory of the price level around this notion seems promising, although it remains to be done. Most of the existing research on the price level has focused instead on the role of regional differences in industry productivities. The next example presents a world where these differences generate a positive association between income and the price level.

Example 3.1.2: Consider a world where $Z_{c,it} = \pi_{c,Ht}^{1-\alpha_i} \cdot \pi_{c,Kt}^{\alpha_i}$ for all $i \in T_t$ and $c \in C$, and $Z_{c,it}$=1 for all $i \in N_t$, with $\sum_{c \in C} \pi_{c,Ht} \cdot \frac{H_{c,t}}{H_t} = 1$ and $\sum_{c \in C} \pi_{c,Kt} \cdot \frac{K_{c,t}}{K_t} = 1$. The crucial feature of this example is that productivity differences exist only in tradable industries.[68] This world economy is akin to that in Example 2.5.1. For instance, assume that there are H- and K-industries as in Examples 2.4.1 and 3.1.1. Then, we have that:

$$(93)\qquad P_{c,t} = \left[\frac{(\pi_{c,Ht}\cdot \hat{w}_{c,t})^{1-\alpha_H}\cdot(\pi_{c,Kt}\cdot \hat{r}_{c,t})^{\alpha_H}}{\hat{f}_{Ht}}\cdot\frac{P_{Ht}}{\sigma_H}\right]^{\sum_{i\in N_t\cap I^H}\sigma_i}\cdot\left[\frac{(\pi_{c,Ht}\cdot \hat{w}_{c,t})^{1-\alpha_K}\cdot(\pi_{c,Kt}\cdot \hat{r}_{c,t})^{\alpha_K}}{\hat{f}_{Kt}}\cdot\frac{P_{Ht}}{\sigma_H}\right]^{\sum_{i\in N_t\cap I^K}\sigma_i} \quad \text{for all } c\in C$$

where $\hat{f}_{it} = \min_{c\in C}\left\{\hat{w}_{c,t}^{1-\alpha_i}\cdot \hat{r}_{c,t}^{\alpha_i}\right\}$ for all $i\in I$. Since productivity differences in tradable industries are factor augmenting, regions with higher productivities have higher factor prices. Since there are no productivity differences in nontradable industries, regions with higher factor prices have a higher price level. Note that now a region inside the cone with high productivity in the tradable industries could have a higher price level than a region outside the cone with low productivity in the tradable industries.

In the world of this example, the price level is determined by a combination of two elements: how adequate are the region's factor proportions to produce in the nontradable industries; and how high is the region's productivity in the tradable industries relative to the nontradable ones. In the world of Example 3.1.1, this second force was not present and Equation (93) was reduced to Equation (92). We could also eliminate the first force by assuming that all regions belong to the cone, i.e. by assuming that there is conditional factor-price equalization. In this case, the price level is given by:

$$(94)\qquad P_{c,t} = \left[\pi_{c,Ht}^{1-\alpha_H}\cdot\pi_{c,Kt}^{\alpha_H}\cdot\frac{P_{Ht}}{\sigma_H}\right]^{\sum_{i\in N_t\cap I^H}\sigma_i}\cdot\left[\pi_{c,Ht}^{1-\alpha_K}\cdot\pi_{c,Kt}^{\alpha_K}\cdot\frac{P_{Kt}}{\sigma_K}\right]^{\sum_{i\in N_t\cap I^K}\sigma_i} \quad \text{for all } c\in C$$

[68] This assumption makes sense because nontradable industries consist mostly of services, and in the real world productivity differences in services seem small relative to productivity differences in agriculture or manufacturing.

In Equation (94) the only determinant of the price level is the level of productivity in the tradable industries. This special case is known as the Balassa-Samuelson hypothesis of why the price level is positively correlated with income. Higher productivity in the tradable industries is what makes regions both rich and expensive.

In addition to providing a theory of the price level, the world of this section is also useful because it allows us to study a smoother and more realistic version of the globalization process, i.e. a gradual reduction in the size of N_t. This is not only important for quantitative applications of the theory, but it also leads to new insights regarding the effects of globalization on welfare. The next example shows this.

Example 3.1.3: Consider a world economy with H- and K-industries, such that $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume α_i=0 if $i \in I^H$, and α_i=1 if $i \in I^K$; and β_i=0 for all $i \in I$. Within each type there are "advanced" and "backward" industries. A-regions have the highest possible productivity in all industries, regardless of whether they are "advanced" or "backward". B-regions have the highest possible productivity only in "backward" industries. Up this point all the assumptions are as in Example 2.1.2, except that industry factor intensities are more extreme. Assume next that initially some industries are nontradable, i.e. $N_t \neq \varnothing$; and consider a small step in the globalization process: some "advanced" H-industries become tradable, i.e. some elements of the set $N_t \cap I^H$ move into the set $T_t \cap I^H$. What is the effect of this partial reduction in transport costs on regional incomes?

The reduction of transport costs leads to structural transformation: A-regions reduce their production in "backward" H-industries and increase their production in "advanced" H-industries, while B-regions do the opposite.[69] This increases efficiency

[69] Given the extreme assumptions on industry factor intensities, we know that the distribution of production in K-industries will not be affected.

and raises the combined world production of H-industries, lowering the price of their products and therefore wages all over the world. Therefore, a partial reduction of transport costs has two effects: an increase in efficiency that lowers prices and benefits all regions, and a change in relative prices that benefits some regions but hurts others. A-regions with a large enough ratio of human to physical capital are worse off as a result of this partial reduction in transport costs. [70] If coupled with an appropriate transfer scheme, partial globalization still constitutes a Pareto improvement for the world economy. But now this transfer scheme might require inter-regional transfers towards A-regions with large enough human to physical capital ratios.

The world of this sub-section is a simple and yet very useful generalization of the world of section 2. It allows us to study the sources of regional differences in price levels and also permits us to consider smoother versions of the globalization process. Despite this progress, the world of this sub-section fails to capture a central aspect of transport costs because these only affect final goods. When transport costs affect intermediate inputs, they create incentives to agglomerate production in a single location. We study how this works next.

3.2 Agglomeration effects

Consider a world where transport costs apply only to intermediates, and not to final goods. In particular, assume that the costs of transporting inputs are negligible if $i \in T_t$ but prohibitive if $i \in N_t$; while the costs of trading final goods are negligible for all $i \in I$. An implication of this last assumption is that the price level is the same in all regions and market-based and PPP-adjusted incomes coincide. But this

[70] How is it possible that a region have negative gains from globalization? Since relative prices have changed, the region’s trade opportunities have changed also and it might no longer be possible to achieve the level of income and welfare that the region enjoyed before the reduction of transport costs in the H-industries.

does not mean that we are back to the worlds of section 2. The inability of trading intermediate inputs creates an incentive to concentrate all the production of an industry in a single region. Only in this way, production of final goods can fully take advantage from the benefits of specialization. This force towards the agglomeration of economic activity has profound effects on the world income distribution and its dynamics.

The formal description of the model is quite similar to that of section 2.3. Regional analogues to Equations (1)-(3) apply. Since all regions share spending patterns and face the same final goods prices, the price of consumption and investment is the same for all, and we keep Equation (4) as the numeraire rule. Equations (5)-(6) describing technology apply to all regions, with the corresponding factor prices and industry productivities. The only difference with the model of section 2.2 is that, even after globalization, producers of intermediate inputs in nontradable industries remain sheltered from foreign competition. As a result, in these industries Equations (9)-(10) apply to producers of intermediates in all regions and not only to the lowest-cost ones. Also Equation (8) applies to each region separately since only the demand from local producers of final goods matters for the producers of intermediate inputs. Thus, Equation (44) no longer applies to the producers of intermediate inputs in nontradable industries, and Equation (43) must be modified as follows:

$$(95)\quad I_{c,t} \equiv \left\{ i \in I \;\middle|\; c \in \underset{c' \in C}{\arg\min} \left\{ \left[\frac{1}{Z_{c,it}} \cdot \left(\frac{w_{c,t}}{1-\alpha_i} \right)^{1-\alpha_i} \cdot \left(\frac{r_{c,t}}{\alpha_i} \right)^{\alpha_i} \right]^{1-\beta_i} \cdot \left[\int_0^{M_{c,it}} p_{c,it}(m)^{1-\varepsilon_i} \cdot dm \right]^{\frac{\beta_i}{1-\varepsilon_i}} \right\} \right\} \quad \text{for all } c \in C$$

Equation (95) simply recognizes that the number of intermediate inputs available and their prices can vary across regions.[71] Finally, the market clearing conditions in Equations (11), (45) and (46) apply.

This completes the formal description of the model. For any admissible set of capital stocks, i.e. $K_{c,0}$ for all $c \in C$, and sequences for the vectors of savings, human capital and industry productivities, i.e. $S_{c,t}$, $H_{c,t}$, and $A_{c,it}$ for all $c \in C$ and for all $i \in I$; and a sequence for the set N_t (or T_t); an equilibrium of the world economy after globalization consists of a sequence of prices and quantities such that the equations listed above hold at all dates and states of nature. Like the other worlds we have studied up to this point, there might be multiple geographical patterns of production that are consistent with world equilibrium. Unlike the worlds we have studied up to this point however, there might also be multiple prices and world aggregates that are consistent with world equilibrium. This is, in fact, the most prominent feature of this world.

As usual, we start the analysis by defining the set of factor distributions that allow the world economy to replicate the integrated economy. This set is now as follows:

$$FPE_t \equiv \left\{ d_t \in D_t \;\middle|\; \begin{array}{l} \exists x_{c,it}(m) \geq 0,\ x^F_{c,it} \geq 0 \text{ with} \sum_{c \in C} x_{c,it}(m) = 1,\ \sum_{c \in C} x^F_{c,it} = 1 \text{ and} \\ x_{c,it} = (1-\beta_i) \cdot x^F_{c,it} + \frac{\beta_i}{M_{it}} \cdot \int_0^{M_{it}} x_{c,it}(m) \cdot dm; \text{ such that:} \\ \text{(R1) } x_{c,it} = 0 \text{ if } Z_{c,it} < \max_{c \in C}\{Z_{it}\}; \\ \text{(R2) } H_{c,t} = \sum_{i \in I} x_{c,it} \cdot H_{it} \text{ and } K_{c,t} = \sum_{i \in I} x_{c,it} \cdot K_{it}; \text{ and} \\ \text{(R3) } x_{c,it} \in \{0,1\}\ i \in I \end{array} \right\} \tag{96}$$

[71] Equation (95) assumes that regions always produce intermediates with the lowest indices. This simplifies notation a bit and carries no loss of generality.

When comparing this set to those in Definitions (36) and (86), we observe that requirement (R3) is much stronger now. While Definitions (36) and (86) only required that the entire production of each intermediate were located in a single region, Definition (96) requires that the entire production of each industry (i.e. all intermediates plus final goods) be located in a single region. This is a direct implication of the assumption that intermediate inputs are nontradable. Naturally, this strengthening of requirement R3 reduces the size of FPE_t.[72] Therefore, this set is always smaller than the set in Definition (36). But it need not be smaller than the set in Definition (86), since requirement R4 no longer applies when final goods are tradable.

Assume that industries are "small" and regions are "large" so that requirement R3 is not binding. Then, it is straightforward to see that the equilibria studied in section 2 still apply. If $d_t \in FPE_t$, there exists an equilibrium in which factor prices are equalized across regions and the world economy operates at the same level of efficiency as the integrated economy despite factor immobility and goods market segmentation. If $d_t \notin FPE_t$, the world economy cannot operate at the same level efficiency as the integrated economy and factor prices differ across regions. All the equilibria analyzed in sub-sections 2.4 and 2.5 are also equilibria for the world of this section, and all the results and intuitions we learned in these sub-sections remain valid without qualification.

There is however a major difference between this world and the ones we studied in section 2. While the equilibria described in section 2 were unique in the worlds analyzed there, they are only one among many in the world of this section. The next example makes this point very clear:

[72] The set FPE_t is never empty, but it is smaller than the set of all the factor distributions that are equilibria of the integrated economy. The reason is that some of these equilibria split industries across regions.

Example 3.2.1: Consider a world where all industries are nontradable, i.e. $N_t = I$. Then, any collection of sets $I_{c,t}$ (with $I_{c,t} \neq \emptyset$ for all $c \in C$) that constitutes a partition of I is part of an equilibrium of the world economy.[73] This follows immediately from Equations (5) and (8), which now apply to each region, and Equation (95). Equation (5) shows that the cost of production of final good producers in a given region depends on the number of available inputs. But Equation (8) shows the number of inputs produced in a given region depends on the demand by local producers of final goods.

This world economy exhibits a very strong form of agglomeration effects, as a result of backward linkages in production.[74] If there are no input producers in a region, the cost of producing final goods is infinity and no final goods producer will choose to locate in the region. But if there are no final goods producers in a region, there is no demand for inputs and no input producer will choose to locate in the region. In this world economy, these forces for agglomeration are so strong that they dwarf comparative advantage. It is possible that a given industry locates in a region offering cheap factors and high productivity, but it is also possible that it ends up locating in region offering expensive factors and low productivity.

The world income distribution can be written as follows:

$$(97) \qquad \frac{Y_{c,t}}{Y_t} = \sum_{i \in I_{c,t}} \sigma_i \qquad \text{for all } c \in C$$

[73] This world economy also has equilibria in which industries are split across regions. In these equilibria, all the regions that host a given industry have the same costs of producing the final goods but possibly different numbers and prices of inputs.
[74] Helpman and Krugman [1985] define a backward linkage as a situation in which a final good producer demands many inputs; and a forward linkage as a situation in which many final good producers demand the same input.

Equation (97) is formally very similar to Equation (76). Remember that the latter described the world income distribution in Example 2.5.2 where differences in industry productivities were so strong so as to single-handedly determine comparative advantage. The formal similarity between these two worlds follows because both exhibit an extreme form of specialization. The difference, of course, is the underlying force that determines this specialization. While in Example 2.5.2 regions specialize in a given industry because of their high productivity, in Example 3.2.1 regions specialize in a given industry only because of luck. While in Example 2.5.2 the shape and evolution of the world income distribution reflects only the distribution of industry productivities, in Example 3.2.1 it reflects only randomness.[75]

Example 3.2.1 is extreme because it assumes all industries are nontradable. Assume instead that $T_t \neq \emptyset$, and let $I_{c,t}^{N_t} = N_t \cap I_{c,t}$. As a result of agglomeration effects, any collection of sets $I_{c,t}^{N_t}$ (with $I_{c,t}^{N_t} \neq \emptyset$ for all $c \in C$) that constitutes a partition of N_t is an equilibrium of the world economy. Let again $H_{c,t}^{T}$ and $K_{c,t}^{T}$ be the factor endowments used in the production of tradable goods, i.e. all final goods and the intermediate inputs of tradable industries. It follows that:[76]

$$\text{(98)} \qquad H_{c,t}^{T} = \max\left\{0, H_{c,t} - \sum_{i \in I_{c,t}^{N_t}} (1-\alpha_i) \cdot \sigma_i \cdot \frac{Y_t}{w_{c,t}}\right\} \quad \text{for all } c \in C$$

$$\text{(99)} \qquad K_{c,t}^{T} = \max\left\{0, K_{c,t} - \sum_{i \in I_{c,t}^{N_t}} \alpha_i \cdot \sigma_i \cdot \frac{Y_t}{r_{c,t}}\right\} \quad \text{for all } c \in C$$

[75] Given our assumption of full depreciation of inputs, nothing prevents the pattern of production to shift randomly from generation to generation. This model therefore is consistent with any dynamics for the world income distribution. If inputs depreciated slowly, initial randomness would persist for some time.
[76] Here, I am assuming that industries do not split across regions. As mentioned in an earlier footnote, this is possible too.

Equations (98)-(99) show the factor supplies that are left after subtracting from aggregate factor supplies the factors used in nontradable industries. These Equations are analogous to Equations (87)-(88) of sub-section 3.1. One can use Equations (98)-(99) and a given collection of sets $I_{c,t}^{N}$ to generalize the theory of sections 2.4 and 2.5. For instance, in Example 2.4.1 Equations (47)-(52) still apply provided that we write $H_{c,t}^{T}$ and $K_{c,t}^{T}$ instead of $H_{c,t}$ and $K_{c,t}$.

The effects of this generalization of the theory are hard to assess given the multiplicity of equilibria and the inherent difficulty of finding a "respectable" selection criteria.[77] It is always possible to find perverse equilibria in which regions specialize in the "wrong" industries, i.e. industries in which they do not have comparative advantage. Naturally, all the equilibria of section 2 in which regions specialize in the industries in which they have comparative advantage still apply if requirement R3 is not binding (as we have assumed so far). But there is no compelling reason to choose them over some of the alternatives. Moreover, if requirement R3 is violated or is binding, the equilibria studied in section 2 no longer apply to this world economy. The following example, inspired by Krugman and Venables [1995], relaxes the assumption that industries are "small" and clearly illustrates this point:

Example 3.2.2: Consider a world with two industries I={A,M} and two regions C={N,S}. Assume that both industries have the same factor intensities, i.e. $\alpha_i=\alpha$ for all $i \in I$; but different sizes $\sigma_A<0.5<\sigma_M$ (remember that $\sigma_A+\sigma_M=1$). Also assume that both regions are identical, i.e. they have the same savings, human capital, industry productivities and initial condition. Assume next that the world starts in autarky and globalization proceeds in two stages: in the first one industry A becomes tradable, i.e. N_t={M} for $0 \leq t<T$; and in the second stage also industry M becomes tradable, i.e.

[77] Matsuyama [1991], Krugman [1991] and Fukao and Benabou [1993] study some interesting ways of resolving this indeterminacy.

$N_t=\varnothing$ for t≥T. In the world of autarky, both regions have the same income and the question that I shall address here is: How does globalization affect the world income distribution?

At date t=0, all transport costs disappear except for those that affect the intermediate inputs of industry M. There are two possible patterns of production and trade that can emerge as a result of this. The first one consists of both regions producing the same they did in autarky and not trading between them. Since both regions would have the same goods and factor prices, there would be no incentive for any producer to deviate from this equilibrium. The second possible pattern of production and trade that can emerge consists of each region specializing in a different industry. For instance, assume N specializes in industry M. The absence of other local producers in industry M means that producers in S have no incentive to produce in industry M. Since spending on industry M is more than half of world spending, factor prices are higher in N and therefore producers in N cannot compete in industry A.[78]

It follows from this discussion that the first stage of globalization generates world inequality and world instability. In the world of autarky, both regions had the same income level and income volatility was driven by volatility in fundamentals, i.e. savings, human capital and industry productivities. Globalization generates divergence in incomes because in the equilibrium with specialization the region that "captures" industry M has higher income than the region that is "stuck" producing in industry A. The world income distribution is determined by Equation (97). One effect of this inequality is faster physical capital accumulation in N than in S. Globalization also generates instability, since the pattern of specialization can now change capriciously just as a result of a change in expectations. At any time the specialization pattern can change to the detriment of N and to the advantage of S.

[78] The assumption that industry M is large is crucial in reducing the number of equilibria to three. If there were many "small" M-industries there would also be additional equilibria that split these industries between regions in many different ways.

This constitutes an additional source of income volatility that goes beyond fundamentals.

At date t=T, transport costs for the intermediate inputs of industry M vanish. Although the pattern of production and trade is not uniquely determined, we know that factor prices and incomes are uniquely determined.[79] Moreover, since we have assumed that both industries have the same factor intensities, the world income distribution is now given by Equation (70). It follows that the second stage of globalization starts a slow process of convergence in incomes that eventually restores equality across regions. Throughout this process, expectations no longer play any role and the only sources of income volatility are fluctuations in fundamentals.

This example features a combination of agglomeration effects and "large" industries that underlies most of the work known as economic geography.[80] This research has focused on explaining how income differences can arise among regions that initially have the same fundamentals. The view of globalization and development that arises from this literature is colorful and suggestive, although it has not been subjected yet to serious empirical analysis.

Not surprisingly, globalization might lead to a Pareto-inferior outcome in the world of this section. The following example, which is related to Examples 2.1.2 and 3.1.3, shows this:

---

[79] When $N_t=\emptyset$, we are back to the world of section 2. The reason why the pattern of production is indeterminate is because I have assumed that industry A and M have the same factor intensities. Otherwise we would be in the case of Example 2.1.1.

[80] See Fujita, Krugman and Venables [1999] and Baldwin, Forslid, Martin, Ottaviano and Robert-Nicaud [2003].

Example 3.2.3: Consider a world economy with H- and K-industries, such that $I^H \cup I^K = I$ and $I^H \cap I^K = \varnothing$. Assume α_i=0 if $i \in I^H$, and α_i=1 if $i \in I^K$; and β_i is small (but not zero) for all $i \in I$. Within each type there are "advanced" and "backward" industries. A-regions have the highest possible productivity in all industries, regardless of whether they are "advanced" or "backward". B-regions have the highest possible productivity only in "backward" industries. Assume next that after globalization all industries are non-tradable. This world is just a special case of Example 3.2.1. We know therefore that there is an equilibrium in which A-regions specialize in "backwards" industries while B-regions specialize in "advanced" industries. This equilibrium can be easily shown to deliver equal or less income and welfare than autarky. Since β_i is small for all $i \in I$, the benefits from an increase in market size are negligible. Since the allocation of production worsens relative to autarky, production and income go down as a result of globalization. Therefore, it is not possible to find a transfer scheme that ensures that globalization benefits all.

Although this is real a theoretical possibility, it is not clear yet how seriously should we take the possibility that globalization worsens the world allocation of production and reduces welfare. How important empirically are these agglomeration effects? What is the relative importance of randomness and comparative advantage in determining the pattern of production and trade? The answers to these questions are critical in determining whether the basic policy prescription that simply opening up to trade leads to development really applies or not. In the worlds of this section, opening up to trade can lead to miracles and disasters alike. A miracle is nothing but a lucky region that attracts a large number of industries exhibiting agglomeration effects. A disaster is an unlucky region that cannot do so. Opening up to trade is therefore a gamble. It opens the door for industries to come into the region and enrich it, but it also opens the door for industries to leave the region and impoverish it. Naturally, the temptation to change the odds of this gamble using industrial policies and protectionism might be overwhelming. The prescriptions for development are therefore easy to spot but not pleasant. This is a world characterized by negative international spillovers and strong temptations to use "beggar-thy-neighbor" policies.

Despite the presence of transport costs, differences in regional market size still play no role in determining the world income distribution in the worlds of this sub-section and the previous one. If intermediate inputs are tradable, all regions use the same set of specialized inputs and enjoy the same level of industry specialization or technology to produce final goods. If final goods are tradable, industries concentrate their production in one or few regions and all regions buy their final goods at the same prices. The ability to trade intermediates and/or final goods therefore implies that regional differences in market size cannot be a source of regional differences in incomes. We next turn to a world that features some industries in which neither intermediates nor final goods can be traded. This brings back market size effects as a determinant of the world income distribution.

3.3 The role of local markets

We turn next to a world in which the costs of trading intermediate inputs and final goods are prohibitive if $i \in N_t$, but negligible if $i \in T_t$. As in all the worlds considered in this chapter, the benefits of developing specialized inputs depend on the size of the industry's market. For tradable industries, this market is the world economy. For nontradable industries, this market is the region. As a result, regional differences in market size will be translated into regional differences in the degree of specialization or technology of nontradable industries. [81]

[81] There is little empirical evidence that regional differences in market size are an important determinant of income differences. When one interprets the data from the vantage point of the world of autarky, this observation implies that market size effects are weak and sustained growth is not possible. This has led many researchers to spend a substantial effort in developing autarky models where sustained growth is possible without market size effects. Somewhat ironically, once one takes a world equilibrium view of the growth process what requires a substantial effort is to develop models where regional differences in market size do affect the world income distribution.

Formally, this model is very similar to the one in sub-section 3.1. Equations (81)-(82) describe investment and consumption, while Equations (83)-(84) still provide the numeraire rule and the price level. Naturally, Equation (3) describing spending patterns still applies to all regions, and Equations (5)-(6) describing technology apply to all regions, with the corresponding factor prices and industry productivities. The only difference with the model of section 3.1 is when Equations Equations (7)-(10) describing pricing policies, input demands and the free-entry condition apply. For tradable industries, these Equations apply only to those regions where the lowest-cost producers are located. For nontradable industries, these Equations apply to all regions and not only to the lowest-cost ones. Thus, Equation (44) no longer applies to producers in nontradable industries, and Equation (93) must be replaced by Equation (95). Market clearing conditions are also the same as in the model of section 3.1, and consist of Equations (45)-(46) describing market clearing in regional factor markets, Equation (11) describing market clearing in global markets for tradable industries, and Equation (85) describing market clearing in regional markets for nontradable industries.

This completes the formal description of the model. For any admissible set of capital stocks, i.e. $K_{c,0}$ for all $c \in C$; sequences for the vectors of savings, human capital and industry productivities, i.e. $S_{c,t}$, $H_{c,t}$, and $A_{c,it}$ for all $c \in C$ and for all $i \in I$; and a sequence for the set N_t (or T_t); an equilibrium of the world economy after globalization consists of a sequence of prices and quantities such that the equations listed above hold in all dates and states of nature. Like other worlds we have studied up to now, there might be multiple geographical patterns of production that are consistent with world equilibrium. But unlike the world of the previous sub-section (and like the worlds of section 2 and sub-section 3.1), prices and world aggregates are uniquely determined.

In this world economy, the set FPE_t is empty. Since intermediate inputs that are produced in a region cannot be used in another region, the world economy cannot reach the level of efficiency of the integrated economy.[82] Despite this, it is relatively straightforward to analyze this world. Define again $H^T_{c,t}$ and $K^T_{c,t}$ as the factor endowments devoted to the production of tradable goods, i.e. all intermediate inputs and final goods of tradable industries. Straightforward algebra shows that: [83]

$$(100)\quad H^T_{c,t} = \max\left\{0, H_{c,t} \cdot \left(1 - \sum_{i \in N_t} (1-\alpha_i) \cdot \sigma_i\right) - K_{c,t} \cdot \left(\frac{w_{c,t}}{r_{c,t}}\right)^{-1} \cdot \sum_{i \in N_t} (1-\alpha_i) \cdot \sigma_i\right\} \quad \text{for all } c \in C$$

$$(101)\quad K^T_{c,t} = \max\left\{0, K_{c,t} \cdot \left(1 - \sum_{i \in N_t} \alpha_i \cdot \sigma_i\right) - H_{c,t} \cdot \frac{w_{c,t}}{r_{c,t}} \cdot \sum_{i \in N_t} \cdot \alpha_i \cdot \sigma_i\right\} \quad \text{for all } c \in C$$

Since factor supplies are well behaved, all the results in sections 2.4 and 2.5 regarding market-based incomes and factor prices still go through in the presence of nontradable industries. As in sub-section 3.1, the only important difference between the world of this sub-section and the one in section 2 is that there is a discrepancy between market-based and real incomes and factor prices. In particular, we can write the price level of region c as follows:

$$(102)\quad P_{c,t} = \prod_{i \in N_t} \left\{ \frac{1}{\sigma_i} \cdot \left[\frac{1}{Z_{c,it}} \cdot \left(\frac{w_{c,t}}{1-\alpha_i}\right)^{1-\alpha_i} \cdot \left(\frac{r_{c,t}}{\alpha_i}\right)^{\alpha_i} \right]^{1-\beta_i} \cdot \left[\int_0^{M_{c,it}} p_{c,it}(m)^{1-\varepsilon_i} \cdot dm \right]^{\frac{\beta_i}{1-\varepsilon_i}} \right\}^{\sigma_i} \quad \text{for all } c \in C$$

---

[82] The set FPE_t might be non-empty in the limiting case where $\beta_i \to 0$ (or $\varepsilon_i \to \infty$) for all $i \in I$. But that limiting case brings us to the world of sub-section 3.1.

[83] To see this, note that the shares of human and physical capital devoted to producing the final good of the i^{th} nontradable industry are now $(1-\alpha_i)$ and α_i. Add over industries and note that the share of spending in the i^{th} industry is $\sigma_i \cdot Y_{c,t}$.

The only difference between this Equation and Equation (91) is that the number and price of intermediate inputs varies across regions. Using Equations (6) and (10), we can transform Equation (103) into the following:

$$(103)\quad P_{c,t} = \prod_{i \in N_t} \left\{ \frac{1}{\sigma_i^{\mu_i} \cdot A_{c,it}} \cdot \left(\frac{w_{c,t}}{1-\alpha_i} \right)^{(1-\alpha_i)\cdot\mu_i} \cdot \left(\frac{r_{c,t}}{\alpha_i} \right)^{\alpha_i \cdot \mu_i} \cdot Y_{c,t}^{1-\mu_i} \right\}^{\sigma_i}$$

Basically, this model brings another element to the theory of the price level. To the extent that nontradable industries exhibit increasing returns, regions with larger markets have lower price levels and higher real incomes.

It is straightforward to re-do some of the previous examples in the context of this world. But I shall not do this. The picture that this world generates is clear and unappealing form an empirical standpoint: regional differences in market size are reflected in regional differences in price levels. Ceteris paribus, larger local markets do not lead to higher market-based incomes and factor prices. But they do lead to lower price levels and, as a result, to higher real incomes and factor prices. This is clearly counterfactual.

4. Final remarks

This chapter has developed a unified and yet tractable framework that integrates many key insights of the fields of international trade and economic growth. Its distinguishing feature is that it provides a global view of the growth process, that is, a view that treats different regions of the world as parts of a single whole. This framework incorporates the standard idea that economic growth in the world economy is determined by a tension between diminishing returns and market size effects to capital accumulation. A substantial effort has been made to show how trade frictions of various sorts determine the shape of the world income distribution and its dynamics.

Despite the length of this chapter, some important topics have been left out. The first and most glaring omission is asset trade. This type of trade allows the world economy to redirect its investment towards regions that offer the highest risk-adjusted return.[84] To the extent that patterns of trade are determined by comparative advantage, these are the regions where capital is scarce and productive and this raises efficiency in the world economy. To the extent that patterns of trade are determined by luck, asset trade magnifies the effect of this randomness and this could either raise or lower the efficiency of the world economy. If this were all there is to asset trade, it would not be too difficult to add to this chapter a section on asset trade in which we endow the world economy with a complete set of asset markets. But asset trade does not seem to work as the standard theory of complete markets would suggest. Empirically asset trade seems both much smaller and much more volatile than it would be warranted by its fundamentals, i.e. savings, human capital and industry productivities. To understand these aspects of asset trade it seems necessary to incorporate to the theory features such as sovereign risk, asymmetric information and asset bubbles. Although this is a very important task, it would require another chapter of this magnitude and must therefore be left for future work.[85]

A second important omission of this chapter is government policy. A central aspect of globalization so far has been its imbalanced nature. While economic integration has proceeded at a relatively fast pace, political integration is advancing at a slower pace or not advancing at all. The world economy today features global (or semi-global) markets but local governments. In this context, globalization can lead to a decline in growth and income through a reduction in the quality of policies.

[84] Naturally, asset trade also allows for a better risk sharing and this raises welfare. Better risk sharing might also increase investment and growth. See Obstfeld [1994].

[85] Among the many papers that study the behavior of financial markets in world equilibrium models, see Gertler and Rogoff [1990], Acemoglu and Zilibotti [1997], Ventura [2002], Matsuyama [2004], Martin and Rey [2002, 2004], Kraay, Loayza, Servén and Ventura [2005] and Broner and Ventura [2005].

International spillovers eliminate the incentives to adopt good but costly policies. Trade also "bails out" regions with bad policies since they can spare some of their costs by specializing in industries where bad policies have little effects. As a result of these forces, globalization could create a "race to the bottom" in policies that lowers savings, human capital, and industry productivities. And this could potentially mitigate or even reverse the benefits from economic integration.[86] Understanding the circumstances under which this "race to the bottom" can happen and the appropriate policy corrections that are required to allow the world economy to take full advantage of globalization is another important task. But this task would also require another chapter of this magnitude and cannot be undertaken here.

At first sight, factor movements might seem a third important omission. But I think it is less so. As mentioned in section 2, the notion that physical and human capital is geographically immobile seems a fair description of reality. Moreover, the benefits of factor mobility might be reaped without factors having to move at all. What is really important about factor movements is that they permit factors located in different regions to work together and produce. Advances in telecommunications technology and the standardization of software allow producers around the world to combine physical and human capital located in different regions in a single production process. We can always think of this situation as one in which the production process has been broken down into intermediate inputs. An increased ability to combine factors located in different regions could therefore be modeled as an increase in the tradability of intermediate inputs, or as an increase in the share of intermediate inputs, or as the development of additional inputs with more extreme factor intensities. All of these possibilities could be (and some have already been) analyzed within the framework developed in this chapter.[87]

---

[86] See Levchenko [2004] for a situation in which globalization leads to a "race to the top" in government policies, though.

[87] An increase in the tradability of inputs corresponds to a gradual increase in T_t in the models of section 3.2 and 3.3. An increase in the share of intermediate inputs corresponds to a gradual increase in β_i, while the development of inputs with more extreme factor intensities corresponds to a gradual change in α_i. I have assumed throughout that industry characteristics are time-invariant only for simplicity. All the formulas in this chapter remain valid if we instead assume that industry characteristics vary, perhaps stochastically, over time.

The goal of this chapter has been to convey a global way of thinking about the growth process. To claim success, you should be persuaded by now that developing and systematically studying world equilibrium models is a necessary condition to gain a true understanding of the growth process. By "true", I mean the sort of understanding that allows us to frame clear and unambiguous hypotheses about why some countries are richer than others or what are the main forces that drive economic growth in the world economy. To claim success, you should also be convinced by now that much is already known about the structure of world equilibrium models. But you should also be aware that the global view of economic growth that these models reveal is still somewhat fuzzy and blurred. Sharpening this view is a major challenge for growth and trade theorists alike.

5. References

Acemoglu, D. [2003]: "Patterns of Skill Premia", *Review of Economic Studies*, Vol. 70, pp. 199-230

Acemoglu, D. and Ventura, J. [2002]; "The world Income Distribution," *Quarterly Journal of Economics* CXVII, pp. 659-694

Acemoglu, D. and Zilibotti, F. [2001]: "Productivity Differences," *Quarterly Journal of Economics* CXVI, pp. 563-605.

Acemoglu, D. and Zilibotti, F. [1997]: "Was Prometheus Unbound by Chance? Risk, Diversification and Growth", *Journal of Political Economy* CV, pp.709-751

Ades, A. and Glaeser, E. L. [1999]: "Evidence on Growth, Increasing Returns, and the Extent of the Market," *Quarterly Journal of Economics*, 114 (3), pp.1025-45.

Alcalá, F. and Ciccone, A. [2003]: "Trade, Extent of the Market, and Economic Growth 1960-96," mimeo, Universitat Pompeu Fabra

Alcalá, F. and Ciccone, A. [2004]: "Trade and Productivity," *Quarterly Journal of Economics*, *119* (2), pp. 613-46.

Alesina, A., Spolaore, E., and Wacziarg, R. [2000]: “Economic Integration and Political Disintegration,” *American Economic Review*, *90* (5), pp. 1276-96

Alvarez, F. and Lucas, R. [2004]: “General Equilibirum Analysis of the Eaton-Kourtum model of International Trade,” manuscript, University of Chicago.

Aoki, M. [1986]: “Dynamic Adjustment Behaviour to Anticipated Supply Shocks in a Two-Country Model”, *Economic Journal*, XCVI, pp. 80-100

Arnold, L. G. [2002]: “On the Growth Effect of North-South Trade: the Role of Labor Market Flexibility,” *Journal of International Economics* LVIII, pp. 451-466

Atkeson, A. and Kehoe, P. [1998]: "Paths of Development for Early and Late Bloomers in a Dynamic Hecksher-Ohlin Model", *Federal Reserve Bank of Minneapolis Staff Report* 256

Backus, D., Kehoe, P., and Kydland, F. [1992]:, “International Real Business Cycle”, Journal of Political Economy, C, pp.745-775

Baldwin, R. [1992]: “Measurable Dynamic Gains from Trade” *Journal of Political Economy* C, pp. 162-174

Baldwin, R. and Forslid R. [2000]: “Trade Liberalization and Endogenous Growth. A q-theory Approach,” *Journal of International Economics* L, pp. 497-517

Baldwin, R., Martin, P., and Ottaviano G. I. P. [2001]: “Global Income Divergence, Trade, and Industrialization: The Geography of Growth Take-Offs”, *Journal of Economic Growth*, V pp. 5–37

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicaud: "Economic Geography and Public Policy ", Princeton University Press 2003 ISBN 0-691-10275-9

Bardham, P. [1965]: “Optimal Accumulation and International Trade,” *Review of Economics Studies* XXXII, pp. 241-244

Bardham, P. [1965a]: “Equilibrium Growth in the International Economy,” *Quarterly Journal of Economics* LXXIX, pp. 455-464.

Bardham, P. [1966]: “On Factor Accumulation and the Pattern of International Specialization,” *Review of Economics Studies* XXXIII, pp. 39-44.

Barro, R. J. [1991]: “Economic Growth in a Cross Section of Countries”, *Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407-443

Basu, S. and Weil P. [1998]: “Appropriate Technology and Growth,” *Quarterly Journal of Economics* CXIII, pp. 1025-1053.

Baxter, M. [1992]: “Fiscal Policy, Specialization in the Two-Sector Model: the Return of Ricardo?” *Journal of Political Economy* C, pp. 713-744

Bourguignon, F and Morrison, C. [2002]: “Inequality Among World Citizens: 1820-1992” *American Economic Review*, Vol. 92, No. 4, pp. 727-744

Brems, H. [1956]: “The Foreign Trade Accelerator and the International Transmission of Growth,” *Econometrica* XXIV, pp. 223-238.

Brems, H. [1970]: “ A Growth Model of International Direct Investment,” *American Economic Review* LX, pp. 320-331.

Brezis, S.L., Krugman, P., and Tsiddon D. [1993]: “Leapfrogging in International Competition: A theory of Cycles in National Technological Leadership,” *American Economic Review* LXXXIII, pp. 1211-1219

Broner, F, and Ventura J. [2005]: “Managing Financial Integration,” manuscript, CREI.

Buiter, W.H. [1981]: “ Time Preference and International Lending and Borrowing in an Overlapping-Generations Model” *Journal of Political Economy* LXXIX, pp. 769-797

Caselli, F., Esquivel, G. and Lefort, F. [1996]: “Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics” *Journal of Economic Growth*

Chang, R. [1990]: “International Coordination of Fiscal Deficits”, *Journal of Monetary Economics*, XXV, pp.347-366

Chui, M., Levine, P., and Pearlman, J. [2001]: “Winners and Losers in a North-South Model of Growth, Innovation and Product Cycles,” *Journal of Development Economics* LXV, pp. 333-365.

Cuñat, A. and Mafezzoli, M. [2004]: “Heckscher-Ohlin Business Cycles”, *Review of Economic Dynamics* 7(3), pp. 555-85.

Cuñat, A. and Mafezzoli, M. [2004a]: “Neoclassical Growth and Commodity Trade”, *Review of Economic Dynamics* 7(3), pp. 707-36.

Davis, D. [1995]: “Intra-industry Trade: A Heckscher-Ohlin-Ricardo Approach,” *Journal of International Economics*, XXXIX, pp. 201-226.

Deardorff, A.V. [2001]: “Rich and poor countries in neoclassical trade and growth”, *Economic Journal*, CXI, pp.277-294.

Devereux, M. and Lapham, B. J. [1994]: “ The Stability of Economic Integration and Endogenous Growth,” *Quarterly Journal of Economics* CIX, pp. 299-305.

Devereux, M. B. and Saito, M. [1997]: “Growth and Risk-Sharing with Incomplete International Assets Markets,” *Journal of International Economics* XLII, pp. 453-481.

Devereux, M. B. and Shi, S. [1991]: “Capital Accumulation and the Current Account in a Two-Country Model,” *Journal of International Economics* L, pp. 1-25.

Dinopoulos, E. and Segerstrom, P. [1999]: “A Schumpeterian Model of Protection and Relative Wages” *American Economic Review*, Vol. 89, No. 3, pp. 450-472
.
Dinopoulos, E. and Syropoulos, C. [1997]: “Tariffs and Schumpeterian Growth,” *Journal of International Economics* XLII, pp. 425-452.

Dixit, A.K., and V. Norman: *Theory of International Trade*, Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney 1989, ISBN 0521234816

Dollar, D. [1986]: “Technological Innovation, Capital Mobility, and the Product Cycle in North-South Trade,” *American Economic Review* LXXVI, pp. 177-190

Dollar, D. and Kraay, A. [2003]: “Institutions, Trade, and Growth,” *Journal of Monetary Economics*, *50* (1), pp. 133-162

Dornbusch, R., Fischer, S., and Samuelson, P. A. [1977]: “Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods”, *American Economic Review*, Vol. 67 No.5, pp. 823-839

Dornbusch, R., Fischer, S., and Samuelson, P. A. [1980]: “Heckscher-Ohlin Trade Theory with a Continuum of Goods”, *Quarterly Journal of Economics*, Vol. 95, No. 2, pp. 203-224

Eaton, J. and Kortum, S. [2002]: "Technology, Geography, and Trade”, *Econometrica* 70: pp. 1741-1779

Elkan, R. van [1996]: “Catching Up and Slowing Down: Learning and Growth Patterns in an Open Economy,” *Journal of International Economics* XLI, pp. 95-111.

Feenstra, R. C. [1996]: “Trade and Uneven Growth,” *Journal of Development Economics* IL, pp. 229-256

Findlay, R. [1978]: “An ‘Austrian’ Model of International Trade and Interest Rate Equalization”, *Journal of Political Economy*, LXXXVI, pp. 989-1007

Findlay, R. [1980]: “The Terms of Trade and Equilibrium Growth in the World Economy,” *American Economic Review* LXX, pp. 291-299

Findlay, R and Kierzkowski, H. [1983]: “International Trade and Capital: A Simple General Equilibrium Model” *Journal of Political Economy* XCI, pp. 957-978

Fisher, O'N. [1995]: "Growth, Trade, and International Transfers," *Journal of International Economics* IXL, pp. 143-158

Flam, H. and Helpman, E. [1987]: "Vertical Product Differentiation and North-South Trade," *American Economic Review* LXXVII, pp. 810-822

Frankel, J. A. and Romer, D. [1999]: "Does Trade Cause Growth?" *American Economic Review*, Vol. 89 No.3, pp. 379-399

Francois, J. F. [1996]: "Trade, Labour Force Growth and Wages", *Economic Journal*, 439, pp. 1586-1609

Frenkel, J. and Razin, A. [1985]: "Government Spending, Debt, and International Economic Interdependence", *Economics Journal*, XCV, pp.619-636

Frenkel, J. and Razin, A. [1986]: "Fiscal Policies in the World Economy", *Journal of Political Economy*, XCIV, pp. 564-594

Fujita, M. P., Krugman, P. R., and Venables, A. J.: "The Spatial Economy", MIT Press 1999 ISBN 0-262-06204-6

Fukao, K., and R. Benabou [1993]: "History vs. Expectations: A Comment," *Quarterly Journal of Economics*, CVIII, pp. 535-542.

Gale, D. [1971]: "General Equilibrium with Imbalance of Trade," *Journal of International Economics* I, pp. 141-158.

Galor, O. and Polemarchakis, H.M. [1987]: "Intertemporal Equilibrium and the Transfer Paradox", *Review of Economic Studies*, LIV, pp.147-156

Gancia, G. [2003]: "Globalization, Divergence and Stagnation" *IIES Working Paper* #720

Gertler, M. and Rogoff, K. [1990]: "North-South Lending and Endogenous Domestic Capital Market Inefficiencies", *Journal of Monetary Economics*, XXVI, pp. 245-266

Glass, A.J. and Saggi, K. [1998]: "International Technology Transfer and the Technology Gap," *Journal of Development Economics* LV, pp. 369-398

Glass, A.J. and Saggi, K. [2002]: "Intellectual Property Rights and Foreign Direct Investment," *Journal of International Economics* LVI, pp. 387-410

Greenwood, J., Williamson, S.D. [1989]: "International Financial Intermediation and Aggregate Fluctuations under Alternative Exchange Rate Regimes", *Journal of Monetary Economics*, XXIII, pp. 401-431

Grossman, G. and Helpman, E. [1989]: “Product Development and International Trade”, *Journal of Political Economy* XCVII, pp. 1261-1283

Grossman, G.M. and Helpman E. [1990]: “Comparative Advantage and Long-Run Growth,” *American Economic Review* LXXX, pp. 796-815

Grossman, G.M. and Helpman E. [1991]: “Quality Ladders and Product Cycles,” *Quarterly Journal of Economics* CVI, pp. 557-586

Grossman, G. and Helpman, E. [1991a]: “Quality Ladders in the Theory of Growth”, *Review of Economic Studies* LVIII, pp. 43-61

Grossman, G. M. and Helpman, E.: “Innovation and Growth in the Global Economy”, MIT Press, Cambridge MA, 1991 ISBN 0-262-07136-3

Hall, R. E. and Jones, C. I. [1999]: “Why Do Some Countries Produces More Output per Worker than Others?” *Quarterly Journal of Economics*

Heathcote, J. and Perri, F. [2002]: “Financial Autarky and International Business Cycles”, *Journal of Monetary Economics*, XLIX, pp. 601-627

Helpman, E. [1993]: “Innovation, Imitation, and Intellectual Property Rights,” *Econometrica* LXI, pp. 1247-1280

Helpman, E. and Krugman, P. R.: *Market Structure and Foreign Trade*, MIT Press, Cambridge, Massachusetts; London, England 1985 ISBN 0-262-08150-4

Howitt, P. [2000]: “Endogenous Growth and Cross-Country Income Differences,” *American Economic Review* CXC, pp. 829-846

Jensen, R. and Thursby, M. [1987]: “A Decision Theoretic Model of Innovation, Technology Transfer and Trade”, *Review of Economic Studies*, LIV, pp.631-647

Klenow, P. J. and Rodriguez-Clare A. [1997]: "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" NBER Macroeconomics Annual, MIT Press, Cambridge, MA, pp. 73-102

Klundert, T. van de and Smulders, S. [1996]: “North-South Knowledge Spillovers and Competition: Convergence versus Divergence,” *Journal of Development Economics* L, pp. 213-232

Klundert, T. van de and Smulders, S. [2001]: “Loss of Technological Leadership of Rentier Economies: a Two-Country Endogenous Growth Model,” *Journal of International Economics* LIV, pp. 211-231

Kraay, A., Loayza, N., Servén L. and Ventura J. [2005]: “Country Portfolios”, *Journal of the European Economic Association*, III, forthcoming

Krugman, P. [1979]: "A Model of Innovation, Technological Transfer and the World Distribution of Income", *Journal of Political Economy*, LXXXVII, pp. 253-266

Krugman, P. [1981]: "Trade, Accumulation, and Uneven Growth," *Journal of International Economics* VIII, pp. 149-161

Krugman, P. [1987]; "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher," *Journal of Development Economics* XXVII, pp. 41-55

Krugman, P. [1991] "Increasing Returns and Economic Geography", *Journal of Political Economy*, IC, pp. 483-499

Krugman, P. [1991]: "History vs. Expectations", *Quarterly Journal of Economics*, CVI, pp. 651-667.

Krugman, P., and A. Venables [1995]: "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, CIX, pp. 857-880.

Lai, E.L.-C. [1995]: "The Product cycle and the World Distribution of Income. A Reformulation," *Journal of International Economics* IXL, pp. 369-382

Lai, E. L.-C. [1998]; "International Intellectual Property Rights Protection and the Rate of Product Innovation" *Journal of Development Economics* LV, pp. 133-153

Levchenko, A. [2004]: "Institutional Quality and International Trade", manuscript MIT

Loayza N. V., Knight M., and Villanueva D. [1993]: "Testing the Neoclassical Theory of Economics Growth: A Panel Data Approach" *IMF Staff Papers*

Lundborg, P. and Segerstrom, P.S. [2002]: "The Growth and Welfare Effects of International Mass Migration," *Journal of International Economics* LVI, pp. 177-204

Mankiw, N. G., Romer, D., and Weil, D. N. [1992]: "A Contribution to the Empirics of Economic Growth" *Quarterly Journal of Economics* Vol. 107, No. 2, pp. 407-437

Martin, P. and Rey H. [2002]: "Globalization and Emerging Markets: With or Without Crash", CEPR working paper DP3378

Martin, P. and Rey H. [2004]: "Financial Super-Markets: Size Matters for Asset Trade," *Journal of* I*nternational Economics*, LXIV, pp. 335-661.

Matsuyama, K. [1991]: "Increasing Returns, Industrialization and Indeterminacy of Equilibrium," *Quarterly Journal of Economi*cs, CVI, pp. 617-650.

Matsuyama, K. [2000]: “A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade”, *Journal of Political Economy*, CVIII, pp. 1093-1120

Matsuyama, K. [2004]: “Financial Market Globalization, Symmetry-Breaking and Endogenous Inequality of Nations”, *Econometrica*, LXXII, No. 3, pp.853-884

Matsuyama, K. [2004a]: “Beyond Icebergs: Modeling Globalization as Biased Technical Change” *Working Paper* Northwestern University

Modigliani, F. and Ando, A. [1963]: "The Life Cycle Hypothesis of Saving", *American Economic Review* Vol. 53, No. 1, pp. 55-84

Molana, H., Vines, D. [1989]: “North-South Growth and the Terms of Trade: A Model on Kaldorian Lines”, *Economic Journal*, IC pp. 443-453

Mountford, A. [1998]: “Trade, Convergence and Overtaking,” *Journal of International Economics* XLVI, pp. 167-182.

Myers, M. G. [1970]: ”Equilibrium Growth and Capital Movements Between Open Economies,” *American Economic Review* LX, pp. 393-397.

Dixit, A.K., and V. Norman: *Theory of International Trade*, Cambridge University Press, Cambridge, New York, Port Chester, Melbourne, Sydney 1989, ISBN 0521234816

Obstfeld, M. [1989]: “Fiscal Deficits and Relative Prices in a Growing World Economy”, *Journal of Monetary Economics*, XXIII, pp. 461-484

Obstfeld, M. [1994]: “Risk-Taking, Global Diversification, and Growth,” *American Economic Review* LXXXIV, pp. 1310-1329.

Oniki, H. and Uzawa H. [1965]: “Patterns of Trade and Investment in a Dynamic Model of International Trade” *Review of Economics Studies* XXXII, pp. 15-38.

Ono, Y. and Shibata, A. [1991]: “Spill-Over Effects of Supply-Side Changes in a Two-Country Economy with Capital Accumulation,” *Journal of International Economics* XXXIII, pp. 127-146.

Ortega, F. [2004]: “Immigration policy and the Welfare State”, manuscript UPF

Pritchett, L. [1997]: “Divergence, Big Time”, *Journal of Economic Perspectives,* Vol. 11, No. 3, pp. 3-17

Puga, D., Venables, A. J. [1999]: “Agglomeration and Economic Development: Import Substitution vs. Trade Liberalisation”, *Economic Journal*, CIX, pp. 292-311

Rauch, J. E. [1991]: “Reconciling the Pattern of trade with the Pattern of Migration,” *American Economic Review* LXXXI, pp. 775-796

Rivera-Batiz, L. and Romer P. A. [1991]: “Economic Integration and endogenous Growth,” *Quarterly Journal of Economics* CVI, pp. 531-555.

Rodriguez, F. and Rodrik, D. [2000]: “Trade Policy and Economic Growth: A Skeptic’s Guide to the Cross-National Evidence” *NBER Macroeconomic Annual*

Ruffin, R.J. [1979]: “Growth and the Long-Run Theory of International Capital Movements,” *American Economic Review* LXIX, pp. 832-842

Sachs; J. and Warner, A. [1995]: "Economic Reform and the Process of Global Integration" Brookings Papers on Economic Activity, No.1

Samuelson, P. A. [1948]: “International Trade and The Equalization of Factor Prices”, *Economic Journal*, LVIII, pp. 163-84

Samuelson, P. A. [1949]: “International Trade and The Equalization of Factor Prices, Again”, *Economic Journal*, Vol. 59, No. 234, pp. 181-197

Sauré, P. [2004]: “Revisiting the Infant Industry Argument”, manuscript UPF

Sauré, P. [2004]: “How to Use Industrial Policy to Sustain Trade Agreements”, manuscript UPF

Segerstrom, P. S., Anant, T.C.A., and E. Dinopoulos [1990]: “A Schumpeterian Model of the Product Life Cycle,” *American Economic Review* LXXX, pp. 1077-1091.

Şener, F. [2001]: “Schumpeterian Unemployment, Trade and Wages,” *Journal of International Economics* LIV, pp. 119-148.

Sibert, A. [1985]: “Capital Accumulation and Foreign Investment Taxation”, *Review of Economic Studies*, LII, pp. 331-345

Solow, R. M. [1956]: “A Contribution to the Theory of Economic Growth”, *Quarterly Journal of Economics*, Vol. 70, No. 1, pp. 65-94

Stiglitz, J. E. [1970]: “Factor Price Equalization in a Dynamic Economy,” *Journal of Political Economy* LXXVIII, pp. 456-488

Stokey, N. L. [1991]: “The Volume and Composition of Trade Between Rich and Poor Countries”, *Review of Economic Studies*, LVIII, pp. 63-80

Taylor, M. S. [1993]: “Quality Ladders and Ricardian Trade,” *Journal of International Economics* XXXIV, pp. 225-243.

Taylor, M. S. [1994]: “”Once-Off and Continuing Gains from Trade”, *Review of Economic Studies*, LXI, pp. 589-601

Trefler, D. [1993]: "International Factor Price Differences: Leontief was Right!" *Journal of Political Economy* Vol. 101, No. 6, pp. 961-987

Vanek, J. [1971]: “Economic Growth and International Trade in Pure Theory,” *Quarterly Journal of Economics* LXXXV, pp. 377-390.

Ventura, J. [1997], “Growth and Interdependence,” *Quarterly Journal of Economics*, CXII, 1, pp. 57-84.

Ventura, J. [2002], “Bubbles and Capital Flows” NBER Working Paper No. 9304

Wang, J.-Y. [1990]: “Growth, Technology transfer, and the Long-Run Theory of International Capital Movements,” *Journal of International Economics* XXIX, pp. 255-271.

Wilson, C. [1980]: “On the General Structure of Ricardian Models with a Continuum of Goods: Applications to Growth, Tariff Theory, and Technical Change” *Econometrica*. 48: pp. 1675-1702

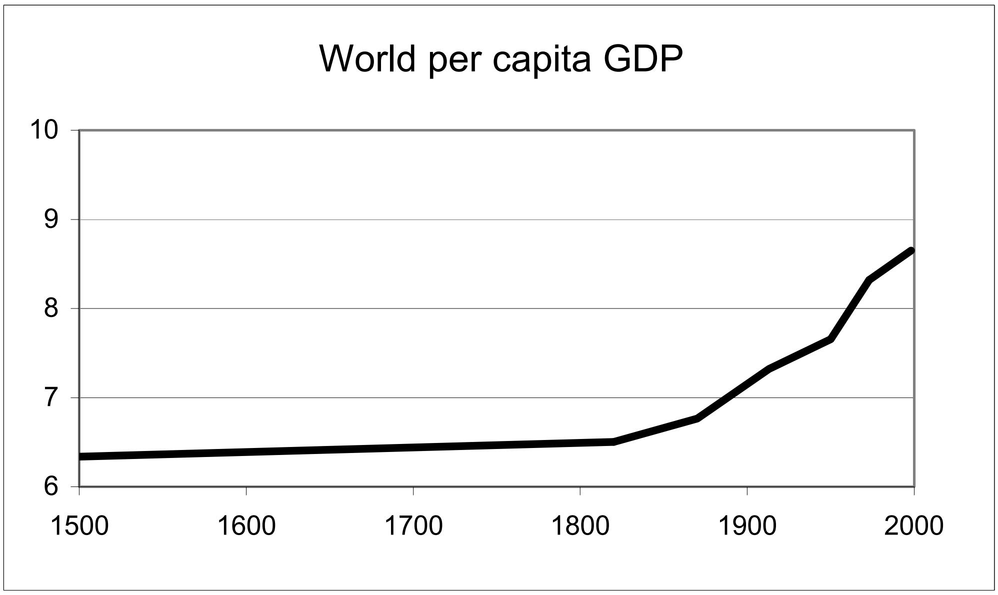
Yanagawa, N. [1996]: “Economic Development in a World with Many Countries,” *Journal of Development Economics* IL, pp. 271-288.

Yang, G. and Maskus K.E. [2001]: “Intellectual Property Rights, Licensing, and Innovation in an Endogenous Product-Cycle Model,” *Journal of International Economics* LIII, pp. 169-387.

Young, A. [1991]: “Learning by Doing and the Dynamic Effects of International Trade” *Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 369-405

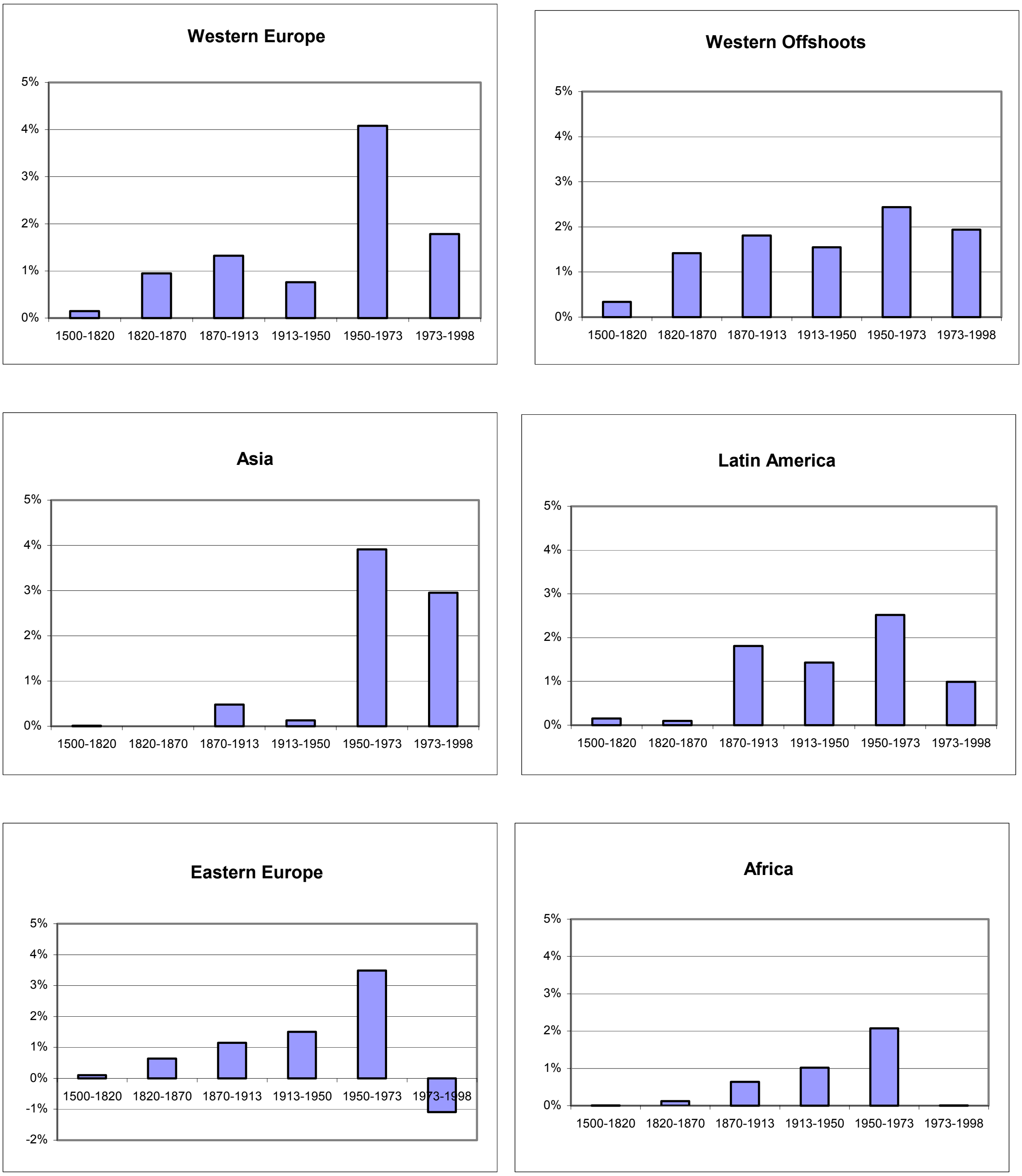
Young, A. [1995]: “The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience” *Quarterly Journal of Economics*, Vol. 110, No. 3, pp. 641-680

Figure 1



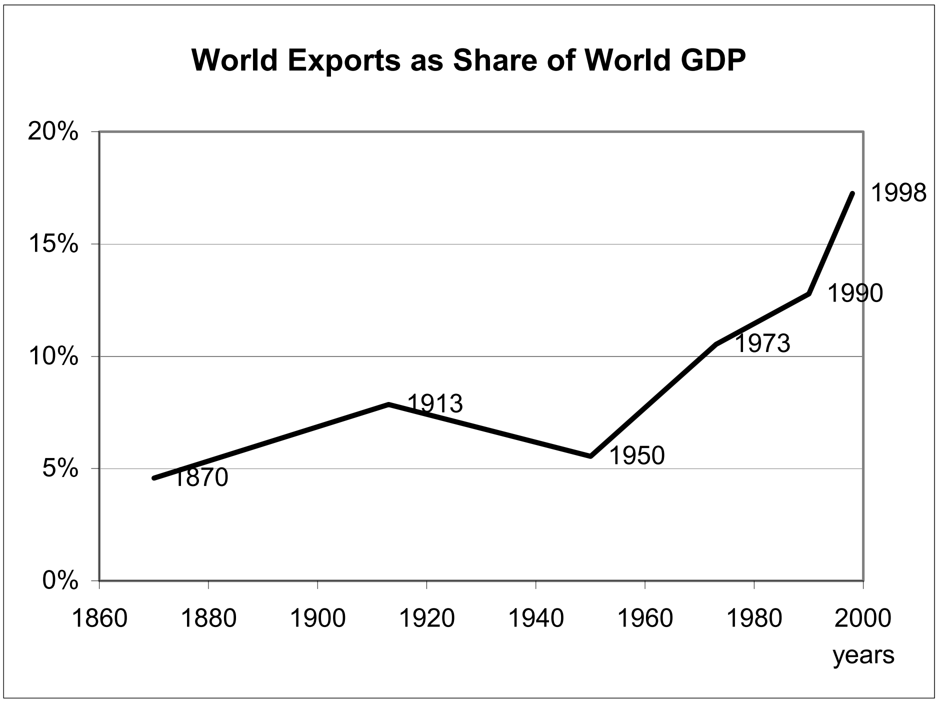
Notes: This figure shows the dynamics of world per capita GDP for the selected years 1500, 1820, 1870, 1913, 1950, 1973, and 1998 (in log of 1990 US$). Data are from Angus Maddison, “The World Economy – A Millennial Perspective” Table 3-1b page126.

Figure 2
Per capita GDP Growth



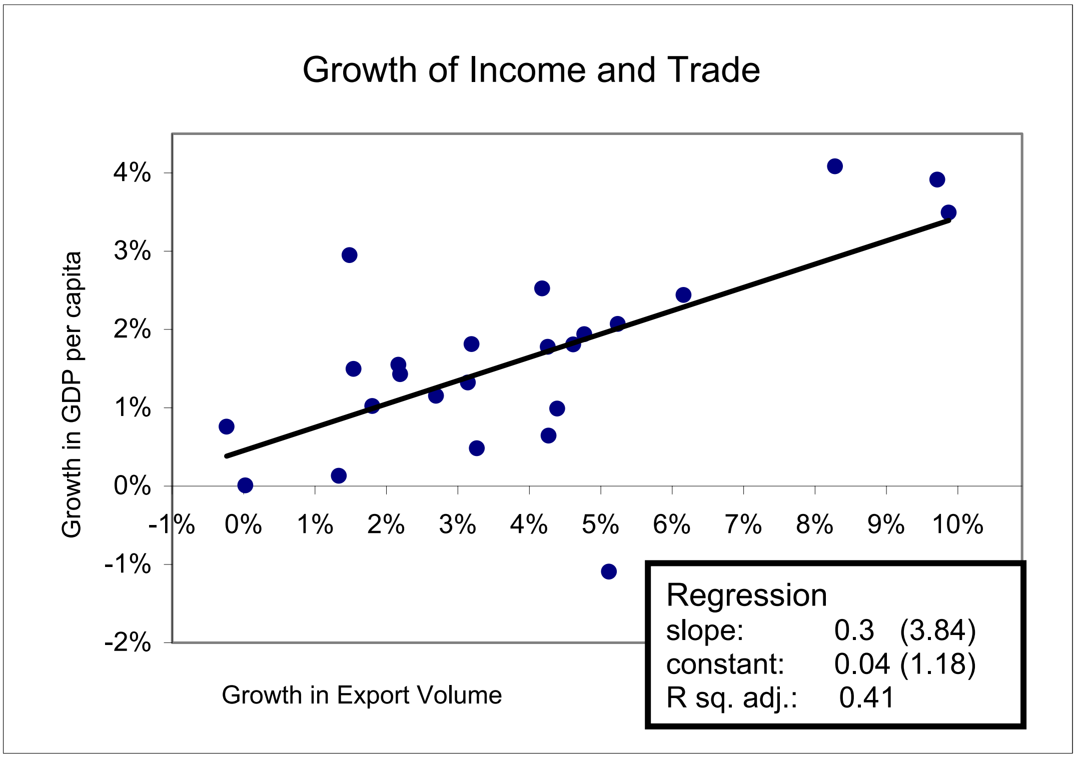
Notes: This figure shows average annual growth rates by major world regions for selected periods. Data are derived from Angus Maddison, "The World Economy – A Millennial Perspective" Table 3-1b page126. (Western Europe contains Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, UK, Portugal, Spain, Greece and 13 small countries; Western Offshoots are United States, Canada, Australia and New Zealand; Asia is China, India, Japan, Korea, Indonesia, Indochina, Iran, Turkey and Other East and West Asian countries; Latin America includes Brazil, Mexico, Peru, and Others; Eastern Europe contains Albania, Bulgaria, Hungary, Poland, Romania and territories of former Czechoslovakia and Yugoslavia; Africa is Egypt and Others.)

Figure 3



Notes: The figure shows Volume of World Exports over World GDP (in constant US$) for selected dates. Data are from Tables 3-1b, A1-b, A2-b, A3-b, A4-b, pages126, 184, 194, 214, and 223 in Angus Maddison, "The World Economy – A Millennial Perspective".

Figure 4

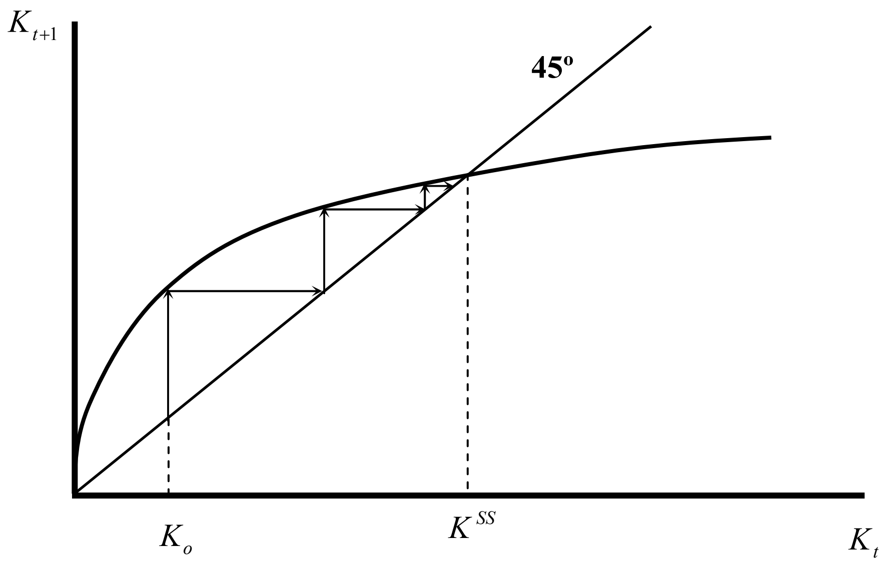


<u>Notes</u>: This figure plots annualized rate of trade growth against annualized rate of per capita GDP growth for major world regions and selected periods. The Regions are Western Europe, Western Offshoots, Eastern Europe and former USSR, Latin America, Asia, and Africa. Periods are 1870-1913, 1913-1950, 1950-1973 and 1973-1998. Each data point stands for one region during one period. The solid line represents the prediction of a linear regression. The estimated regression are reported in the box, t-statistics are in brackets. Data are from Angus Maddison, “The World Economy – A Millennial Perspective”. Data for GDP growth are obtained from Table 3-1b page126 and Table B-10 page 241 (to include Japan). Data for export growth are derived from Table F-3 page362 and Tables A1-b, A2-b, A3-b, and A4-b, pages 184, 194, 214 and 223, respectively.

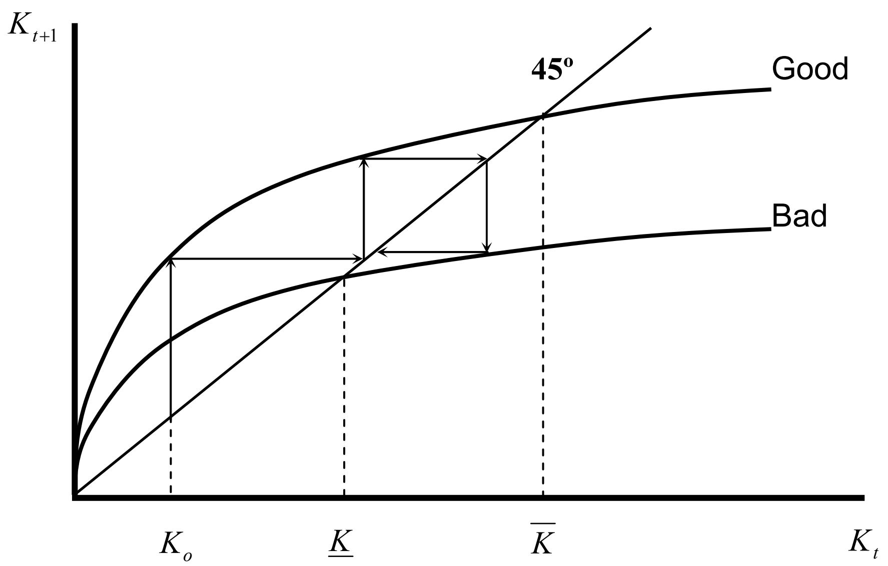
Figure 5

$\alpha\mu + \upsilon < 1$

The "deterministic" case



The "stochastic" case

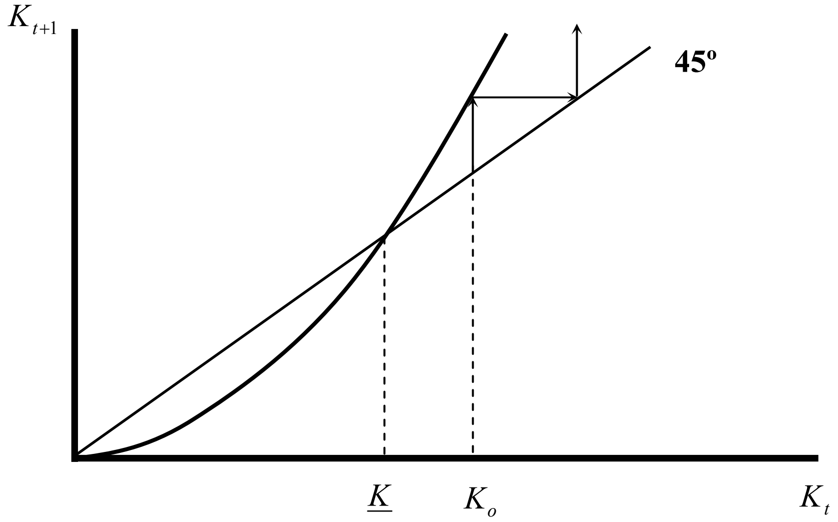


Notes: This figure shows the case of strong deminishing returns and weak market size effects. In the top panel, the stock of physical capital converges monotonically to its unique steady state. The bottom panel shows the stochastic case, where the stock of physical capital converges to the steady state interval $[\underline{K}, \overline{K}]$ within which it fluctuates according to the states of the world.

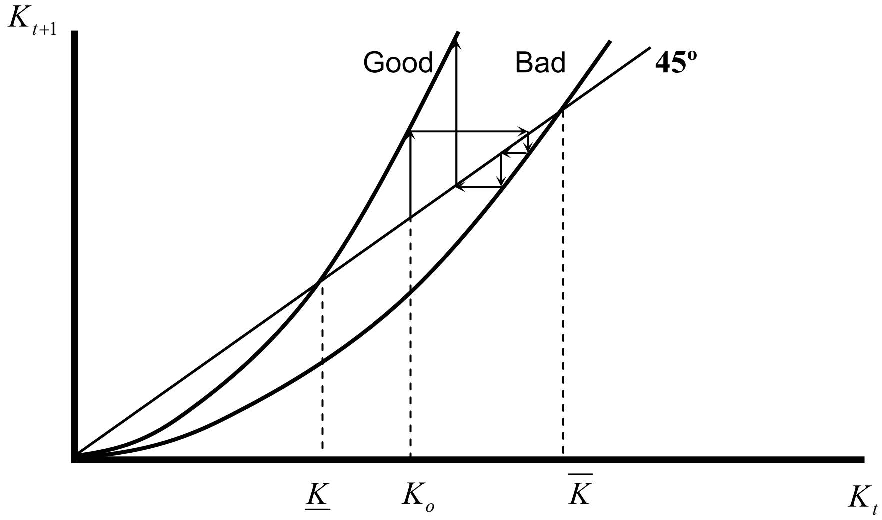
Figure 6

$$\alpha\mu + \upsilon > 1$$

The "deterministic" case



The "stochastic" case

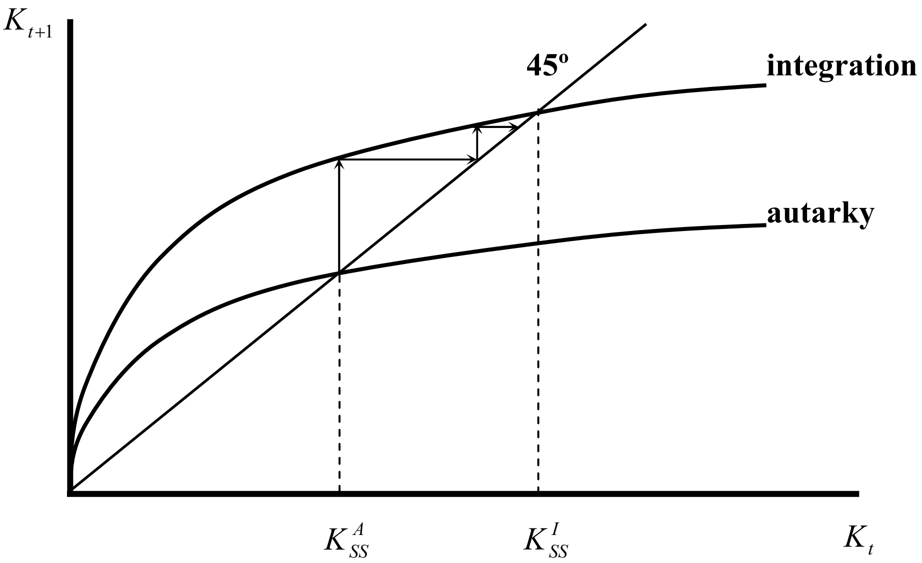


Notes: This figure shows the case of weak diminishing returns and strong market size effects. In the top panel, the stock of physical capital grows at increasing rates since $K_o > \underline{K}$. In the bottom panel the stock of physical capital fluctuates between $\underline{K}$ and $\overline{K}$ according to the states of the world, until it eventually leaves this range.

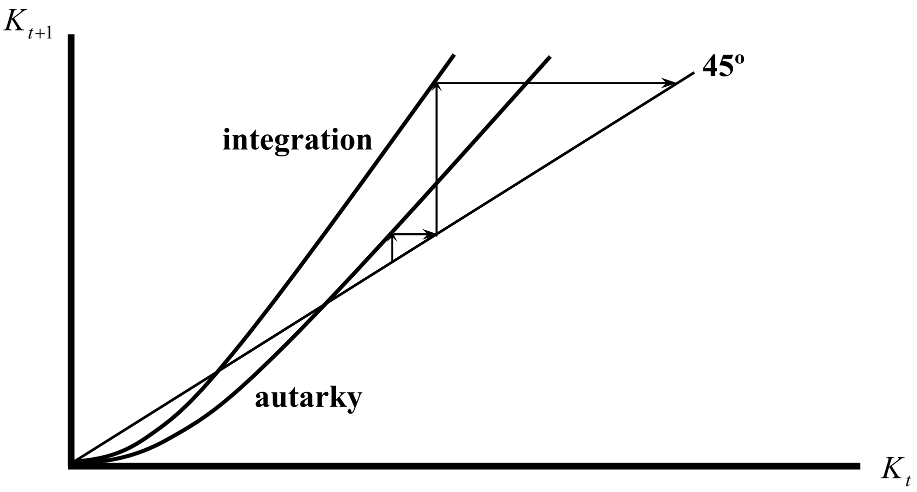
Figure 7

Effects of Economic Integration

$\alpha\mu + \upsilon < 1$

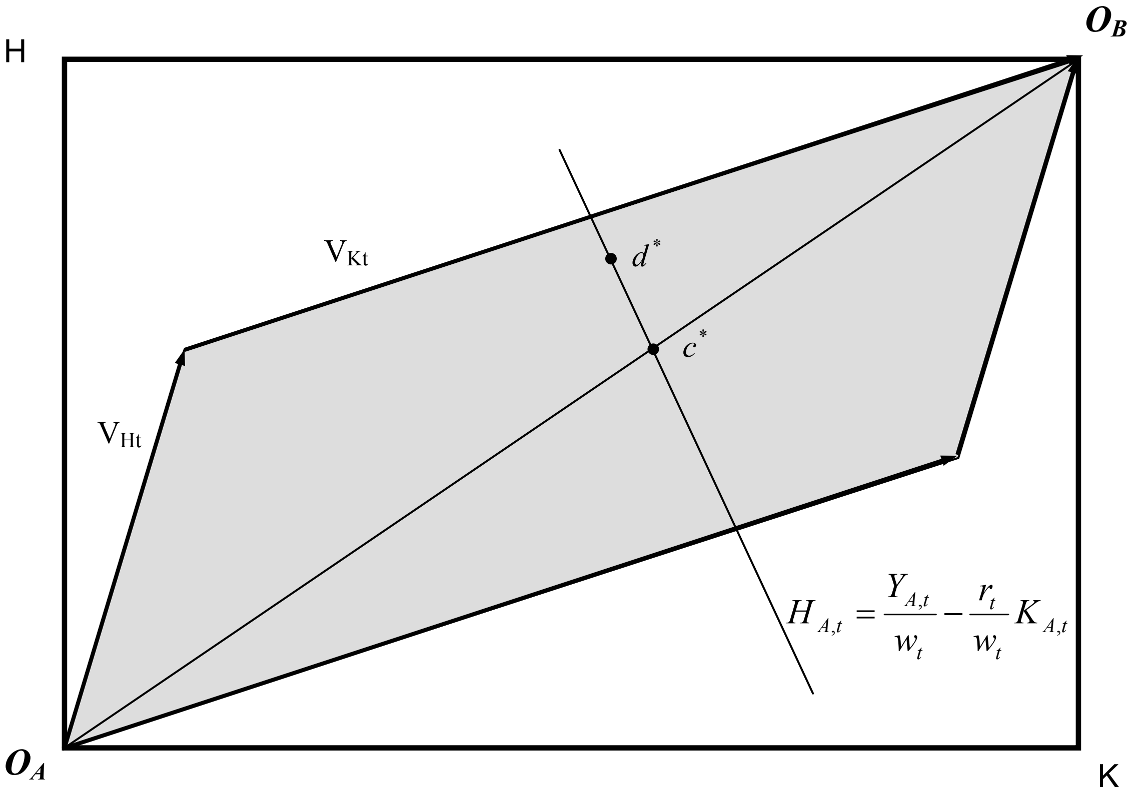


$\alpha\mu + \upsilon > 1$



Notes: This figure illustrates the effects of economic integration. The top panel shows that, if $\alpha \cdot \mu + \upsilon < 1$, economic integration has level effects on income. The bottom panel shows that, if $\alpha \cdot \mu + \upsilon > 1$, economic integration has growth effects on incomes.

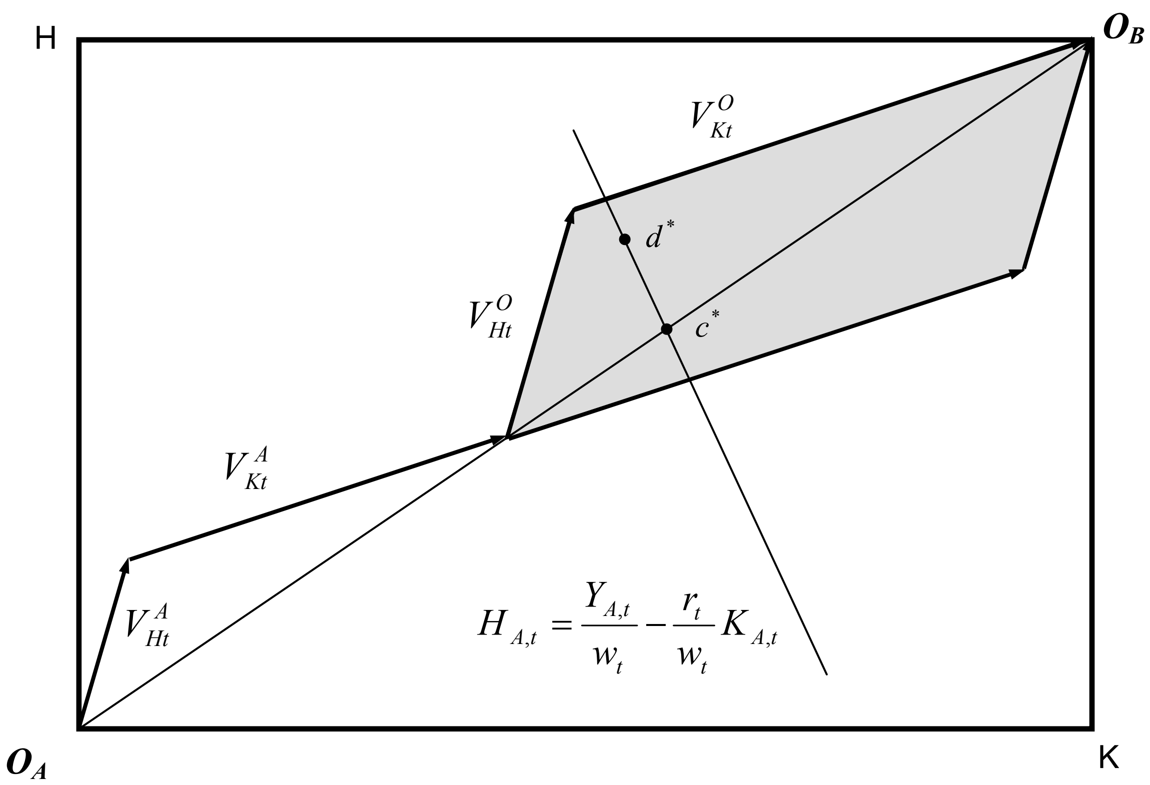
Figure 8



Notes: The box in this figure is a geometrical representation of the set D_t, as each element of this set is a point in the box and vice versa. For instance, d^* is a factor distribution such that A-regions have more human and physical capital than B-regions; but human capital is relatively more abundant in A-regions than in B-regions. The box also contains a set of vectors that represent the factor usage per industry that would apply in the integrated economy. For instance, the vector V_{it} has height H_{it} and width K_{it}. The set FPE_t is the gray area. Since all regions have the same industry productivities, production trivially takes place only in regions with the highest possible productivity (requirement R1). Each of the points in the gray area can be generated as a convex combination of the integrated economy's vectors of factor usage per industry (requirement R2). Since β_i=0, trivially there are no fixed costs of production that are incurred twice (requirement R3). Points outside of the shaded area do not have this property and therefore do not belong to FPE_t.

The factor content of production is given by the regions' factor endowments, i.e. d^*. Since all regions have the same spending shares and use the same techniques to produce all goods, the factor content of consumption lies in the diagonal, i.e. c^*.
In A-regions, the H-industry is a net exporter while the K-industry is a net importer. The opposite occurs in B-regions.

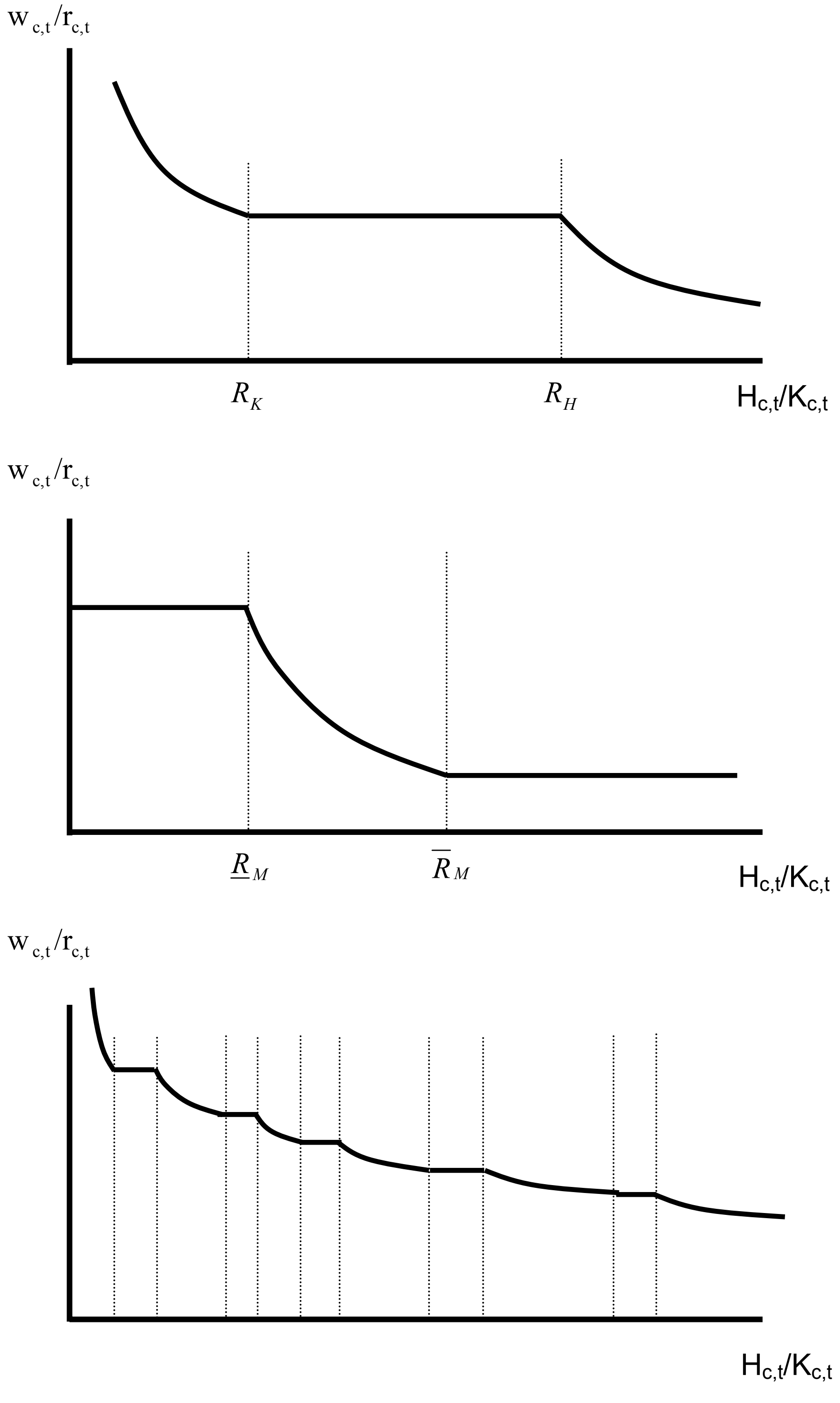
Figure 9



Notes: The box in this Figure is a geometrical representation of the set D_t, as each element of this set is a point in the box and vice versa. For instance, d^* is a factor distribution such that A-regions have more human and physical capital than B-regions; but human capital is relatively more abundant in A-regions than in B-regions. There are four different industries, "advanced" physical (human) capital intensive and "backward" physical (human) capital intensive. The A-countries have a highest productivity in the "advanced" industries; technologies in the "backward" industries are equal in all countries. The vectors V_{it}^{X} have height H_{it}^{X} and width K_{it}^{X} and represent the factor content of the *X*-industries, where *X=A,B* stands for "advanced" or "backward" industries. The set FPE_t is the shaded area. In this set, all "advanced" industries must be located in the A-countries (requirement R1). Once this requirement is satisfied, each of the points in the shaded area can be generated as a convex combination of the integrated economy's vectors of factor usage of the "backward" industries (requirement R2). Since β_i=0, trivially there are no fixed costs of production that are incurred twice (requirement R3). Points outside of the shaded area do not have both properties and therefore do not belong to FPE_t.

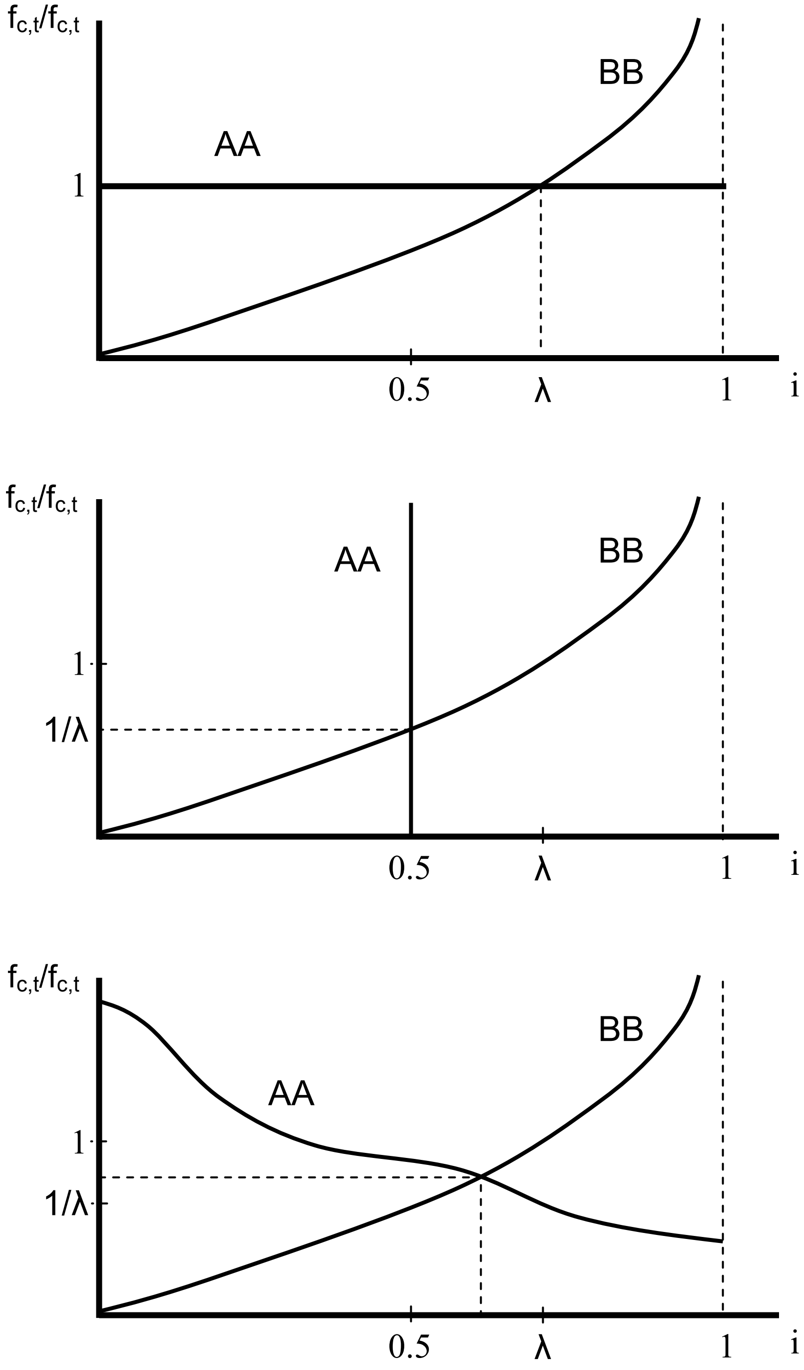
The factor content of production is given by the regions' factor endowments, i.e. d^*. Since all regions have the same spending shares and use the same techniques to produce all goods, the factor content of consumption lies in the diagonal, i.e. c^*. In H-regions, the H-industry is a net exporter while the K-industry is a net importer. The opposite occurs in K-regions.

Figure 10



Notes: This figure shows how the wage-rental ratio varies with the factor proportions. The top panel represents a two-goods, one-cone world where countries with extreme factor proportions are outside the cone (Example 2.4.1). The middle panel represents a three-good, two-cone world where countries with intermediate factor proportions lie outside the cone (Example 2.4.2). The bottom panel shows a world with multiple goods and cones.

Figure 11



Notes: This figure shows how pattern of production and trade (i*) and relative factor costs ($f_{N,t}/f_{S,t}$) are determined in Example 2.5.3. The top panel shows the case of arbitrarily small differences in industry productivities. The middle panel shows the case of arbitrarily large differences in industry productivities. The bottom panel shows the intermediate case.