Globalization, Gains from Specialization and the World Distribution of Output

Paul Beaudry and Fabrice Collard*
University of British Columbia and NBER
University of Toulouse CNRS–GREMAQ and IDEI

July 2003

— Abstract —

This paper proposes and evaluates a model which links change in the world distribution of output-per-worker with the process of globalization. In particular, the model shows how the reduction of trade frictions changes the world distribution of income when specialization of production has productivity enhancing effects. The main claim of the paper is that such a model of globalization – which emphasizes gains from specialization and its interaction with capital accumulation– provides a concise and empirically relevant explanation to recent changes in the cross–country distribution of output–per–worker. In particular, we evaluate the relevance of the model by examining whether its central mechanism finds support in the data. Overall, the analysis highlights how the introduction of gains to specialization into a model of international trade offers new sights into an important puzzle.

Key Words: Income distribution, Cross–Country Growth, Globalization.

JEL Class.: O33, O41.

Corresponding Author: Paul Beaudry, Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver, B.C., Canada, V6T 1Z1.
Email: paulbe@interchange.ubc.ca

*This paper was begun while the second author was visiting the Department of Economics, University of British Columbia. The authors would like to thank Patrick Francois, David Green and Ashok Kotwal for very helpful discussions.

Introduction

Over the last forty years, the volume of international trade has increased enormously, vastly outpacing the growth in world GDP. This observation is often regarded as a defining element of the process of globalization. Interestingly, over the same period, the distribution of income across countries has changed considerably. For example, over the period 1960 to 1998, the distribution of output–per–worker across countries hollowed–out substantially, as mass moved away from the mean of the distribution towards two emerging modes (see Quah [1997], Jones [1997], Beaudry, Collard and Green [2002] among others), thereby giving rise to a bi-modal or *Twin–Peaked* distribution. This simultaneity among the two phenomena is rather intriguing and it is natural to ask whether they may be related. In particular, it is relevant to ask whether the different growth performances underlying the change in the world distribution may be the result of an unequal distribution of the gains associated with globalization. This issue is the object of study of the paper.

As noted by Krugman and Venables [1995], one potential explanation for why countries may vastly differ in their benefits from globalization has to do with the size of their home markets. For example, if certain sectors of the economy exhibit increasing returns to scale (as in Ethier [1979]), then big countries could reap a disproportionate share of the gains associated with the reduction of trade frictions, since they would tend to specialize in the sectors with increasing returns. Alternatively, pecuniary externalities associated with agglomeration may also allow large economies to appropriate much of the gains of free trade.[1] Under either of these scenarios, it would not be surprising to see the world distribution of income change during a period of globalization as large and small countries would have very different economic performances. However, the main weakness of any scale based explanations of observed changes in the cross–country distribution has to do with its predictions with respect to the growth–size relationship. Indeed, such explanations generally imply that during the process of globalization there should emerge a strong positive relationship between measures of size and economic growth. But, as we will review, the cross–country data over the period 1960–98 do not provide much evidence in favor of such a mechanism. Hence, it seems that some other process must be at work.

In this paper, we present and analyze a model of globalization in which there are productivity gains associated with specialization, but no scale effects. The main claim of the paper is that such a model provides a simple explanation for the observed changes in the world distribution of income over the last forty years and that, especially important, we show that its central

[1] The literature that examines links between some type of economy of scale and trade is vast, for example see Ethier [1979, 1982], Krugman [1979], Helpman and Krugman [1985], Grossman and Helpman [1991], Fujita, Krugman and Venables [1999]. Within this literature, there are important distinctions between the sources of economies of scale. For example the New Economic Geography (as surveyed by Neary [2001]) emphasizes pecuniary externalities, while the earlier literature often emphasized non–pecuniary externalities.

mechanism receives considerable support in the data. Although the model is standard on many fronts, it departs from the literature by allowing sectorial productivity to be influenced by the degree to which a country specializes in a particular sector. Hence it is not scale that is relevant in the model, but the degree of specialization. It should be noted that the structure of our model has similarities with many models in the economic geography literature (see for example Ciccone and Hall [1996]) which emphasize productivity gains associated with the density of activity rather than the scale of activity.

The benchmark model we develop is a two good, two factor model with dynamic comparative advantage. The dynamics are driven by an interaction between *(i)* the gains to specialization and *(ii)* the process of capital accumulation. The gains to specialization arise in the capital intensive sector and take the form of improved quality for intermediate goods directed at the capital intensive sector. As we shall show, if these gains to specializations are sufficiently strong, then the move from an autarkic world equilibrium to a free trade equilibrium generates a hollowing–out of the middle of the distribution and the emergence of more than one mode in the world distribution of output.

One of the key elements in our model is the interaction between gains to specialization and the process of capital accumulation. In particular in our model, the main determinant of whether a country gains disproportionately from the opening up of trade depends on its tendency to accumulate capital. For example, a country with a high saving rate and a low rate of labor force growth will tend to specialize in the high capital intensive sector and thereby improve its productivity in that sector. This aspect of the model leads to implications that, during the process of globalization, one should witness the emergence of an abnormally strong link between a country growth rate and its tendency to favor capital deepening (as measured by its investment rate or saving rate, as well as its rate of labor force growth). As we shall show, this prediction of the model finds considerable support in the data.

The remaining sections of the paper are as follows. In section 2 we review a set of observations related to changes in the world distribution over the period 1960–98 (these summarize many of the observations documented in Beaudry et al. [2002], hereafter BCG) and we present a simple conceptual framework for organizing these observations. In section 3, we present the benchmark model. In Section 4, we examine some of the empirical implications of the model. Since many of the assumption in the benchmark model are rather extreme, in Section 5 we discuss generalizations. Finally, the last section offers concluding remarks.

1 Changes in the Distribution of Output–per–Worker

In this section, we first review the salient changes in the cross–country distribution of the output–per–worker that occurred over the period 1960–1998.[2] Then we present a simple framework that helps clarify the different types of explanations which could be behind such changes. Finally, we use the framework to illustrate why theories emphasizing scale effects do not find much support in the data. In the later sections, we will apply a similar approach to evaluate our proposed model.

1.1 The Emergence of Twin–Peaks

Figure 1 reports the distribution of (log) output per worker across the set of Non–Sub–Saharan African countries. The data are taken from the world Penn tables 6.0 for both the years 1960 and 1998.[3] We choose to highlight here the set of Non–Sub–Saharan African countries as to emphasize changes in the distribution of income which are not simply the results of the well–known poor growth performance of the Sub–Saharan Africa countries. The plotted distributions are kernel density estimates based on a Gaussian kernel.[4] Both distributions are expressed as deviations from the given year's mean in order to emphasize changes in the shape of the distribution. It should be noted that the actual distribution shifted substantially to the right from 1960 to 1998. The average output–per–worker increased by 134% between 1960 and 1998 for the 75 countries we consider, implying an average annualized rate of growth of 2.27%. Figure 1 also reports the points in the distribution associated with the interquartile ranges in 1960 and 1998.[5] As can be seen from the figure, the shape of the distribution changed considerably from 1960 to 1998. In effect, it was clearly uni–modal — and close to normal — in 1960, and it hollowed–out in the middle as to become bi–modal. This observation corresponds to what Quah [1993] and Jones [1997] call the *twin–peaks* phenomenon.[6] As can also be seen from Figure 1, the change in the distribution is captured by a substantial widening of the interquartile range. Indeed, the interquartile range expanded by more than 25% between 1960 and 1998.

[2] Most of the observations discussed in this section can also be found in other work such as Quah [1993] Jones [1997], and Beaudry et al. [2002].

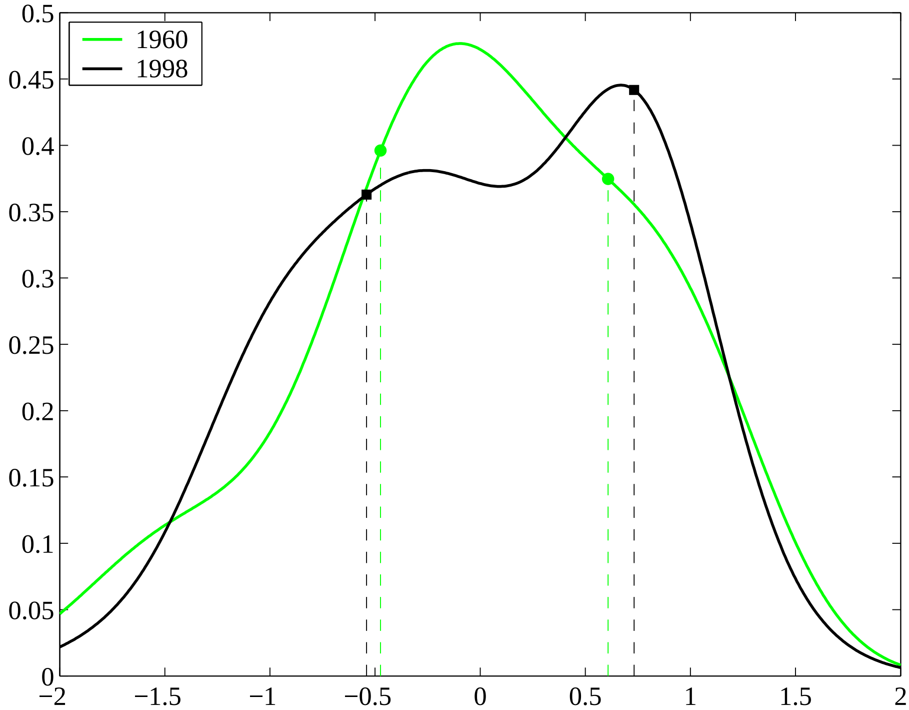
[3] See Appendix A.3 for the list of countries.

[4] The density estimates are computed using the Rosenblatt–Parzen kernel density estimator. We used a Gaussian kernel, with an optimal bandwidth parameter chosen as $h = 1.0592\sigma N^{-1/5}$ where σ is the standard deviation of the data and N is the number of observations.

[5] The points of the interquartile range are calculated from the raw data and not from the kernel estimates of the density function.

[6] To see changes in the distribution that include the Sub–Saharan countries, see Beaudry et al. [2002].

Figure 1: Across-Country (log–)Income Distribution: 1960–1998



Note: The dashed lines indicate the upper and lower bounds delimiting the interquartile range.

1.2 Explaining Changes in Distributions: a Simple Framework

Before examining whether any particular model can explain the change in the distribution observed in Figure 1, it is helpful to ask the following question: *What characteristics must a model possess to explain a hollowing–out of a distribution and the potential emergence of twin peaks?* To answer this question, let us consider the following simple framework where output–per–worker (in log) is represented by y_t and x_t represents a set of exogenous variables that influence y_t. The relationship between y_t and x_t can be expressed as follows.

$$y_t = g_t(x_t) \tag{1}$$

where $g_t(\cdot)$ is a potentially time–varying function, as captured by the index t. This structure allows the relationship between output–per–worker and any exogenous force to potentially change over time, as could be induced for example by a reduction of trade frictions. Assuming that x_t is distributed according to the probability density function (pdf hereafter) $\mu_t^x(.)$, the standard change of variable formula implies that the pdf of y_t, denoted $\mu_t(.)$, is[7]

$$\mu(y_t) = \frac{\mu_t^x[g_t^{-1}(y_t)]}{|g_t'(g_t^{-1}(y_t))|} \tag{2}$$

The conditions which will lead to a hollowing–out (i.e. less mass) in the middle of the distribution can most easily be seen by examining the behavior of the distribution of y_t around its median. Let us therefore define $\widehat{x}_t$ and $\widehat{y}_t$ as the deviations of, respectively, x and y from their median x^m and $y^m = g(x^m)$. Let us also consider two consecutive periods, referred to as 0 and 1 hereafter. The distribution over the first period of log output–per–worker may then be written as

$$\mu_0^y(y) = \frac{\mu_0^x[g_0^{-1}(\widehat{y} + g_0(x^m))]}{|g_0'(g_0^{-1}(\widehat{y} + g_0(x^m)))|} \tag{3}$$

and over the second period as

$$\mu_1^y(y) = \frac{\mu_1^x[g_1^{-1}(\widehat{y} + g_1(x^m))]}{|g_1'(g_1^{-1}(\widehat{y} + g_1(x^m)))|} \tag{4}$$

The hollowing–out in the middle of the distribution can then be studied by focusing on what happens at the median — *i.e.* when $\widehat{y} = 0$ — in which case, from (3) and (4)

$$\mu_0^y(y) = \frac{\mu_0^x(x^m)}{|g_0'(x^m)|} \text{ and } \mu_1^y(y) = \frac{\mu_1^x(x^m)}{|g_1'(x^m)|} \tag{5}$$

As it should be clear from (5), a hollowing–out in the middle of the distribution may result from *(i)* a hollowing–out of the distribution of the driving forces (x_t), as defined by:

$$\mu_1^x(x^m) \leqslant \mu_0^x(x^m)$$

[7] In order to keep notations to the minimum, we are deriving the distribution under the assumption that X is a scalar. The generalization is trivial.

or *(ii)* a modification of the determination of y — *i.e.* an increased sensitivity of output–per–worker and the driving force, as defined by:

$$|g_1'(x^m)| \geqslant |g_0'(x^m)|$$

Hence, in light of this simple framework, one can infer that any relevant model of the observed change in the distribution of output–per–worker should either *(i)* explain and document a hollowing–out of the distribution of a driving variable that affects y_t or *(ii)* explain and document an increase in the marginal effect of a variable x_t on y_t. A nice feature of an explanation in the first category is that it can be directly examined by looking at the distribution. An equally nice feature of the second class of explanations is that it can be examined by simple use of regression analysis (or even more appropriately by a median based regression) since such analysis should detect changes in the function $g(\cdot)$. Since, as documented in BCG, the distributions of most variables that one may conceive as affecting output–per–worker do not seem to have hollowed–out[8], it seems most likely that an explanation to the observed change in the distribution will need to fall in the second category. Accordingly, in this paper we present and examine an explanation to the observed change in the distribution falling in category *(ii)*.

In concluding this subsection, let us note that the above framework can also give insight into the a bimodal distribution around the median. For example, assuming that the distribution of driving forces, $\mu_t^x(.)$, is unimodal and is time–invariant, then the emergence of twin peaks of the type exhibited in Figure 1 will arise if, for some δ,

$$g_1'(x^m - \delta) \leqslant g_1'(x^m) \text{ and } g_1'(x^m) \geqslant g_1'(x^m + \delta) \tag{6}$$

This condition corresponds to the emergence of a local non–concavity of the $g(.)$ function around the median. Hence, even if $\mu_t^x(.)$ is unimodal, it is possible to explain the emergence of bimodality by explaining the emergence of a strong local non–concavity in the $g(\cdot)$ function.

Our goal in Section 3 will be to illustrate how, in the presence of gains from specialization, the process of globalization (defined as a move from an autarkic world equilibrium to free trade equilibrium) can trigger a hollowing–out of the cross–country distribution of output–per–worker. However, before presenting our model, we will briefly overview why we need to look beyond models that emphasize scale effects if we want to explain the change in the distribution.

1.3 Why Economies of Scale is not the Solution?

As we mentioned in the introduction, one potential class of explanations for the observed change in the cross–country distribution of income comes from models which emphasize economies of

[8]Typical examples are the distribution of education levels and investment rates.

scale. In this class of models, the process of globalization may change the distribution of income by changing the relevance of scale. Using the notation introduced above, this class of models suggests that a relevant variable to consider for x is a measure of scale and that the marginal effect of this variable should be seen to increase over the process of globalization. To examine the empirical relevance of this idea, it is helpful to consider a partial adjustment model for (log–) output–per–worker y_t, where y_t converges at rate λ towards its long run equilibrium given by the time varying function $g_t(x_t)$. Then the growth in output–per–worker can be expressed as

$$y_{i,t} - y_{i,t-1} = \lambda g_t(x_{i,t}) + (1-\lambda)y_{i,t-1}$$

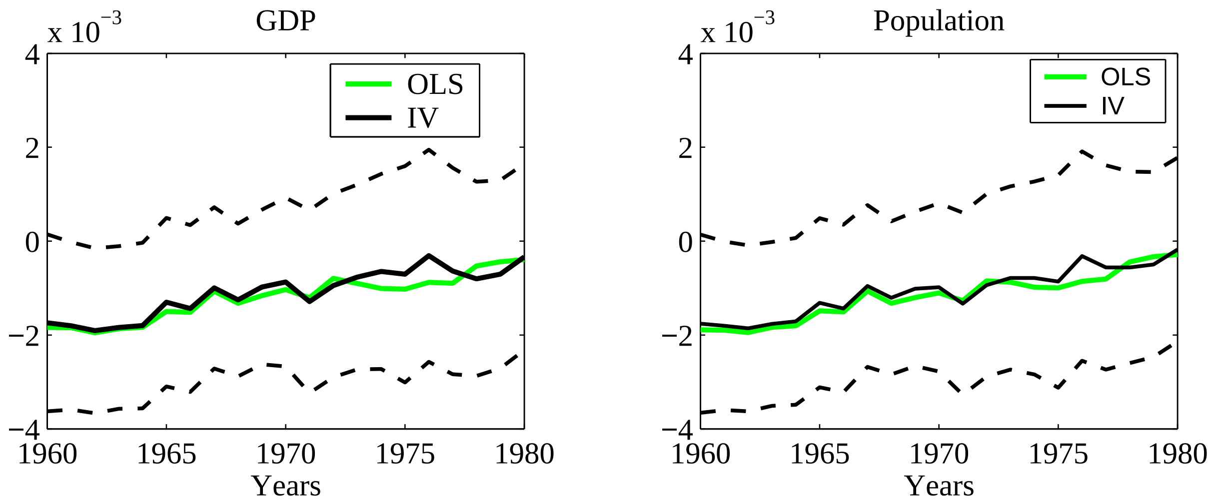
If we further assume that $g_t(\cdot)$ can be approximated by a linear function with time–varying parameters $\Upsilon_{\cdot,t}$, we have the following regression model of output growth in country i:

$$y_{i,t} - y_{i,t-1} = \Upsilon_{0,t} + \Upsilon_{1,t} x_{i,t} + (1-\lambda)y_{i,t-n} \tag{7}$$

where $\Upsilon_{0,t}$ and $\Upsilon_{1,t}$ are respectively a time varying intercept and a time varying slope coefficient obtained from the linear approximation of $g_t(\cdot)$. More generally, (7) can be seen as a special case of a growth regression with time–varying coefficients. This equation indicates that the relevance of models with economies of scale in explaining the hollowing–out of the world distribution income can be gauged by examining whether the parameter for scale in a growth regression increased in importance over the period 1960–98. To explore this possibility, we run a series of 20 rolling regressions of 19 years in length, where we regress growth of output–per–worker over each eighteen year window on the initial value of output–per–worker, a measure of scale and other controls.[9] Figure 2 reports the results obtained when we control for the investment rate, the rate of adult–population growth, and educational investment. Two series of coefficients on scale are reported in the figure: one for OLS estimates (grey line) and the other for IV estimates (dark line). The confidence interval associated to IV estimates is also reported (dashed line). The first coefficient corresponds to the regression estimates over the period 1960–1978, and the last coefficient corresponds to the results over the period 1980–1998. The two measures of scale used are the total GDP in the initial year and the total population in the initial year. As it can be seen from the figures, the coefficient associated with the scaling effect remained *(i)* insignificant and *(ii)* steady over the years. These graphical results indicate that the effects of scale did not contribute significantly to the determination of growth over the last 40 years of the twentieth century nor did it change in importance over time. In order to assess the latter issue more formally, we also perform a test of whether the coefficients on scale change between the period 1960–1978 versus the period 1978–1998 for the different specifications considered in Figure 2. Table 1 reports the stability test statistics as well as the associated p–values. As can be seen from the table, the data do not support any evidence for a structural break in the scale

[9] The results are robust to alternative size of the window.

Figure 2: The Scale Effect: Rolling Regression



Note: The dashed lines are the upper and lower bound for 95% confidence interval (IV).

Table 1: Stability Tests

	OLS	IV	OLS (Educ.)	IV (Educ.)
	Scale variable: GDP			
$\mathcal{Q}$(Scale)	1.503	0.834	0.461	0.059
	[0.220]	[0.361]	[0.497]	[0.808]
	Scale variable: Population			
$\mathcal{Q}$(Scale)	1.786	1.361	0.666	0.373
	[0.181]	[0.243]	[0.414]	[0.541]

Note: p–value in brackets.

effect. Actually, the coefficient associated to the scaling variable is never found to be significant. This implies that these data do not provide empirical support for scale type stories.

Having established that something other than scale effects likely affected the cross country distribution of income, we now propose an explanation based on gains to specialization.

2 The model

This section describes a simple model of globalization, which is aimed at explaining the hollowing–out of the world distribution of output–per–worker. The model consists of a continuum of countries that produce and exchange tradable goods. We first present the behavior of agents in each economy, insisting on the production side, and then turn our attention to the determination of the equilibrium.

2.1 Individual behaviors

We consider a world with a continuum of countries indexed by i, $i \in [1, N]$. In each country, there are two potentially tradeable goods, with quantities of goods produced in country i at time t denoted by $Z_{1,i,t}$ and $Z_{2,i,t}$.

The Z_1 good is produced using intermediate goods M_1 according to the production function:

$$Z_{1,i,t} = M_{1,i,t} \tag{8}$$

where M_1 is a composite of intermediate goods $Q_{i,t}(\ell)$, such that

$$Z_{1,i,t} = \left(\int_0^1 Q_{1,i,t}(\ell)^\rho \mathrm{d}\ell \right)^{\frac{1}{\rho}} \text{ with } \rho < 1 \tag{9}$$

In contrast to good Z_1, the Z_2 good is assumed to be produced using intermediate goods and capital according to:

$$Z_{2,i,t} = AK_{i,t}^{\alpha} M_{2,i,t}^{1-\alpha} \tag{10}$$

with M_2 being composed by the following quality weighted CES aggregate[10] of intermediate goods $Q_{i,t}(\ell)$:

$$M_{2,i,t} = \left(\int_0^1 \left(\theta_{i,t}^{\mathrm{s}}(\ell) Q_{2,i,t}(\ell) \right)^\rho \mathrm{d}\ell \right)^{\frac{1}{\rho}} \text{ with } \rho < 1 \tag{11}$$

[10] Note that the index of intermediate goods is characterized by the same elasticity in the two sectors. This restriction can be relaxed without any difficulty. This would only complicate the analysis by multiplying sub–cases without altering much the results.

In (11), $\theta^{s}_{i,t}(\ell)$, $\ell \in (0,1)$, denotes a *subjective* quality assigned by the firm in sector 2 to the intermediate good $Q_{i,t}(\ell)$ sold by firm ℓ. This index reflects the possibility that the intermediate good $Q(\ell)$ can be more or less productive for sector 2.

The important aspects that differentiate the two tradeable goods are their capital intensities (with good Z_2 being the capital intensive good) and the sensitivity of good Z_2 is sensitive to the quality of intermediate goods.

Assuming firms in the tradeable goods sector take prices as given, then the demand for each intermediate good by sector 1 is given by

$$p_{1,i,t}(\ell) = Q_{1,i,t}(\ell)^{\rho-1} M_{1,i,t}^{1-\rho} \tag{12}$$

where $p_{1,i,t}(\ell)$, $\ell \in (0,1)$ is the price of each intermediate good ℓ sold to firms in sector 1, and the good Z_1 is taken as the numéraire. Accordingly, the demand for each intermediate good directed to sector 2, is given by

$$p_{2,i,t}(\ell) = \theta^{s}_{i,t}(\ell)^{\rho} Q_{2,i,t}(\ell)^{\rho-1} p_{i,t} A K_t^{\alpha} M_{2,i,t}^{1-\alpha-\rho} \tag{13}$$

where $p_{i,t}$ is the price of good $Z_{2,i,t}$ and $p^2_{i,t}(\ell)$ is the price of each intermediate good ℓ sold to firms in sector 2.

We assume that each intermediate good $Q_{i,t}(\ell)$, $\ell \in (0,1)$, is produced by a monopolistic competitive firm by means of effective units of labor. Specificly, one unit of effective labor is required to produce one unit of good, that is,

$$Q_{i,t}(\ell) = Q_{1,i,t}(\ell) + Q_{2,i,t}(\ell) = \Gamma_t L_{i,t}(\ell) \tag{14}$$

where $L_{i,t}(\ell)$ is the labor employed by intermediate producer ℓ in country i at time t, and Γ_t denotes labor augmenting technological progress which is assumed to grow at the exogenous rate $\gamma \geq 0$ ($\Gamma_t = (1+\gamma)\Gamma_{t-1}$). It may be helpful to think of $Q^1_{i,t}(\ell)$ and $Q^2_{i,t}(\ell)$ as slightly different intermediate products sold by the same firm.

We now address the issue of the determination of the quality of intermediate goods going into the production of Z_2. What we want to allow for is the quality of an intermediate good used by producers of Z_2 to depend on the degree of specialization of an intermediate producer in terms of its supply to sector 2 relative to sector 1. To introduce this notion as simply as possible, let us assume the following relationship between the share of total output that producer ℓ devotes to sector 2, $\frac{Q_{2,i,t}(\ell)}{Q_{1,i,t}(\ell)+Q_{2,i,t}(\ell)}$, and let $\theta_{i,t}(\ell)$ be defined as follows

$$\theta_{i,t} = \overline{\theta} \left(\frac{Q_{2,i,t}(\ell)}{Q_{1,i,t}(\ell) + Q_{2,i,t}(\ell)} \right)^{\varepsilon} \tag{15}$$

The above determination of θ_i implies that labor productivity in sector 2 increases as intermediate good firms direct a greater share of their production to sector 2. This reflects a type of quality–by–specialization phenomenon.[11] Although we do not model the internal structure of the firm that gives rise to this effect, we believe it is a description of technology that is worth exploring. In our main analysis, we assume that firms that produce Z_2 cannot directly monitor quality and hence they demand intermediate goods based on their expectation of quality. Given this assumption, intermediate good firms will not try to manipulate quality, and quality will act as an externality.[12] Also let us note that, as in much of the trade literature which emphasizes economies of scale in the capital intensive sector, our model has economies of specialization only in the capital intensive sector.

Given the demands from both sector 1 and 2, the problem of firm $\ell \in (0,1)$ can be seen as to maximize profits by choosing prices and allocating labor to production directed at the different sectors. Denoting by $\sigma_{i,t}(\ell) \in [0,1]$ the share of total labor, $L_{i,t}$, allocated to the production of good 1, the program of an intermediate firm is given by

$$\max_{\{L_{i,t}(\ell),p_{1,i,t}(\ell),p_{2,i,t}(\ell),\sigma_{i,t}(\ell)\}} p_{1,i,t}(\ell)\sigma_{i,t}(\ell)\Gamma_t L_{i,t}(\ell) + p_{2,i,t}(\ell)(1-\sigma_{i,t}(\ell))\Gamma_t L_{i,t}(\ell) - W_{i,t}L_{i,t}(\ell) \quad (16)$$

subject to (12)–(13), where $W_{i,t}$ is the wage rate in economy i.

In order to complete the description of this economy, it is necessary to specify how physical capital is accumulated. To this end, we assume that there is a final good Y, obtained by combining the two tradeable goods according to a Cobb-Douglas technology of the form:

$$Y_{i,t} = (Z_{1,i,t} - X_{1,i,t})^{\varphi}(Z_{2,i,t} - X_{2,i,t})^{1-\varphi} \text{ with } \varphi \in (0,1) \quad (17)$$

where $X_{1,i,t}$ and $X_{2,i,t}$ denote net exports of tradable goods 1 and 2. The price of this final good is denoted by $p^{Y}_{i,t}$. Final good producers take prices as given and select the combination of tradeable goods that maximizes profits.

This final good can be either consumed or invested by households to create capital. For simplicity, we assume, following Solow [1956], that investment represents a constant share of final output, $s_i \in (0,1)$, such that $I_{i,t} = s_i Y_{i,t}$. Therefore, consumption is given by $C_{i,t} = (1-s_i)Y_{i,t}$ and the law of motion of capital is given by

$$K_{i,t+1} = s_i Y_{i,t} + (1-\delta)K_{i,t}, \; K_{i,0} > 0 \text{ given} \quad (18)$$

The labor force in each economy is assumed for now to grow at a common rate $n \geqslant 0$ such that

$$L_{i,t+1} = (1+n)L_{i,t}, \; L_{i,0} \text{ given.} \quad (19)$$

[11]It may be more natural to allow the quality to change slowly over time. For example by setting $\theta_{i,t+1}(\ell) = \left(\overline{\theta}\left(\frac{Q_{2,i,t}(\ell)}{Q_{1,i,t}(\ell)+Q_{2,i,t}(\ell)}\right)^{\varepsilon}\right)^{\lambda}\theta_{i,t}(\ell)^{1-\lambda}$ with $\lambda \in (0,1)$. However, since we will focus on steady states, such a dynamic formulation would not change anything in the long–run. Hence, for sake of simplicity and in order to save on notations, we will keep with the proposed formulation.

[12]Since this assumption is questionable, it is worth noting that many of our results are robust to allowing intermediate good firms to internalize the value of quality as we will discuss in footnotes.

Hence, countries are allowed to differ only in terms of their saving rate. In fact, it will be helpful to think of countries being differentiate according to the index $\nu_i = \frac{s_i}{(1+n)(1+\gamma)-(1-\delta)}$, which can be referred to as the country's propensity to accumulate capital.[13] The variable ν is assumed to be distributed according to the probability density function $\mu^{\nu}(\cdot)$ across countries, which does not change over time.

In what follows, it will considerably reduce notation to abstract from labor augmenting technological change by setting $\gamma = 0$ and $\Gamma_0 = 1$. An attractive feature of this simplification is that it will allow us to disregard the distinctions between output–per–worker and output–per–effective worker, which greatly eases presentation. However, let us emphasize that setting $\gamma = 0$ is without loss of generality as it essentially leaves our results unchanged up to some minor algebraic complications. We will accordingly denote per–worker variables by lowercase letters, for example y and k for output and capital. Also, it will be convenient to adopt the following normalization for $\overline{\theta} = (\frac{\phi}{\phi+(1-\alpha)(1-\phi)})^{-\epsilon}$.

2.2 Equilibrium

We are interested in comparing two types of equilibria for this economy: an autarkic equilibrium and a free trade equilibrium.

A world equilibrium consists of a sequence of prices $\mathscr{P} = \{W_{i,t}, q_{i,t}, p_{i,t}, p^{\mathrm{Y}}_{i,t}; i = 1, \ldots, N\}_{t=0}^{\infty}$, where $q_{i,t}$ is the real rental rate of capital and a set of allocations $\mathscr{Q} = \{K_{i,t}, \sigma_{i,,t}(\ell), L_{i,t}(\ell), Q^{j}_{i,t}(\ell), Y_{i,t}, Z^{j}_{i,t}, X^{j}_{i,t}; \ell \in (0,1); j = 1,2; i = 1, \ldots, N\}_{t=0}^{\infty}$ such that,

1. Given the sequence of prices $\mathscr{P}$, $\mathscr{Q}$ solves the firms' problem;
2. Given a sequence of allocations, $\mathscr{Q}$, $\mathscr{P}$ clears the markets;
3. Capital accumulates according to equation (18);
4. International markets clear in the sense that
   - In the absence of trade (autarky), $X^{j}_{i,t} = 0$
   - If there is free trade, $p_{i,t}$ is independent of i and trade is balanced in each economy — *i.e.* $X^{1}_{i,t} + p_t X^{2}_{i,t} = 0$

For future reference, it is useful to note that when we refer to a country's level of output we mean the quantity given by

$$C_{i,t} + I_{i,t} = \frac{1}{p^{\mathrm{Y}}_{i,t}} Z_{1,i,t} + \frac{p_{i,t}}{p^{\mathrm{Y}}_{i,t}} Z_{2,i,t}$$

[13] This is a natural interpretation of this quantity as it represents the country's long run capital–output ratio.

This is the value of production calculated in units of the final good and is the value of national income.

3 From Autarky to Free Trade

This section examines how when moving from an autarkic equilibrium to a free trade equilibrium — which we refer to as the process of globalization — output–per–worker is affected by potential productivity gains associated with specialization. We start by assuming that the world economy is composed of a continuum of closed economies that all live in autarky and show that each economy behaves as the standard Solow [1956] growth model with Cobb–Douglas technology. In particular, we show that the only exogenous cross–difference in the model, which is the propensity to accumulate capital, just retrieves its standard level effect on output. More precisely, if the distribution of ν, the propensity to accumulate capital, is uni–modal, the distribution of output–per–worker (and also its logarithm) is also uni–modal. We then open the economies and show how free trade will lead to a hollowing–out of the distribution of output–per–worker and the emergence of new modes when there are gains to specialization.

For each case — the autarkic and free trade equilibrium — we first examine the determination of output taking a country's capital labor ratio as given. Then we study properties of the cross–country distribution when physical capital is allowed to adjust to its steady state value. As it can be expected, the resulting distributions will be affected by the distribution of ν. We should immediately emphasize that, for clarity of presentation, we are assuming countries differ only in terms of their propensity to accumulate capital (ν). Obviously, this assumption is at odds with the data since it is well known that differences in ν can only account for a fraction of the cross–country differences in income levels. One way to remedy this failure would be to allow workers across countries to differ in terms of their effective units of labor. However, for our purposes, this would only complicate matters without providing extra insight. Hence, we choose to focus on how differences in propensities to accumulate capital provide insight to changes in the cross–country distribution of income over time, knowing that such a focus leaves unexplained much of the cross–section distribution at a point in time.

3.1 Autarky

In the absence of international trade, the model we have presented takes a very simple form. In particular, as stated in Proposition 1, aggregate output in the country is given by a simple Cobb–Douglas production function.[14]

[14] The interested reader is referred the Appendix for proofs of the Propositions.

Proposition 1 *In the absence of international trade, the relationship between output–per–worker and capital–per–worker is given by*

$$y_t = Bk_t^{\alpha(1-\varphi)} \tag{20}$$

with $B \equiv \frac{\varphi^{\varphi}((1-\alpha)(1-\varphi))^{(1-\alpha)(1-\varphi)}}{(\varphi+(1-\alpha)(1-\varphi))^{\varphi+(1-\alpha)(1-\varphi)}} A^{1-\varphi}$.

Two aspects are worth emphasizing about Proposition 1. First, to understand this and later proposition, it is helpful to note that the amount of labor time being allocated by each intermediate good firm between the production of type 1 versus type 2 type intermediate goods does not change as the capital–labor ratio changes. In effect, as the capital–labor ratio increases the production of the type 2 good increases, but this leaves unchanged the ratio of marginal productivity of labor across the two activities. Therefore, as the capital–labor ratio increases, a country does not become more specialized in the production of the capital intensive good. Accordingly, there is no interaction between the degree of gains to specialization and capital–labor ratios. Hence, as can be seen in Proposition 1, changes in ϵ do not affect the aggregate production function.[15] Second, Proposition 1 makes it clear that the model is essentially a collection of standard Solow growth models and therefore shares its main features. This can be easily understood by noting that, in the absence of trade, the world economy essentially consists of the collection of closed economies that all produce and sell the final good on their local market. A direct implication of this result is that the model generates the standard negative relationship between cross–country growth in output per worker and its initial level, and a positive (negative) relationship with the saving rate (population rate of growth). It also has strong implications for the cross–country distribution of (log–) output–per–worker.

We now turn to the determination of the cross–country distribution of output–per–worker when we allow capital to adjust to its steady state level. Given the capital accumulation equation (18) and the aggregate production function given in Proposition 1, one can immediately notice that the closed economy model behaves as a standard neoclassical growth model. Indeed, the steady state value of y_i in a country with a propensity to accumulate capital given by ν_i is :

$$y_{i,t} = B^{\frac{1}{1-\alpha(1-\varphi)}} \nu_i^{\frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}} \tag{21}$$

It is now straightforward to derive the steady state distribution[16] of y, which is given in the next proposition.

[15]The result that ϵ does not affect the aggregate production function is in part the result of the convenient normalization in the specification of the gains to specialization.

[16]Note that since the logarithmic function is strictly monotonic, the the distribution of log output–per–worker inherits the shape (in particular the uni– or bi–modality) of the distribution of output–per–worker.

Proposition 2 *In the absence of international trade, the steady state distribution of output–per–worker is given by*

$$\mu^y = \frac{1-\alpha(1-\psi)}{\alpha(1-\varphi)} \frac{g^{\mathrm{A}}(y)}{y} \mu^\nu(g^{\mathrm{A}}(y))$$

where $g^{\mathrm{A}}(y) = B^{1\frac{1}{\alpha(1-\varphi)}} y^{\frac{1-\alpha(1-\varphi)}{\alpha(1-\varphi)}}$ *and* $\mu^\nu(\cdot)$ *denotes the distribution of* ν*.*

As can be seen in Proposition 2, the distribution of output–per–worker essentially consists in a re–scaling of the distribution of ν by a monotonic function, $g^{\mathrm{A}}(y)$, of y. A direct implication of this linearity is that if the propensity to accumulate capital, ν, is distributed according to a unimodal distribution, then the distribution of output–per–worker inherits this unimodality under autarky. Hence, unless the accumulation forces have twin–peaked distribution — for which we do not have much evidence in the data — the model generates a unimodal distribution of output–per–worker in the long run.

3.2 Free trade

We now examine the implications of opening up trade for the equilibrium distribution of output–per–worker. To derive these implications, we need to resolve two distinct questions. First, we need to determine how a country's aggregate level of output changes when it can trade both type of goods (Z_2 and Z_1) in the world market at relative price p. Second, we need to determine the world relative price p. Our approach is to focus on the first question and adopt a convenient approximation for the second. In effect, we assume that world price of Z_2 in terms of Z_1 under free trade can be approximated by the median price under autarky. That is, the world price under free trade is taken to be the equilibrium relative price $p_{i,t}$ that arises in an autarkic economy with $\nu = \nu^m$. Although this is a strong assumption, we believe that it is a quite reasonable approximation. For later reference, it is convenient to denote by $k^\star$ the steady state capital–labor ratio for an autarkic economic with $\nu = \nu^m$.

In order to better characterize the main properties of the economy, it is helpful to determine the conditions for full specialization under free trade. This is undertaken in Lemma 1.

Lemma 1 *There exists a level of capital per efficient worker, denoted by* $\widetilde{k}(p)$*, above which a small open economy specializes in the production of good 2. Furthermore the threshold level,* $\widetilde{k}(p)$*, is a decreasing function of the elasticity,* ε*, and*

$$\lim_{\varepsilon \to \frac{\alpha}{1-\alpha}} \widetilde{k}(p) = k^\star$$

The intuition for Lemma 1 is standard from international trade theory. As in any Hecksher and Ohlin type model, economies specialize in the production of the good which is intensive in the factor they have in abundance. Hence, when the economy has accumulated a sufficiently high level of capital per efficient worker, it specializes in the production of good 2, which is capital intensive. Given this notation, we can now present the relationship between a country's capital labor ratio and its level of production. In all that follows, it will be helpful to assume that $\varepsilon < \frac{\alpha}{1-\alpha}$ as to ensure that a country's production possibility frontier is continuous.

Proposition 3 *If $0 \leqslant \varepsilon < \frac{\alpha}{1-\alpha}$, then under free trade a country's level of output–per–worker is given by*

$$y = \begin{cases} \Phi p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right) & \text{if } k \leqslant \widetilde{k}(p) \\ \Phi p^{\varphi}A\overline{\theta}^{1-\alpha}k^{\alpha} & \text{if } k > \widetilde{k}(p) \end{cases}$$

where $\widetilde{k}(p) \equiv \left[A(1-\alpha)\overline{\theta}^{1-\alpha}p\right]^{\frac{-1}{\alpha}}$ *and* $\Phi = \varphi^{\varphi}(1-\varphi)^{1-\varphi}$.

This proposition is illustrated in Figure 3, where the production function is depicted under both autarky and free trade. In the autarkic situation, the technology is strictly concave, as implied by Proposition 1. When world trade is available, Proposition 3 indicates that as long as ϵ is positive, the production function will display a local convexity stemming from the productivity effects of specialization. This can be seen from the fact that the exponent on k is $\frac{\alpha}{\alpha-\varepsilon(1-\alpha)} > 1$ when $k \leqslant \widetilde{k}(p)$.[17] In order to understand this local convexity, it is useful to first consider the case with no gains to specialization ($\varepsilon = 0$). Panel (a) of Figure 3 corresponds to this situation. When $\epsilon = 0$, international trade causes the aggregate production function of the economy to become a linear function of the capital labor ratio in the zone where both goods Z_2 and Z_1 are produced ($k \leqslant \widetilde{k}(p)$). This result stems from the ability of firms to reallocate labor between the 2 sectors, therefore avoiding decreasing returns to capital. In effect, any change in the capital labor ratio when $k \leqslant \widetilde{k}(p)$ induces a reallocation of labor between the two sectors. Since this can be done while simultaneously maintaining both a fixed relative price and a fixed capital–labor ratio, there are no decreasing returns to capital. In contrast, when the capital–labor ratio is high enough — *i.e.* when $k \geqslant \widetilde{k}(p)$ — the economy fully specializes in the production of the capital intensive good and aggregate production becomes concave. In the case with gains to specialization, $\varepsilon > 0$, the effect of reallocating labor across sectors as the capital labor ratio

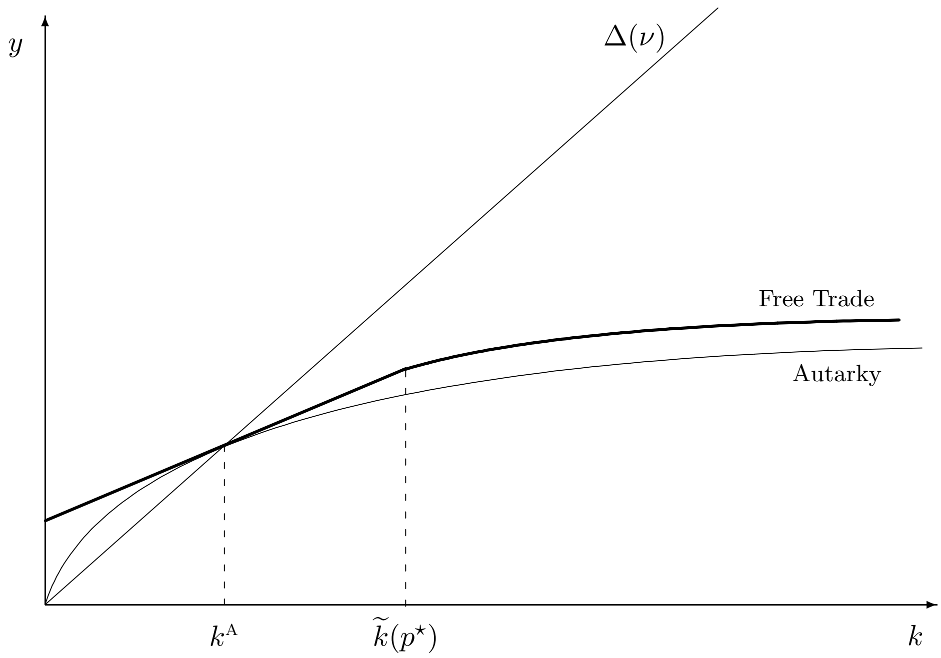
[17] At this point, one may think that the local convexity is created solely by the presence of the externality. However, it is straightforward to verify that if the returns to specialization were internalized by the firm, the proposition would still hold but the technology would take the form

$$y = \begin{cases} \Phi p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}(A(1-\alpha)(1+\varepsilon)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right) & \text{if } k \leqslant \widetilde{k}(p) \\ \Phi p^{\varphi}A(1+\varepsilon)\overline{\theta}^{1-\alpha}k^{\alpha} & \text{if } k > \widetilde{k}(p) \end{cases}$$

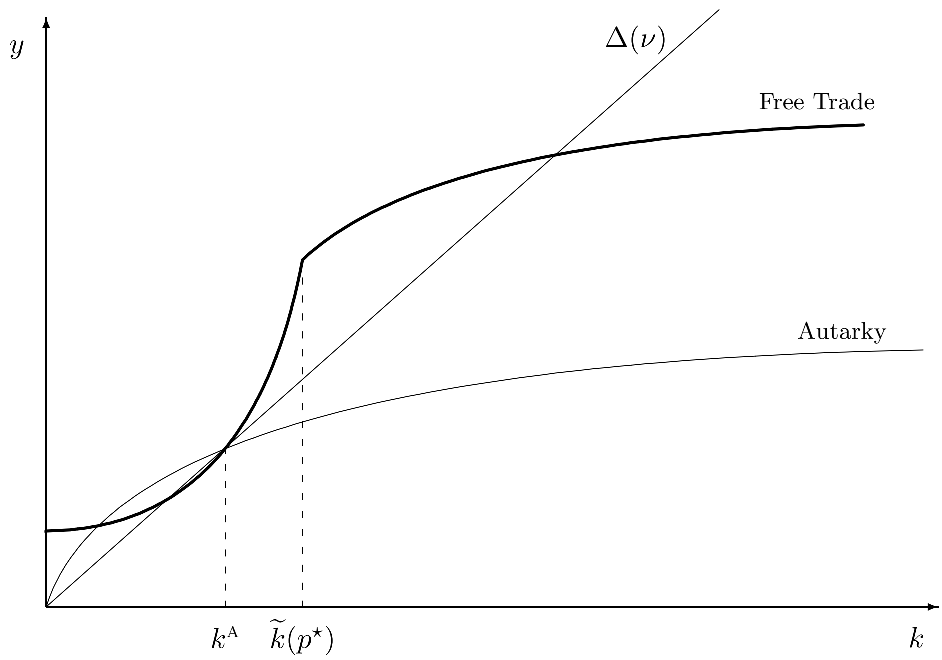
where $\widetilde{k}(p) \equiv \left[A(1-\alpha)(1+\varepsilon)\overline{\theta}^{1-\alpha}p\right]^{\frac{-1}{\alpha}}$ and $\Phi = \varphi^{\varphi}(1-\varphi)^{1-\varphi}$. This therefore makes it clear that what matters to the result is the openness of the economy rather than the externality case. In other words, the local convexity of the production frontier essentially originates in the globalization phenomenon and therefore trade.

Figure 3: Production function and local convexities

(a) No returns to specialization ($\varepsilon = 0$)



(b) Returns to specialization ($\varepsilon > 0$)



Note: $\Delta(\nu)$ denotes the combinations of y and k that are consistent with a steady state: $\Delta(\nu) = \{(y, k) \in \mathbb{R}_+ \times \mathbb{R}_+ : k = \nu y\}$.

increases (when $k \leqslant \widetilde{k}(p)$) causes countries with higher capital labor ratios to achieve higher productivity in sector 2. This is due to the fact that they produce good Z_2 more intensively. The gains to specialization effect thereby creates a local convexity in the aggregate production function as illustrated in Panel (b) of Figure 3. When the capital per efficient worker is greater than $\widetilde{k}(p)$, the economy fully specializes in the production of good 2 and exhausts all potential gains stemming from productivity gains specialization. The production function then regains its strict concavity back.

It should be noticed that as the economy increases the share of labor it allocates to the production of intermediate good directed to sector 2 (Q_2), the productivity in the sector increases due to specialization. This rise in productivity creates an additional incentive to reallocate production towards sector 2. Hence, the presence of gains to specialization decreases the threshold above which it is optimal to choose full specialization in the production of good 2. An important aspect to note from Figure 3 is that the presence of gains to specialization causes the slope of the aggregate production function to become steeper when evaluated at $k^{\star}$.

Having established that free trade can create a local convexity in a country's aggregate production function and increase the slope of this function, we now investigate the effect of this property on the steady state distribution of output–per–worker.

3.3 Free Trade and the Distribution of Output–per–Worker

This section shows how the process of globalization can give rise to both a hollowing–out in the distribution of output-per-worker and the emergence of twin–peaks. We consider two cases. We first consider the case of moderate gains to specialization, and show that even in this case twin peaks can emerge. We then show that when gains to specialization are large enough the model exhibits multiple steady state equilibria, none of which are consistent with a unimodal distribution.

3.3.1 The case of moderate gains to specialization

The case of moderate gains to specialization corresponds to a situation where $0 \leqslant \varepsilon < \alpha$. In such a case, the following proposition can be established on the steady state of each economy.

Proposition 4 *If $0 \leqslant \varepsilon < \alpha$, then under free trade all economies possess a unique steady state.*

Proposition 4 indicates that when the productivity gains to specialization are not too strong, then the local non–convexity in the production function economy does not cause the emergence of multiple equilibria, and hence the mapping between ν_i and y_i is well–behaved.

A nice feature of the model is that even in the case of free trade, it is again simple enough to allow for an analytical derivation of the whole distribution of output–per–worker across countries, as given below.

Proposition 5 *For $0 \leqslant \varepsilon < \alpha$, the steady state distribution of output–per–worker is given by*

$$\mu_y(y) = \begin{cases} \mu_s\left(g_1(y)\right) g_1(y) \left[\dfrac{\alpha\Phi - \varepsilon(1-\alpha)p^{1-\varphi}y}{\alpha y(p^{1-\varphi}y - \Phi)}\right] & \text{if } y \leqslant \widetilde{y}(p) \\ \mu_s\left(g_2(y)\right) \dfrac{1-\alpha}{\alpha} \dfrac{g_2(y)}{y} & \text{if } y > \widetilde{y}(p) \end{cases}$$

where $\widetilde{y}(p) = \frac{\Phi p^{\varphi-1}}{1-\alpha}$,

$$g_1(y) \equiv \frac{\Psi}{y}\left(\frac{yp^{1-\varphi} - \Phi}{\Phi}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}} \quad \text{with } \Psi \equiv \left(\frac{1-\alpha}{\alpha}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}} \left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-\frac{1}{\alpha}}$$

and

$$g_2(y) = \left[\Phi p^{\varphi} A\overline{\theta}^{1-\alpha}\right]^{-\frac{1}{\alpha}} y^{\frac{1-\alpha}{\alpha}}$$

Having characterized the general shape of the world distribution of income, we can now study the main implications of free trade on this distribution, and more particularly the effect of the presence of gains to specialization. This analysis is summarized in the next proposition.

Proposition 6 *If $0 < \varepsilon < \alpha$, allowing free trade will cause:*

1. *a hollowing–out of the middle of the distribution of output–per–worker in the sense that the density at the median will be lower under free trade;*
2. *the income distribution to exhibit at least two modes if ε is sufficiently close to α.*

Proposition 6 contains the main implications of the model. The first result on the hollowing–out of the middle of the distribution can be easily understood by referring to Section 1.2 and Figure 3. We saw in Section 1.2 that for the density of output–per–worker to reduce at the median of the distribution — in the absence of a change in the distribution of ν — it must be the case that the marginal effect of the propensity to accumulate, ν, on output–per–worker, y, evaluated at the median ν must increase. Since we saw in Figure 3 that the slope of the aggregate production function evaluated at $k^\star$ increases under free trade, it should come as no surprise that the marginal effect of ν on y (evaluated at median ν^m) increases under free trade. More formally,

$$\left.\frac{\partial y^{\mathrm{FT}}}{\partial \nu}\right|_{\nu=\nu^m} \geqslant \left.\frac{\partial y^{\mathrm{A}}}{\partial \nu}\right|_{\nu=\nu^m} \text{ if } 0 < \epsilon < \alpha$$

where A and FT respectively denote autarky and free trade. The intuition underlying this result can be found in the existence of gains to specialization. This is illustrated in Figure 3, where, in addition to the aggregate production functions, we also included the line that represents the combinations of y and k that are consistent with a steady state. This locus of points is derived directly from the accumulation equation and corresponds to $y = k/\nu$. A steady state equilibrium in this model is simply a crossing point between this locus and the aggregate production function. The figure then illustrates how our hollowing–out result in Proposition 6 depends crucially on the presence of gains to specialization. In effect, in the absence of gains to specialization, $\epsilon = 0$, the model predicts that the free trade would leave the density around the middle of the distribution unchanged. Indeed, when there are no gains to specialization, the aggregate production function is linear for values of the capital stock below the full specialization threshold $\widetilde{k}(p)$. This implies that, around the median, the sensitivity of output–per–worker with respect to ν is not affected by the opening of trade, since the new aggregate production function is tangent to the autarkic production function. This can again be inferred from panel (a) of Figure 3 since the free trade is shown not to affect the slope of the aggregate production function at the median. Conversely, when there exist gains to specialization, $\varepsilon > 0$, not only can the economy avoid decreasing returns by specializing in the production of type 2 good, but it also becomes more productive. Therefore, any change in the accumulation behavior exerts a larger effect on output–per–worker — *i.e.* the sensitivity of output–per–worker to ν is greater. This can be seen from panel (b) of Figure 3 as free trade makes the aggregate production function to display local convexity in a neighborhood of $k^\star$.

The second issue addressed in Proposition 6 relates to whether globalization can create a twin–peaked distribution. Although this issue is closely related to the notion of a hollowing–out of the distribution, it is distinct since it requires that mass away from the middle of the distribution either increases substantially more – or decreases substantially less – than mass around the middle. In the proof of Proposition 6 we show that as ϵ approach α from below, the density near the middle goes to zero and therefore there necessarily emerges more than one mode in the distribution. This type of results can be inferred graphically by noticing that as the gains to specialization get strong enough, the aggregate production function becomes close to tangent to one of the steady state locus. In this case, the sensitivity of output–per–worker to ν becomes infinite when evaluated at this point and the associated density goes to zero.

Although Proposition 6 does not indicate that the process of globalization will necessarily cause twin peaks to emerge, it does show that such a change can arise in the model.

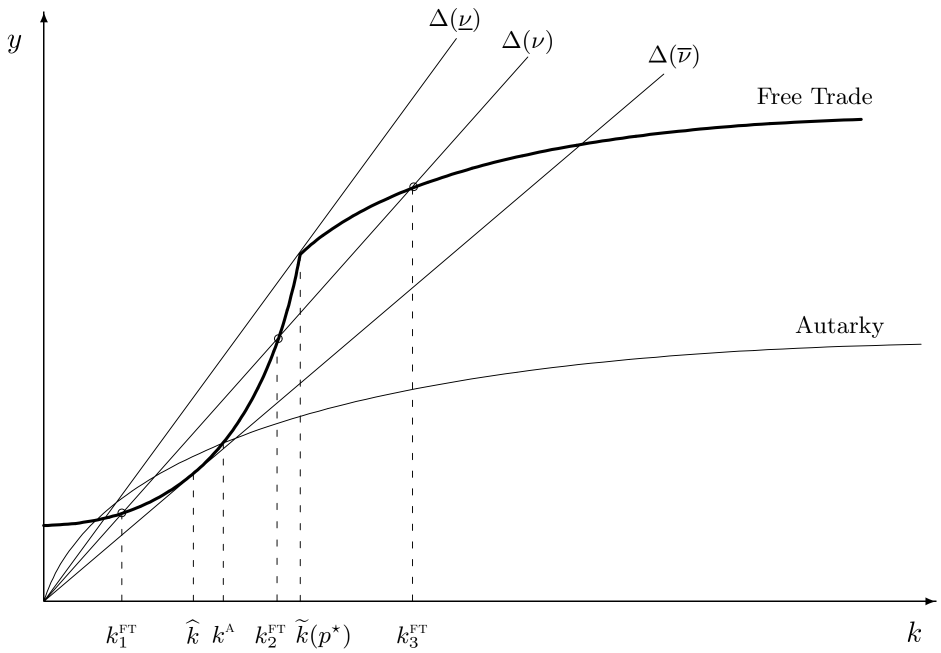
3.3.2 The case of strong gains to specialization

In this section we briefly examine the effects of allowing free trade when the gains to specialization are strong, in the sense that $\epsilon > \alpha$. In this case, as stated in Proposition 7, the possibility of multiple equilibria arises.

Proposition 7 *For* $\alpha < \varepsilon < \frac{\alpha}{1-\alpha}$ *there exist two levels of* ν, $\underline{\nu}$ *and* $\overline{\nu}$, *such that if* $\nu \in [\underline{\nu}; \overline{\nu}]$, *an economy admits three steady states level of output–per–worker, otherwise it admits a unique steady state.*

Proposition 7 establishes that when the productivity effect of specialization is strong enough there exists a range of propensities to accumulate capital which give rise to multiple steady states. This is illustrated in Figure 4. As can be seen from the figure, when the propensity

Figure 4: Multiplicity



Note: $\Delta(\nu)$ denotes the combinations of y and k that are consistent with a steady state: $\Delta(\nu) = \{(y,k) \in \mathbb{R}_+ \times \mathbb{R}_+ : k = \nu y\}$.

to accumulate capital is low ($\nu < \underline{\nu}$), the steady state is unique and corresponds to a partial specialization with a relatively low capital labor ratio. Conversely, when the propensity to accumulate capital is very high ($\nu > \overline{\nu}$), the high tendency to accumulate capital together with the gains to specialization lead such an economy to specialize totally in the production of the capital intensive good. Furthermore, the figure makes it clear that there exists a cone defined by the two propensities to accumulate capital $\underline{\nu}$ and $\overline{\nu}$ within which an economy admits three steady states. The two lowest ones lie on the convex part of the production frontier and correspond to a partial specialization situation. The third steady state lies on the concave part of the frontier and illustrates that such an economy also admits a steady state with full specialization. Indeed, when specialization enhances productivity, any increase in the capital stock exerts two effects. First, an increase in the capital labor ratio triggers a reallocation of labor toward the capital intensive good sector in order to avoid the effects of decreasing returns. Second, it drives productivity upward in that sector through the specialization mechanism. When ε is high enough, this latter mechanism is strong enough to induce a faster pace of accumulation, which further increases productivity. Hence it is the complementarity between accumulation and gains to specialization which gives rise to multiple equilibria.

Proposition 8 documents the stability properties of the three steady state.

Proposition 8 *Under free trade, for $\alpha < \varepsilon < \frac{\alpha}{1-\alpha}$, the two outside steady states are stable, such that for any given propensity to accumulate capital $\nu \in [\underline{\nu}; \overline{\nu}]$ we have*

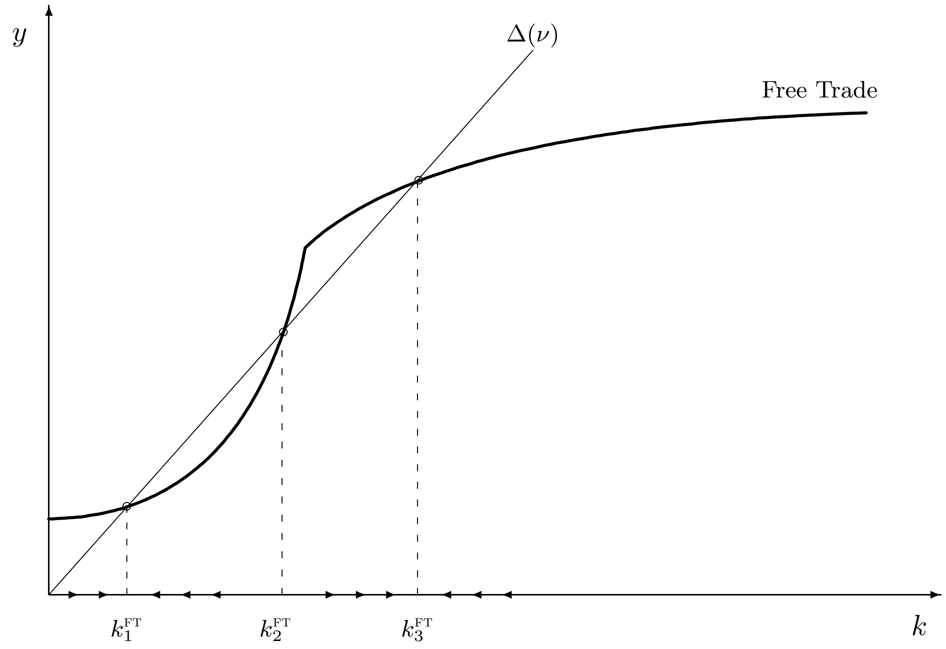
$$\begin{cases} \textit{If } k_0 < k_2^{\mathrm{FT}} \textit{ the economy converges to } k_1^{\mathrm{FT}} \\ \\ \textit{If } k_0 > k_2^{\mathrm{FT}} \textit{ the economy converges to } k_3^{\mathrm{FT}} \end{cases}$$

where $k_1^{\mathrm{FT}}, k_2^{\mathrm{FT}}, k_3^{\mathrm{FT}}$ are the three steady states in ascending order.

Proposition 8 shows that out of the three steady states, only two — the external — are stable. This phenomenon is represented in Figure 5, which illustrates how an economy that starts with a high enough capital per efficient worker ($k_0 > k_2^{\mathrm{FT}}$) will eventually fully specialize and end up with a high capital–labor ratio. In other words, the economies that start above their central steady state, for example because of their autarkic position, will greatly benefit from free trade. Once again the economic intuition for this result is simple and is found in the standard Hecksher and Ohlin argument. If an economy starts with a high capital–labor ratio, it will allocate under free trade a larger share of its workers to the sector that is capital intensive. This induces gains to specializations that raise the marginal efficiency of capital, which then leads the economy to accumulate more capital and therefore reinforces the initial effect. The economy then eventually reaches the highest steady state.

Conversely, let us now consider an economy which is initially poorly endowed in capital ($k_0 < k_2^{\mathrm{FT}}$). Proposition 8 makes it clear that such an economy will not be able to take advantage of specialization, and will end up in the lowest steady state despite the fact it may be characterized by the same propensity to accumulate capital. In other words, the model exhibits hysteresis if $\nu \in [\underline{\nu}, \overline{\nu}]$.

Figure 5: Stability



Note: $\Delta(\nu)$ denotes the combinations of y and k that are consistent with a steady state: $\Delta(\nu) = \{(y, k) \in \mathbb{R}_+ \times \mathbb{R}_+ : k = \nu y\}$.

In the case of strong returns to specialization, it becomes even more apparent how the process of globalization can cause a hollowing–out of the distribution and the emergence of twin peaks. In particular, in this case, there is a whole range of values of output–per–worker, within the support of steady values of output–per–worker, which are not steady state. Hence, the distribution of output–per–worker in that case would become completely hollowed–out and therefore there would necessarily emerge at least two modes.

4 Empirical Evidence

The model presented in the previous section suggests that the opening up of trade could be the central force underlying the observed change in the cross–country distribution of output–per–worker. In this section we want to present evidence in support of the particular mechanism which induces a change in distribution in our model.

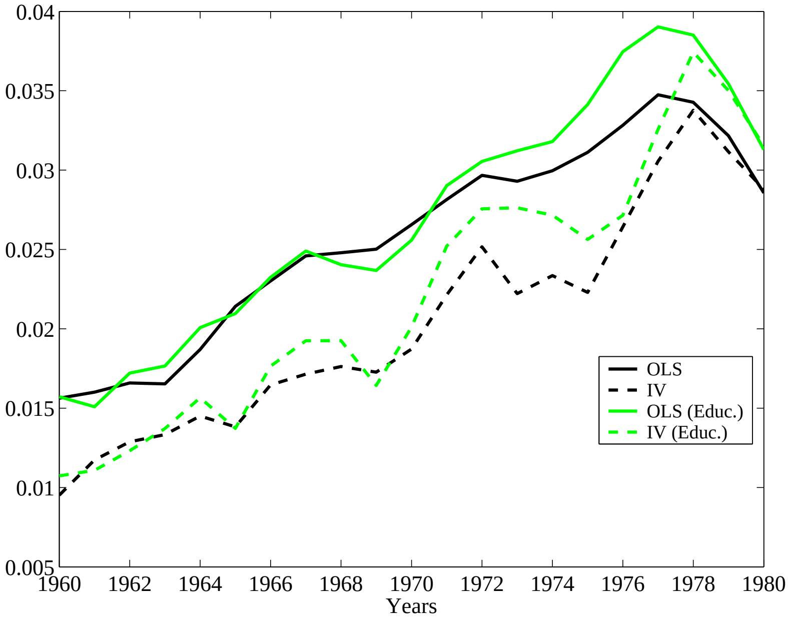
4.1 Assessing the model's direct implications

One of the model's main implication is that, over a period of globalization, we should observe the emergence of an abnormally high relationship between a country's propensity to accumulate capital and its growth rate. We now investigate empirically the relevance of this prediction of the model. To this end, we begin by reporting a set of cross–country regressions relating growth in output–per–worker to its initial level and to its propensity to accumulate capital, as captured by $\mathcal{K}$, which will be our empirical counterpart for $\log(\nu)$

$$\mathcal{K}_{i,t} = \log\left(\frac{s_{i,t}}{(1+n_{i,t})(1+\gamma)-(1-\delta)}\right)$$

In the construction of $\mathcal{K}_{i,t}$, we follow Mankiw, Romer and Weil [1992] and impose a 3% annual depreciation rate of capital ($\delta = 0.03$) and a 2% rate of technological growth ($\gamma = 0.02$).[18] Furthermore, the country's saving rate is proxied by its' investment rate ($s_{i,t} = (i/y)_{i,t}$), and set $n_{i,t}$ to the country's annualized rate of growth of labor force. Note that even though we kept n constant in the model, we are allowing it to vary in this section. The main observation

Figure 6: Effect of propensity to accumulate capital (Rolling Regressions)



that we want to highlight is the extent to which the role of this measure of the propensity to accumulate capital has changed over the 1960–1998 period. To address this issue, we begin by reporting estimates associated with a series of very simple rolling regressions. Our basic regression consists of regressing the growth of output per worker[19] on the average value of $\mathcal{K}$

[18] Setting these parameters to alternative values leaves our main results unaffected.
[19] See appendix A for a definition of the variables

and the initial level of output per worker, with each variable measured in relation to the given window. Each regression is estimated by OLS and IV. We also report results when we control for educational investment. Our measure of educational investment is a weighted average of school enrollment for primary, secondary and tertiary degrees of education, as documented in Barro and Lee [1993].[20] This measure departs from the Mankiw et al. [1992] specification of educational investment and rather follows Klenow and Rodrìguez-Clare [1997] who argue that focusing on this wider educational investment class is more appropriate. However, we should note that our findings are very robust to varying the measure of educational investment. In both case, we will report results from OLS and IV estimation.

We chose a 19 years window, which corresponds to half of our sample length. Figure 6 reports the estimated coefficients on $\mathcal{K}$ obtained from rolling regressions over the whole sample, with each variable measured in relation to the given window.[21] As can be seen in the figure, the importance of this measure of a country's propensity to accumulate capital increase substantially over the period. In effect, it almost tripled irrespectively of the estimation procedure or the inclusion of additional variables (such as education, or scale effects).

In what follows, we will document the robustness of this observation along several dimensions. First, we will allow the two components of $\mathcal{K}$ to enter the regression separately. That is, we will run the regression allowing the country's investment rate and its labor growth to enter separately, instead of bundling it together in $\mathcal{K}$. If the forces emphasized in our model are relevant, we should observe the importance of each of these factors to increase over the process of globalization. Once again, we also allow an additional regressor our measure of educational investment.

Figure 7 reports the estimated coefficients on population growth and investment rate obtained from rolling regressions over the whole sample, with each variable measured in relation to the given window. The figure displays four series corresponding to the choice of estimation procedure (OLS or IV) and inclusion or not of our measure of educational investment. In the IV procedure, the variables used to instrument the investment rate and population rate of growth are the initial level of the investment rate, the average consumption rate over the sub–sample (which proxies the saving rate), and the average growth rate of population over the 15 first periods of the sub–

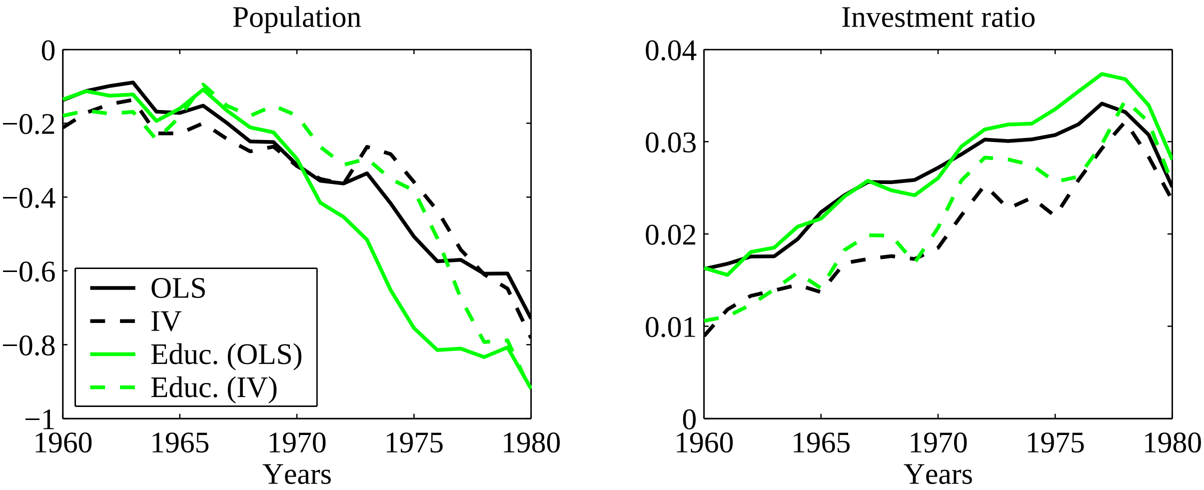
[20] The weights for calculating this measure of educational investment correspond to the average number of years spent in each degree

$$Educ \equiv \log\left(\frac{6\times E^P+6\times E^S+4\times E^T}{16}\right)$$

where E^P, E^S and E^T denote respectively the enrollment rate in the primary, secondary and tertiary sector.

[21] We estimate a series of regressions with the first regression using observations 1 through 19, the second regression using observations 2 through 20, etc... We also explored window sizes between 15 to 25 years, all of which led to similar conclusions. In Beaudry and Collard [2002] this approach was used to examine changes in the growth process among OECD country's.

Figure 7: Importance of Growth factors: Rolling regressions



sample when we consider our baseline regression. The latter instrument is aimed at correcting for the potential endogeneity of fertility within each window, as it takes at least 15 years for fertility to affect the labor force. Hereafter, this set of instrument will be referred to as IV*. When Education is added, IV* is completed by the years of schooling. This is a stock rather than a flow measure of education and is hence less sensitive to endogeneity over short periods. The plain dark line in the figure corresponds to OLS estimation of our benchmark equation where output per worker is regressed on adult-population growth, the investment rate and the initial level of output per worker. The dashed dark line refers to the IV estimation of this benchmark regression. Grey lines corresponds to OLS (plain) and IV (dashed) estimation of our regression where education is introduced.

As can be seen from the figure, there has been a continuous increase in the magnitude of both the population growth effect and the investment effect, even though the latter effect decreases slightly at the end of the period. This suggests that the role of the two traditional factors associated with capital deepening increased since 1960. For example, in the end of the sample — for example the 1980–1998 window — our estimates of the population growth effect is more that 3 times as much as it was in the 1960–1978 period. Likewise, the coefficient of the investment rate is multiplied by about 2.5 over the same period. Furthermore, the size of the coefficients over the different windows is relatively robust to the estimation method and to whether or no we control for educational investment.

In order to examine the statistical significance of this increase, Table 2 reports a series of OLS and IV estimates for the regression of the average yearly growth of output per worker on the initial log-level of output per worker, the average yearly growth in labor force, the log of the average rate of investment and our measure of education investment. Estimates are reported for both the first (1960–1978) and last (1980–1998) window of our sample. Averages for both the regressors and regressants are taken over the respective windows. Let us focus on our

Table 2: Growth regressions

	OLS		IV1		OLS		IV2	
	60–78	80–98	60–78	80–98	60–78	80–98	60–78	80–98
const.	0.185	0.203	0.149	0.199	0.218	0.203	0.197	0.227
	(0.025)	(0.026)	(0.027)	(0.027)	(0.032)	(0.036)	(0.036)	(0.046)
y_0	-0.013	-0.013	-0.011	-0.013	-0.017	-0.013	-0.015	-0.015
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
n	-0.137	-0.730	-0.203	-0.773	-0.082	-0.830	-0.117	-0.830
	(0.164)	(0.152)	(0.172)	(0.156)	(0.164)	(0.149)	(0.171)	(0.153)
i/y	0.016	0.025	0.009	0.024	0.014	0.028	0.006	0.027
	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)
Educ.	–	–	–	–	0.010	-0.007	0.015	0.002
					(0.006)	(0.009)	(0.008)	(0.014)
R^2	0.36	0.54	–	–	0.37	0.59	–	–
$\mathcal{Q}$(Total)	16.436	[0.001]	19.142	[0.000]	26.670	[0.000]	28.502	[0.000]
$\mathcal{Q}(y_0)$	0.005	[0.943]	0.743	[0.389]	0.759	[0.384]	0.000	[0.989]
$\mathcal{Q}(n, i/y)$	13.605	[0.001]	17.332	[0.000]	21.997	[0.000]	21.430	[0.000]
$\mathcal{Q}$(Educ.)	–	–	–	–	2.668	[0.102]	0.686	[0.407]

Note: Standard errors in parenthesis, p–values in brackets. IV1: the variables used in instrument i/y and n are the initial level of (i/y), the average (c/y) over the sub–sample, and the average growth rate of population over the 15 first periods of the sub–sample. 75 observations. IV2: the set of instruments is that of IV1 completed by the years of schooling to instrument Educ. 68 observations.

benchmark regression. Note first that the convergence parameter has remained stable over the two windows. Indeed, the stability test of this parameter, which we denote by $\mathcal{Q}(y_0)$ in the table, does not indicate a rejection of stability at conventional 5% level, whatever the estimation method (p–value=0.943 using OLS, 0.389 using IV).[22] Conversely, the estimates reported in the four first columns of Table 2 clearly show that the magnitude of the population growth and investment/output ratio effects increased quite dramatically between the 1960–1978 and 1980–1998 periods. In particular, over the first sub–period (1960–1978), the coefficient measuring the impact of population growth on the growth process was insignificant at the conventional 5% level whatever the estimation method. For example, when the benchmark equation is estimated by IV, this coefficient amounted to -0.203, implying that a negative 1% differential in adult-population growth in 1960 would have been translated into a positive differential in output per worker of about 4% by 1978. In the second sub–period, this coefficient became significant and rose to -0.773, implying the same negative 1% differential in the rate of growth would have been translated in a positive differential in output per worker of about 14% by 1998 relative to 1978. A

[22] Note that all our stability tests are performed allowing for residuals to be correlated within countries over the two samples and for the variances to differ.

similar pattern is found in the behavior of the coefficient affecting the investment/output ratio. Over the first window, the IV coefficient affecting investment was significant and amounted to 0.009. It rose to 0.024 over the 1980–1998 window. Note that $\mathcal{Q}(n, i/y)$, the test statistic associated with the null hypothesis of stability in these parameters across these windows strongly leads to rejection of the null of stability (p–value=0.000) for both methods. The test of the overall stability of all three coefficients ($\mathcal{Q}$(Total)) is also clearly rejected.

The introduction of education in the regression does not alter our previous findings. In fact, these are even reinforced. Indeed, if we consider the results obtained through IV estimation, the coefficient affecting labor force growth is now multiplied by a factor of about 7 between the two windows shifting from -0.117 over the 1960–1978 window to -0.830 over the 1980–1998 window. A very similar pattern is found for the investment rate whose coefficient is multiplied by 4.5 between the two windows. Accordingly, the joint stability of the factors associated with capital accumulation is again strongly rejected by the data (p–value=0.000). It should be noted that we tested and could not reject the hypothesis that the increased importance of the investment rate was the same as the increased importance of the effect of labor force growth. This is precisely what would be expected if the forces emphasized in our theoretical model were present.

We now want to go one step further and follow recent work of Bernanke and Gürkaynak [2001] who examined whether total factor productivity growth across countries is exogenous with respect to a country's accumulation path, or whether conversely, a country's accumulation affects TFP growth over certain periods, thereby suggesting some form of endogenous productivity gains. In effect, our model suggests that not only we should see a rise in the importance of capital accumulation forces on output–per–worker over the process of globalization, but we should also see this relation emerged for measured productivity. To address this issue, we build TFP series for each country by subtracting from output–per–worker growth the share weighted contribution of physical capital–per–worker growth. Since we do not have good share data for all these countries, we build TFP series using a share of capital of 0.35 — which is close to the average share of capital across–countries.[23] Capital series are constructed using the perpetual inventory method.[24] Table 3 reports a series of OLS and IV estimates for the regression of the average yearly growth of TFP on its initial level, the average yearly growth of the labor force, the log of the average rate of investment and our measure of education investment. Estimates are reported for both the 1960–1978 and 1980–1998 window. When IV regressions are considered, the set of instrument is IV* (completed with years of schooling when education is introduced).

[23] In Beaudry et al. [2002], we have assessed the robustness of these results considering a share of capital of 0.5, which is close to the highest capital shares reported in Bernanke and Gürkaynak [2001]. The results are indeed very similar.

[24] Details of this construction are given in appendix A.

Table 3: TFP regressions

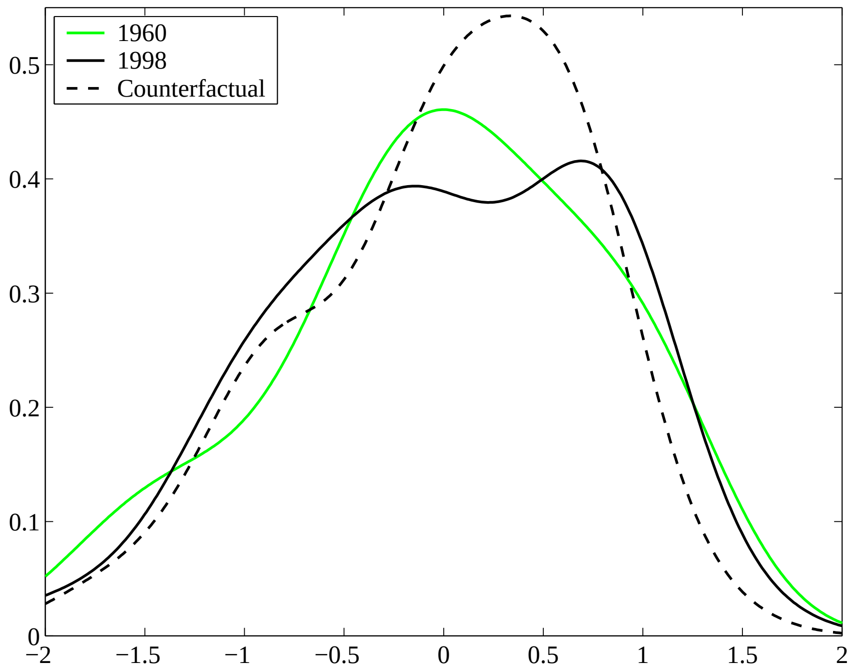
	OLS		IV1		OLS		IV2	
	60–78	80–98	60–78	80–98	60–78	80–98	60–78	80–98
const.	0.118	0.107	0.111	0.110	0.135	0.098	0.142	0.109
	(0.016)	(0.020)	(0.017)	(0.021)	(0.020)	(0.027)	(0.023)	(0.033)
TFP_0	-0.014	-0.011	-0.014	-0.012	-0.017	-0.009	-0.019	-0.011
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
n	-0.073	-0.597	-0.122	-0.616	0.001	-0.684	-0.010	-0.678
	(0.125)	(0.122)	(0.128)	(0.124)	(0.125)	(0.124)	(0.130)	(0.130)
i/y	0.009	0.012	0.005	0.012	0.006	0.015	0.001	0.015
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Educ.	–	–	–	–	0.007	-0.007	0.013	-0.005
					(0.004)	(0.007)	(0.006)	(0.011)
R^2	0.38	0.45	–	–	0.38	0.51	–	–
$\mathcal{Q}$(Total)	11.885	[0.009]	14.505	[0.002]	20.673	[0.000]	22.893	[0.000]
$\mathcal{Q}(\text{TFP}_0)$	0.415	[0.519]	0.262	[0.609]	2.076	[0.150]	1.653	[0.199]
$\mathcal{Q}(n, i/y)$	11.130	[0.004]	13.692	[0.001]	19.235	[0.000]	17.369	[0.000]
$\mathcal{Q}$(Educ.)	–	–	–	–	2.825	[0.093]	1.806	[0.179]

Note: Standard errors in parenthesis, p–values in brackets. IV1: the variables used in instrument i/y and n are the initial level of (i/y), the average (c/y) over the sub–sample, and the average growth rate of population over the 15 first periods of the sub–sample. 75 observations. IV2: the set of instruments is that of IV1 completed by the years of schooling to instrument Educ. 68 observations.

The results presented in Table 3 are intriguingly in lines of those obtained using output–per–worker. If we first focus on the benchmark regression, it appears that the coefficient related to initial TFP over each window is found to be significant and negative, whatever the window and the estimation method. This indicates that there has probably been technological catch–up over each period. But, this catch–up parameter remained stable over the two windows. The stability test of this parameter ($\mathcal{Q}(TFP_0)$) in the table does not indicate a rejection at a conventional 5% level, whatever the estimation method (p–value=0.519 using OLS, 0.609 using IV). Conversely, the estimates of the population growth and investment/output ratio effects clearly indicate that they increased substantially between the 1960–1978 and 1980–1998 periods. For example, in the first sub-period (1960–1978), the coefficient measuring the impact of population growth on the growth process was not significant at the conventional 5% level whatever the estimation method and amounted to -0.073 with OLS and -0.122 with IV estimation. In the last window, this coefficient became significant and rose to, respectively, -0.597 and -0.616. Similarly, over the first window, the IV coefficient affecting investment amounted to 0.005 (0.009 with OLS) and rose to 0.012 (0.012 with OLS) over the 1980–1998 window. Accordingly, the stability test of these parameters ($\mathcal{Q}(n, i/y)$) across the two windows strongly leads to a rejection of the null of stability (p–value=0.000) for both methods. The test of the overall stability of all three coefficients ($\mathcal{Q}$(Total)) is also clearly rejected.

Just as in the case of output–per–worker, the introduction of education in the regression strengthens our results. For example, if we focus on the IV estimates, it appears that the coefficient affecting labor force growth has increased enormously (from -0.010 over the 1960–1978 window to -0.678 over the 1980–1998 window). Likewise, the investment ratio effect is hugely increased between the two windows. Accordingly, the joint stability of the traditional growth factors across the two windows is strongly rejected by the data (p–value=0.000). The results reported in Table 3 therefore provide support to the model's prediction that over a period of globalization, countries's with high propensities to accumulate capital should exhibit above average productivity growth. In order to illustrate the extent to which the observed increase in the importance

Figure 8: Counterfactual Distribution



of capital accumulation forces may be relevant for understanding the observed change in the cross-country distribution of output–per–worker, in Figure 8 we overlay the actual distribution observed in 1998 and 1960 with a counterfactual distribution designed to control for the observed increase in the regression parameters associated with the investment rate and the growth of labor force over the period 1978–1998 relative to the period 1960–1978. More precisely, we create the counterfactual distribution of (log) output–per–worker, denoted y^c_{98} as follows:[25]

$$y^c_{i,98} = y^{78}_{i,78} + (\beta_1 - \beta_2)X_{i,98}$$

where $X_{i,98}$ are the values of the investment rate and the labor force growth for country i over

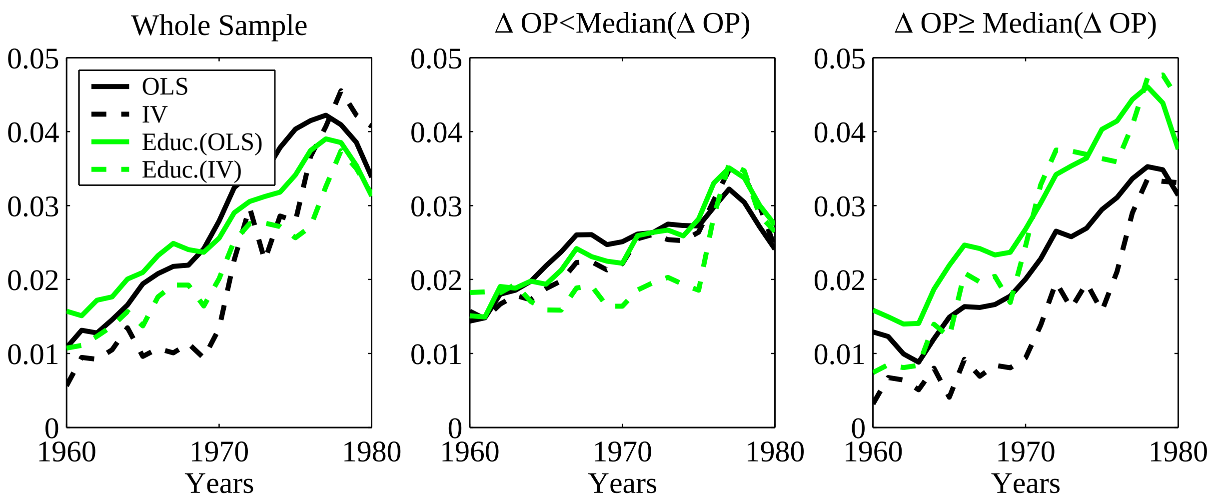
[25]Greater details on the construction of the counterfactual distribution can be found in BCG.

the period 78–98, β_1 is the estimated effect of these two variables over the period 1960–1978 and β_2 is the estimated effect over the period 1978–1998. We can see from the figure that the estimated changes in the β's can account for all the hollowing of the distribution. This indicates that the observed changes in the the role of capital accumulation forces are of the right order to explain the observed change in the distribution, thereby providing support for the central mechanism of the model.

4.2 Looking for a more direct link with trade

Up to now, our empirical evidence has focused on highlighting changes in the growth process that have arisen during a period where there has been a substantial increase in international trade. However, in itself, this does not imply that the change in the growth process is related to globalization. It only supports the theoretical possibility presented in the model. One way to address this issue more directly is to look at whether the change in the increased importance of capital for growth is more pronounced among countries where the process of globalization is most apparent. To this end, we now divide our sample into two sub–samples according to their growth of international trade, where the growth of international trade is measured by the change in the ratio of exports plus imports to GDP (this is the openness measure in the World Penn Tables). The first sample (ΔOP <Median(Δ)) is the set of countries where the growth in international trade between 1960 and 1998 is below the median growth rate of the entire sample. The second sample ($\Delta OP \geqslant$Median(Δ)) is comprised of countries with growth in international trade above the median. For each of the sub–samples, we once again run a set of 20 rolling regressions over 19 years windows starting in the 1960. Figure 9 plots the coefficients

Figure 9: Rolling Regressions: Openness



(on the same scale) on our accumulation variable ($\mathcal{K}$) for the whole sample and the two sub–

samples. In each case, we plot four series of coefficients corresponding to different estimation techniques and different control variables. The initial value of output–per–worker is included in each regression. The differences between the series correspond to whether or not we included education as an additional regressor and whether we estimate by instrumental variables (using the same instruments as discussed previously) or by OLS. To help judge statistical significance, Table 4 reports the point estimates associated with the first and last window for the base case, where the estimation is by OLS and where we do not include an education control. What is

Table 4: Openness (OLS)

	Total		ΔOP <med(ΔOP)		$\Delta OP \geqslant$med(ΔOP)	
	60–78	80–98	60–78	80–98	60–78	80–98
Const.	0.147	0.126	0.094	0.150	0.165	0.092
	(0.022)	(0.023)	(0.029)	(0.033)	(0.034)	(0.029)
y_0	-0.014	-0.014	-0.009	-0.017	-0.016	-0.012
	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
$\mathcal{K}$	0.016	0.031	0.014	0.027	0.014	0.041
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)
R^2	0.34	0.55	0.33	0.49	0.31	0.61
$\mathcal{Q}(\mathcal{K})$	13.006	[0.000]	4.103	[0.043]	16.399	[0.000]

Note: Standard errors in parenthesis, p–values in brackets. 75 observations.

clear from Figure 9 is that the increased importance of physical capital for growth documented previously appears to arise most prominently among the countries which have witnessed the greatest increase in international trade. This is very much in line with the spirit of our model. In effect, for the countries with a high growth in international trade, the effects of capital intensity, as measured by $\mathcal{K}$, appear to have tripled or quadrupled, which is huge. For instance, as shown in the last row of Table 4, the change in importance of physical capital for growth is significant across the two sub–periods for both samples. The coefficient on $\mathcal{K}$ doubled — increasing from 0.014 to 0.027 — for the economies where the process of globalization was less apparent, while the increase was more substantial among the economies that experience more growth in trade — increasing from 0.014 to 0.041. Even though these last results, in conjunction with our previous ones, still cannot unambiguously confirm causality running from globalization to changes in the world distribution, we believe that they provide considerable support for the idea and mechanism presented in our model.

5 Discussion

In this section we briefly discuss two extensions of our model. Our goal is to illustrate that our theoretical results our robust to changes in some of our simplifying assumptions. In particular, we will show that the results are robust to allowing both types of goods to be produced by means of both capital and labor, while maintaining that the type 2 good is more capital intensive. We also show that allowing for international capital flows does not invalidate our results, in fact they accentuate them.

Let us first extend our model to the case where capital is used in the two sectors. The Z_1 good is now produced using intermediate goods M_1 and capital K according to the production function:

$$Z_{1,i,t} = A_1(K_{1,i,t})^{\alpha_1} M_{1,i,t}^{1-\alpha_1} \tag{22}$$

where M_1 is a composite of intermediate goods $Q_{i,t}(\ell)$, such that

$$M_{1,i,t} = \left(\int_0^1 Q_{1,i,t}(\ell)^\rho \mathrm{d}\ell \right)^{\frac{1}{\rho}} \text{ with } \rho < 1 \tag{23}$$

The Z_2 good is assumed to be produced using intermediate goods and capital according to:

$$Z_{2,i,t} = A_2((K_{2,i,t})^{\alpha_2} M_{2,i,t}^{1-\alpha_2} \tag{24}$$

where $\alpha_2 \geqslant \alpha_1$, indicating that sector 2 is more capital intensive than sector 1, and M_2 is composed by the following quality weighted sum[26] of intermediate goods $Q_{i,t}(\ell)$.

$$M_{2,i,t} = \left(\int_0^1 \left(\theta_{i,t}^{\mathrm{s}}(\ell) Q_{2,i,t}(\ell)\right)^\rho \mathrm{d}\ell \right)^{\frac{1}{\rho}} \text{ with } \rho < 1 \tag{25}$$

In (25), $\theta_{i,t}^{\mathrm{s}}(\ell)$, $\ell \in (0, 1)$, denotes the *subjective* quality assigned by the firm in sector 2 to the intermediate good $Q_{i,t}(\ell)$ sold by firm ℓ.

Given the demand from both sector 1 and 2, the problem of firm $\ell \in (0, 1)$ can be seen as maximizing profits by choosing prices and allocating labor to production directed to the different sectors. Denoting by $\sigma_{i,t}(\ell) \in [0, 1]$, the share of total labor, $L_{i,t}$, allocated to the production of good 1, the program of an intermediate firm is given by

$$\max_{\{L_{i,t}(\ell), p_{1,i,t}(\ell), p_{2,i,t}(\ell), \sigma_{i,t}(\ell)\}} p_{1,i,t}(\ell)\Gamma_t\sigma_{i,t}(\ell)L_{i,t}(\ell) + p_{2,i,t}(\ell)\Gamma_t(1-\sigma_{i,t}(\ell))L_{i,t}(\ell) - W_{i,t}L_{i,t}(\ell) \tag{26}$$

taking the form of the demand into account.

[26] Note that the index of intermediate goods is characterized by the same elasticity in the two sectors. This restriction can be relaxed without any difficulty. This would only complicate the analysis by multiplying sub–cases without altering much the results.

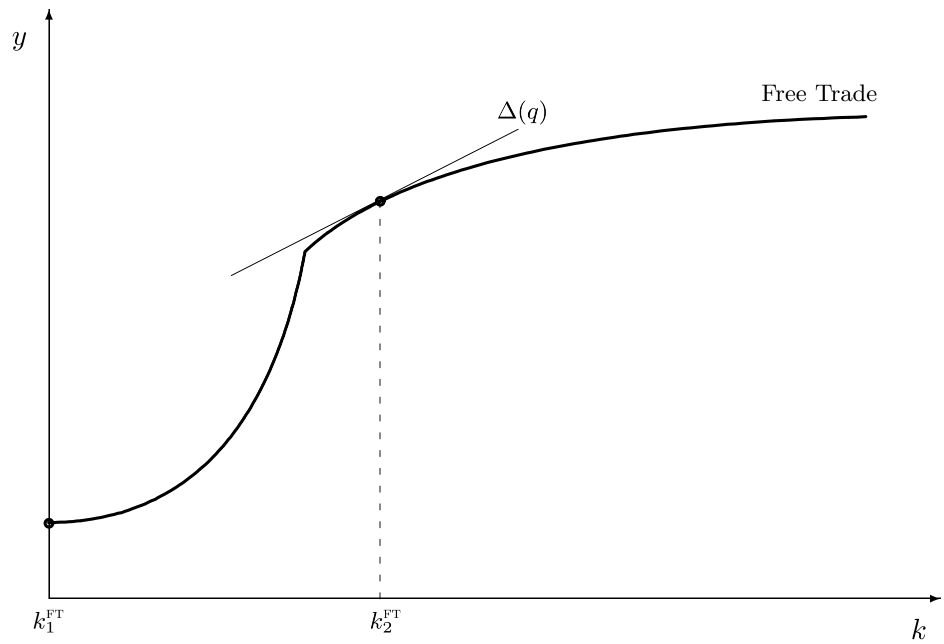
As before, an economy can choose to fully specialize in the production of type 1 good, to fully specialize in the production of type 2 good, or allocate its inputs in both sector to produce both type of goods. We show in the next proposition that our main result regarding the emergence of a local convexity survives this extension.

Proposition 9 (Local convexity) *If* $0 \leqslant \varepsilon < \frac{(\alpha_2-\alpha_1)^2}{\alpha_2(1-\alpha_2)(1-\alpha_1)}$*, then, under free trade, output–per–worker displays local convexity.*

Proposition 9 states that, as long as ϵ is positive (and not too large) the production function displays a local convexity stemming from the productivity effects of specialization. In particular, when there exist gains to specialization, $\varepsilon > 0$, the economy *(i)* can avoid decreasing returns by specializing in the production of type 2 good, and *(ii)* becomes more productive. Therefore, any change in the accumulation behavior exerts a larger effect on output–per–worker, implying that the model supports the emergence of a twin peaked distribution of output–per–worker.

One may also be concerned with the fact that, in our model, the world economy consists of a collection of economies that do not trade in capital markets. How do our explanation to the twin–peaks phenomenon survive the introduction of international capital flows? In fact, if we

Figure 10: Perfect capital mobility



Note: The line $\Delta(q)$ corresponds to the total input cost at market prices. For instance, $\Delta(q) = w_2 + q.k$, where w_2 corresponds to the wage when the economy fully specializes in the production of a type 2 good.

now take international capital flows into account, the forces favoring polarization are actually reinforced. Indeed, when capital is perfectly mobile across countries, perfect competition on the capital market forces real returns to capital to be equalized across country. Then, as illustrated in Figure 10, the distribution of output–per–worker must collapse to two mass points (k_1^{FT} and k_2^{FT}) as either countries specialize in the production of type 1 good for which capital is not used (k_1^{FT}), or specialize in the production of type 2 good (k_2^{FT}).[27] Hence, the opening up of both trade and international capital markets in the presence of gains to specialization would give rise to an extreme version of the twin–peaks phenomenon.

6 Concluding remarks

Over the last forty years, the distribution of output–per–worker across countries has hollowed–out substantially. Two major classes of explanation can be proposed for such a phenomenon. First, it may be the case that the distribution of the exogenous forces that drive the dynamics of output–per–worker has changed over time to yield a bi–modality in the distribution of output–per–worker. Second, it may be that the relationship between these exogenous forces and output–per–worker has changed. In this paper, we presented theory and evidence in support of an explanation of the second type. In particular, we developed a simple dynamic general equilibrium multi–country model where, in the absence of international trade, each economy evolves like in the Solow growth model. However, in the presence of free trade, conforming with the empirical evidence, countries polarize in a manner influenced by their propensity to accumulate physical capital. The main element of the model which drives the result is the presence of productivity gains to specialization. We showed how this effect of specialization induces, under free trade, a local convexity that causes the distribution of output–per–worker across countries to hollow–out in the middle and loose its original uni–modal shape. Although we did not presented an organization theoretic explanation to the form of the firm level gains to specialization used in the model, we believe that our analysis highlights the potential of this approach to give new insight for understanding how the world economy is being affected by globalization.

[27]There exists another point of tangency for a $\Delta(q)$ line in the convex part of the production function. However, this steady state is unstable.

References

Aten, B., A Heston, and R. Summers, *Penn World Table Version 6.0*, Database, Center for International Comparisons at the University of Pennsylvania 2002.

Barro, R. and J.W. Lee, International Comparisons of Educational Attainment, *Journal of Monetary Economics*, 1993, *32* (3), 363–394.

Beaudry, P. and F. Collard, *Why has the Employment–Productivity Tradeoff among Industrialized Countries been so Strong?*, Working Paper 8754, NBER 2002.

———, ———, and D. Green, *Decomposing the Twin Peaks: A Study of the Changing World Distribution of output per worker*, Working Paper 9240, NBER 2002.

Bernanke, B.S. and R.S. Gürkaynak, Is Growth Exogenous? Taking Mankiw, Romer and Weil Seriously, in B.S. Bernanke and K. Rogoff, editors, *NBER Macroeconomics Annual*, Vol. 16, Cambridge MA: MIT Press, 2001, chapter 1.

Ciccone, A. and R.E. Hall, Productivity and the Density of Economic Activity, *American Economic Review*, 1996, *86* (1), 54–70.

Ethier, W.J., Internationally Decreasing Costs and International Trade, *Journal of International Economics*, 1979, *9* (1), 1–24.

———, National and International Returns to Scale in the Modern Theory of International Trade, *American Economic Review*, 1982, *72* (3), 389–406.

Fujita, M., P. Krugman, and A.J. Venables, *The Spatial Economy: Cities, Regions and International Trade*, Cambridge, MA: MIT Press, 1999.

Grossman, G.M. and E. Helpman, Trade, Knowledge Spillovers and Growth, *European Economic Review*, 1991, *35* (2–3), 517–526.

Helpman, E. and P. Krugman, *Market Structure and Foreign Trade*, Cambridge, MA: MIT Press, 1985.

Jones, C.I., On the Evolution of the World Income Distribution, *Journal of Economic Perspectives*, 1997, *11* (3), 19–36.

Klenow, P.J. and A. Rodrìguez-Clare, The Neoclassical Revival in Growth Economics: Has it Gone too Far?, in B.S. Bernanke and J. Rotemberg, editors, *NBER Macroeconomics Annual*, Vol. 12, Cambridge MA: MIT Press, 1997, chapter 2.

Krugman, P., Increasing Returns, Monopolistic Competition, and International Trade, *Journal of International Economics*, 1979, *9* (4), 469–479.

—— and A.J. Venables, Globalization and the Inequality of Nations, *Quarterly Journal of Economics*, 1995, *110* (4), 857–880.

Mankiw, N.G., D. Romer, and D.N. Weil, A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, 1992, *107* (2), 407–437.

Neary, J.P., Of Hype and Hyperbolas: Introducing the New Economic Geography, *Journal of Economic Literature*, 2001, *39* (2), 536–561.

Quah, D., Empirical Cross–Section Dynamics in Economic Growth, *European Economic Review*, 1993, *37* (2/3), 426–434.

—— , Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs, *Journal of Economic Growth*, 1997, *2* (1), 27–59.

Solow, R.M., A Contribution to the Theory of Economic Growth, *Quarterly Journal of Economics*, 1956, *70* (1), 65–94.

— Appendix —

A Data

Two datasets are used in this study. Most of the data are taken from the latest version of the Penn World Table 6.0 downloadable from:
`http://webhost.bridgew.edu/baten/`

Education data are taken from the Barro and Lee [1993] dataset, which is downloadable from:
`http://www.nuff.ox.ac.uk/Economics/Growth/datasets.htm`

A.1 Main data

Our measure of income, y, is the logarithm of real GDP chain per worker (RGDPW in PWT 6.0), where the definition of a worker is based on the economically active population.

Population is POP in PWT 6.0. Workers are computed as the population from 15 to 64 obtained from

$$\texttt{POPW}_t = \frac{\text{real GDP chain per capita}}{\text{real GDP chain per worker}} \times \text{population} = \frac{\texttt{RGDPL}}{\texttt{RGDPW}} \times \texttt{POP}$$

n then denotes the rate of growth of the 15–64 population.

The corresponding annualized average rate of growth for the variable Z within the sub–sample [t;t+n] is computed as

$$\Delta z = \frac{\log(Z_{t+n}) - \log(Z_t)}{n}$$

The share of consumption at constant prices corresponds to the variable KC in the PWT 6.0. In the IV procedure, the average share of consumption c/y over the sub–sample [t;t+n] is computed as

$$\frac{1}{n}\sum_{j=1}^{n} \log(\texttt{KC}_{t+j}/100)$$

The investment ratio at constant prices corresponds to the variable KI in the PWT 6.0 and is divided by 100. In the regressions, i/y then refers to the logarithm of this variable. It is also used to compute our accumulation variable that accounts for the overall accumulation effect, which is given by

$$\mathcal{K} \equiv \log\left(\frac{\texttt{KI}}{(1+\gamma)(1+n)+\delta-1}\right)$$

We follow Mankiw et al. [1992] and assume an annual depreciation rate of $\delta = 0.03$ and a rate of growth of technical progress, γ, of 2%.[28]

Further, the investment share is used to compute the capital stock per worker — needed to compute the TFP — using the permanent inventory scheme

$$(1+n)k_{t+1} = i_t + (1-\delta)k_t$$

[28] We performed robustness checks but did not find any significant effect of a change in either δ or γ.

where $k = K/L$. k_{1960} is obtained as

$$k_{1960} = \frac{\left(\frac{1}{10}\sum_{t=1960}^{1970} \texttt{KI}_t\right)\texttt{RGDPW}_{1960}}{(1+\overline{n}_{60-70})(1+\gamma)+\delta-1}$$

where

$$\overline{n}_{60-70} \equiv \frac{\log(\texttt{POPW}_{1970})-\log(\texttt{POPW}_{1960})}{10}$$

The measure of TFP then iss

$$TFP = \log(\texttt{RGDPW}) - \alpha\log(k)$$

where α is set to 0.35 and 0.5 to check robustness.

A.2 Education data

The education variables are borrowed from Barro and Lee [1993]. We consider essentially 3 measures of human capital. The first one is related to the overall enrollment rate in education. Assuming that, on average, most people sped 6 years in primary schooling, 6 years in secondary schooling and 4 years in higher schooling, we first define the index

$$\log\left(\frac{6\times\texttt{P}+6\times\texttt{S}+4\times\texttt{H}}{16}\right)$$

where P, S and H respectively denote the total gross enrollment ratio for, respectively, primary, secondary and higher schooling.

Our last education variable, Years, should be more understood as a stock since it is given by

$$\log(\texttt{HUMAN})$$

where HUMAN is the average schooling years in the total population over 25.

Note that these measures are only reported every 5 years in the database from 1960 to 1985. We actually used an average of years 1960 and 1965 for the first sub–sample and 1975 and 1980 for the second sub–sample. We restrict ourselves to 2 periods for the average for data availability purposes.

A.3 Composition of the sample

Our restricted sample of 75 countries consists of:

Argentina, Australia, Austria, Belgium, Bangladesh, Bolivia, Brazil, Barbados, Botswana, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Cyprus, Denmark, Dominican Republic, Ecuador, Egypt, Spain, Finland, Fiji, France, United Kingdom, Greece, Guatemala, Guyana, Hong Kong, Honduras, Indonesia, India, Ireland, Iran, Iceland, Israel, Italy, Jamaica, Jordan, Japan, Republic of Korea, Sri Lanka, Lesotho, Luxembourg, Morocco, Mexico, Mozambique, Malaysia, Namibia, Nicaragua, Netherlands, Norway, Nepal, New Zealand, Pakistan, Panama, Peru, Philippines, Papua New Guinea, Portugal, Paraguay, Romania, Singapore, El Salvador, Sweden, Syria, Thailand, Trinidad and Tobago, Tunisia, Turkey, Taiwan, Uruguay, USA, Venezuela, South Africa,

B Proof of Propositions

In this section we will abstract from any reference to the index of the economy, except when strictly necessary, in order to save on notation.

Proposition 1 In the absence of trade, the problem the final good producer has to solve is

$$\max_{K_t,Q_{jt}(\ell);j=1,2,\ell\in(0,1)} Y_t - q_tK_t - \int_0^1 (p^{\mathrm{Q}}_{1t}(\ell)Q_{1t}(\ell) + p^{\mathrm{Q}}_{2t}(\ell)Q_{2t}(\ell))\,\mathrm{d}\ell$$

which yields the following set of first order conditions

$$q_t = \alpha(1-\varphi)\frac{Y_t}{K_t} \tag{27}$$

$$p^{\mathrm{Q}}_{1t}(\ell) = \varphi Y_t Q_{1t}(\ell)^{\rho-1}M_{1t}^{-\rho} \tag{28}$$

$$p^{\mathrm{Q}}_{2t}(\ell) = (1-\varphi)(1-\alpha)Y_t\theta^{\mathrm{s}}_t(\ell)^{\rho}Q_{2t}(\ell)^{\rho-1}M_{2t}^{-\rho} \tag{29}$$

The intermediate firm then maximizes

$$\max_{\{L_t\geqslant 0,\sigma_t(\ell)\in[0,1]\}} p^{\mathrm{Q}}_{1t}(\ell)\sigma_t(\ell)\Gamma_tL_t(\ell) + p^{\mathrm{Q}}_{2t}(\ell)(1-\sigma_t(\ell))\Gamma_tL_t(\ell) - W_tL_t(\ell) \tag{30}$$

subject to (28) and (29). The set of first order conditions is then given by

$$W_t = \rho\varphi Y_t\frac{Q_{1t}(\ell)^{\rho}}{L_t(\ell)}M_{1t}^{-\rho} + \rho(1-\varphi)(1-\alpha)Y_t\theta^{\mathrm{s}}_t(\ell)^{\rho}\frac{Q_{2t}(\ell)^{\rho}}{L_t(\ell)}M_{2t}^{-\rho} \tag{31}$$

$$\rho\varphi Y_t\frac{Q_{1t}(\ell)^{\rho}}{\sigma_t(\ell)}M_{1t}^{-\rho} = \rho(1-\varphi)(1-\alpha)Y_t\theta^{\mathrm{s}}_t(\ell)^{\rho}\frac{Q_{2t}(\ell)^{\rho}}{1-\sigma_t(\ell)}M_{2t}^{-\rho} \tag{32}$$

In a symmetric equilibrium, we have

$$\sigma_t(\ell) = \sigma_t \qquad L_t(\ell) = L_t$$
$$Q_{1t}(\ell) = Q_{1t} \quad Q_{2t}(\ell) = Q_{2t}$$

Furthermore, rational expectations and symmetric equilibrium imply that $\theta^{\mathrm{s}}_t(\ell) = \theta_t(\ell) = \theta_t$, such that $M_{1t} = Q_{1t} = \sigma_t\Gamma_tL_t$ and $M_{2t} = \theta_tQ_{2t} = \theta_t(1-\sigma_t)\Gamma_tL_t$. Hence, the labor allocation choice becomes

$$\rho\varphi Y_t\sigma_t^{\rho-1}(\Gamma_tL_t)^{\rho}(\sigma_t\Gamma_tL_t)^{-\rho} = \rho(1-\varphi)(1-\alpha)Y_t\theta_t^{\rho}(1-\sigma_t)^{\rho-1}(\Gamma_tL_t)^{\rho}(\theta_t(1-\sigma_t)\Gamma_tL_t)^{-\rho}$$

such that

$$\sigma_t = \sigma^{\star} \equiv \frac{\varphi}{\varphi + (1-\alpha)(1-\varphi)}$$

Note that θ_t then reduces to $\theta_t = \overline{\theta}(1-\sigma^{\star})^{\varepsilon} = 1$ since $\overline{\theta} = (1-\sigma^{\star})^{-\varepsilon}$. Plugging the optimal allocation of labor in the definition of total output, we get

$$Y_t = A^{1-\varphi}\sigma^{\star\varphi}(1-\sigma^{\star})^{(1-\varphi)(1-\alpha)}K_t^{\alpha(1-\varphi)}(\Gamma_tL_t)^{\varphi+(1-\varphi)(1-\alpha)}$$

Henceforth, output per efficient worker can be written as

$$y_t = A^{1-\varphi}\sigma^{\star\varphi}(1-\sigma^{\star})^{(1-\varphi)(1-\alpha)}k_t^{\alpha(1-\varphi)} = Bk_t^{\alpha(1-\varphi)}$$

where $B \equiv A^{1-\varphi}\sigma^{\star\varphi}(1-\sigma^{\star})^{(1-\varphi)(1-\alpha)}$

Hence the dynamics of the economy — in intensive form — may be summarized by

$$(1+\gamma)(1+n)k_{t+1} = sBk_t^{\alpha(1-\varphi)} + (1-\delta)k_t$$

which admits

$$k^{\star} = \left(\frac{sB}{(1+\gamma)(1+n)+\delta-1}\right)^{\frac{1}{1-\alpha(1-\varphi)}}$$

as steady state.

Q.E.D □

Proposition 2 Let us recall that the steady state value of the capital stock in country i is given by

$$k_i^\star = \left(\frac{s_i B}{(1+\gamma)(1+n)+\delta-1}\right)^{\frac{1}{1-\alpha(1-\varphi)}}$$

such that the steady state level of output can be written as

$$y_i^\star = \left(\frac{s_i}{(1+\gamma)(1+n)+\delta-1}\right)^{\frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}} B^{\frac{1}{\alpha(1-\varphi)}}$$

or in logarithm

$$\log(y_i^\star) = \frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}\log\left(\frac{s_i}{(1+\gamma)(1+n)+\delta-1}\right) + \frac{1}{\alpha(1-\varphi)}\log(B)$$

or

$$\log(y_i^\star) = \frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}\nu_i + \frac{1}{\alpha(1-\varphi)}\log(B)$$

Therefore, the deviation from the median level is given by

$$\widehat{y}_i = \frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}(\nu_i - \nu^m) = g(\nu_i - \nu^m)$$

Making use of the change of variable formula, and denoting by $\mu^\nu(\cdot)$ the distribution of ν, we have

$$\mu(\widehat{y}) = \frac{\mu^\nu(g^{-1}(y))}{g'(g^{-1}(y))}$$

Since $g(\cdot)$ takes the simple linear form $g(\nu) = \nu/\psi$ with $\psi = \alpha(1-\varphi)/(1-\alpha(1-\varphi))$, it follows that

$$\mu(\widehat{y}) = \psi\mu^\nu(\psi y + \nu^m)$$

which can be restated as

$$\mu(\widehat{y}) = \psi\mu^{\widehat{\nu}}(\psi y)$$

Q.E.D □

Lemma 1 In the small open economy, each firm takes the price of goods as given, such that it solves

$$\max_{K_t, Q_{jt}(\ell); j=1,2, \ell\in(0,1)} Z_{1t} + p_t Z_{2t} - q_t K_t - \int_0^1 \left(p^{\mathrm{Q}}_{1t}(\ell)Q_{1t}(\ell) + p^{\mathrm{Q}}_{2t}(\ell)Q_{2t}(\ell)\right)\mathrm{d}\ell$$

which yields the following set of first order conditions

$$q_t = p_t\alpha\frac{Z_{2t}}{K_t} \quad (33)$$

$$p^{\mathrm{Q}}_{1t}(\ell) = \Gamma_t Q_{1t}(\ell)^{\rho-1}M_{1t}^{1-\rho} \quad (34)$$

$$p^{\mathrm{Q}}_{2t}(\ell) = (1-\alpha)p_t\theta^{\mathrm{s}}_t(\ell)^\rho Q_{2t}(\ell)^{\rho-1}AK_t^\alpha\Gamma_t^{1-\alpha}M_{2t}^{1-\alpha-\rho} \quad (35)$$

The intermediate firm then maximizes

$$\max_{\{L_t\geqslant 0, \sigma_t(\ell)\in[0,1]\}} p^{\mathrm{Q}}_{1t}(\ell)\sigma_t(\ell)\Gamma_t L_t(\ell) + p^{\mathrm{Q}}_{2t}(\ell)(1-\sigma_t(\ell))\Gamma_t L_t(\ell) - W_t L_t(\ell) \quad (36)$$

subject to (34) and (35). The set of first order conditions is then given by

$$W_t = \rho\frac{Q_{1t}(\ell)^\rho}{L_t(\ell)}M_{1t}^{1-\rho} + \rho(1-\alpha)p_t\theta^{\mathrm{s}}_t(\ell)^\rho\frac{Q_{2t}(\ell)^\rho}{L_t(\ell)}AK_t^\alpha\Gamma_t^{1-\alpha}M_{2t}^{1-\alpha-\rho} \quad (37)$$

$$\rho\frac{Q_{1t}(\ell)^\rho}{\sigma_t(\ell)}M_{1t}^{-\rho} \leqslant \rho(1-\alpha)p_t\theta^{\mathrm{s}}_t(\ell)^\rho\frac{Q_{2t}(\ell)^\rho}{1-\sigma_t(\ell)}AK_t^\alpha\Gamma_t^{1-\alpha}M_{2t}^{1-\alpha-\rho} \quad (38)$$

In a symmetric equilibrium, we have

$$\begin{array}{cc}\sigma_t(\ell) = \sigma_t & L_t(\ell) = L_t \\ Q_{1t}(\ell) = Q_{1t} & Q_{2t}(\ell) = Q_{2t}\end{array}$$

Furthermore, rational expectations and symmetric equilibrium imply that $\theta_t^s(\ell) = \theta_t(\ell) = \theta_t$, such that $M_{1t} = Q_{1t} = \sigma_t\Gamma_t L_t$ and $M_{2t} = \theta_t Q_{2t} = \theta_t(1-\sigma_t)\Gamma_t L_t$. Hence, the labor allocation choice becomes

$$\rho\sigma_t^{\rho-1}(\Gamma_t L_t)^{\rho}(\sigma_t\Gamma_t L_t)^{1-\rho} \leqslant \rho(1-\alpha)p_t\theta_t^{\rho}(1-\sigma_t)^{\rho-1}(\Gamma_t L_t)^{\rho}AK_t^{\alpha}(\theta_t(1-\sigma_t)\Gamma_t L_t)^{1-\alpha-\rho}$$

or

$$\rho\Gamma_t L_t \leqslant \rho(1-\alpha)p_t AK_t^{\alpha}\Gamma_t^{1-\alpha}\theta_t^{1-\alpha}(1-\sigma_t)^{-\alpha}L_t^{1-\alpha}$$

which rewrites

$$1 \leqslant (1-\alpha)p_t Ak_t^{\alpha}\theta_t^{1-\alpha}(1-\sigma_t)^{-\alpha}$$

Taking the definition of θ_t into account ($\theta_t = \overline{\theta}(1-\sigma_t)^{\varepsilon}$), it should be clear that the economy specializes in the production of good 2, $\sigma_t = 0$, as soon as

$$k_t \geqslant \widetilde{k}(p_t) \equiv \left((1-\alpha)A\overline{\theta}^{1-\alpha}p_t\right)^{-\frac{1}{\alpha}}$$

Noting that $\overline{\theta} = (1-\sigma^{\star})^{-\varepsilon}$, we can calculate the the effect of ε on $\widetilde{k}(p_t)$ as

$$\frac{\partial\widetilde{k}(p_t)}{\partial\varepsilon} = \frac{1-\alpha}{\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p_t\right)^{-\frac{1}{\alpha}}\log(1-\sigma^{\star}) < 0$$

To prove the second component of the proposition, note that in the limit, we have

$$\overline{\theta} = \left(\frac{\varphi+(1-\alpha)(1-\varphi)}{(1-\alpha)(1-\varphi)}\right)^{\frac{\alpha}{1-\alpha}} = (1-\sigma^{\star})^{-\frac{\alpha}{1-\alpha}}$$

Therefore

$$\widetilde{k}(p) = [(1-\alpha)pA]^{-1/\alpha}(1-\sigma^{\star})$$

From the program of the final good producer in the median economy, we have

$$p = \frac{1-\varphi}{\varphi}\frac{z_1}{z_2} = \frac{1-\varphi}{\varphi}\frac{\sigma^{\star}}{A(1-\sigma^{\star})^{1-\alpha}{k^{\star}}^{\alpha}}$$

Plugging this result in the former equation, we get

$$\widetilde{k}(p) = \left[(1-\alpha)\frac{1-\varphi}{\varphi}\frac{\sigma^{\star}}{A(1-\sigma^{\star})^{1-\alpha}{k^{\star}}^{\alpha}}A\right]^{-1/\alpha}(1-\sigma^{\star})$$

by definition of $\sigma^{\star}$, this rewrites

$$\widetilde{k}(p) = \left[\frac{1-\sigma^{\star}}{\sigma^{\star}}\frac{\sigma^{\star}}{A(1-\sigma^{\star})^{1-\alpha}{k^{\star}}^{\alpha}}A\right]^{-1/\alpha}(1-\sigma^{\star})$$

which simplifies to

$$\widetilde{k}(p) = k^{\star}$$

Q.E.D □

Proposition 3 We have to study two cases:

$k \leqslant \widetilde{k}(p)$: In such a case, $\sigma > 0$, so that the first order condition holds with equality and

$$(1-\alpha)pAk^{\alpha}\overline{\theta}^{1-\alpha}(1-\sigma)^{\varepsilon(1-\alpha)-\alpha} = 1$$

from which we get

$$\sigma = 1 - \left[(1-\alpha)pA\overline{\theta}^{1-\alpha}\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}$$

Since $z_1 = \sigma$ and $z_2 = Ak^{\alpha}(\overline{\theta}(1-\sigma)^{(1+\varepsilon)})^{1-\alpha}$ in the long–run, and given that $p_y y = z_1 + pz_2$, we have

$$\begin{aligned} p_y y &= 1 - \left[(1-\alpha)pA\overline{\theta}^{1-\alpha}\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}} \\ &\quad + pAk^{\alpha}\left(\overline{\theta}\left(\left[(1-\alpha)pA\overline{\theta}^{1-\alpha}\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right)^{(1+\varepsilon)}\right)^{1-\alpha} \end{aligned}$$

This reduces to

$$p_y y = 1 + \frac{\alpha}{1-\alpha}\left[A(1-\alpha)\overline{\theta}^{1-\alpha}p\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}$$

Finally, note that from the final good producer, we get $p_y = p^{1-\varphi}/\Phi$ where $\Phi \equiv (\Phi)$. Plugging this result in this equation, we get

$$y = \Phi p^{\varphi-1}\left[1 + \frac{\alpha}{1-\alpha}\left[A(1-\alpha)\overline{\theta}^{1-\alpha}p\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right]$$

Note that

$$\frac{\partial y}{\partial k} = \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\Phi p^{\varphi-1}\left[\frac{\alpha}{1-\alpha}\left[A(1-\alpha)\overline{\theta}^{1-\alpha}p\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}-1}\right]$$

such that y is increasing in k iff $\varepsilon < \frac{\alpha}{1-\alpha}$. Furthermore

$$\frac{\partial^2 y}{\partial k^2} = \left(\frac{\alpha}{\alpha-\varepsilon(1-\alpha)} - 1\right)\frac{1}{k}\frac{\partial y}{\partial k} > 0$$

as long as $\varepsilon < \frac{\alpha}{1-\alpha}$, therefore creating local convexities.

$k > \widetilde{k}(p)$: In such a case, $\sigma_t = 0$, so that the long–run production function reduces to

$$p_y y = pz_2 = pAk^{\alpha}\overline{\theta}^{1-\alpha}$$

Plugging the definition of p_y in the last equation, we get

$$y = \Phi p^{\varphi}A\overline{\theta}^{1-\alpha}k^{\alpha}$$

Let us now prove the second part of the proposition. First note that

$$\left.\frac{\partial y^{\mathrm{A}}}{\partial k}\right|_{k=k^{\star}} = \alpha(1-\varphi)Bk^{\star\alpha(1-\varphi)-1}$$

Likewise, let us recall that

$$\frac{\partial y^{\mathrm{FT}}}{\partial k} = \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\Phi p^{\varphi-1}\left[\frac{\alpha}{1-\alpha}\left[A(1-\alpha)\overline{\theta}^{1-\alpha}p\right]^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}-1}\right]$$

We know that

$$p = \frac{1-\varphi}{\varphi}\frac{z_1}{z_2}$$

such that

$$p\Big|_{k=k^\star} = \frac{1-\varphi}{\varphi}\frac{\sigma^\star}{A(1-\sigma^\star)^{1-\alpha}k^{\star\alpha}}$$

Plugging this in the derivative, and remembering that

$$\overline{\theta} = (1-\sigma^\star)^{-\varepsilon}, \Phi = \varphi^\varphi(1-\varphi)^{1-\varphi} \text{ and } \frac{\sigma^\star}{1-\sigma^\star} = \frac{\varphi}{(1-\alpha)(1-\varphi)}$$

we get

$$\left.\frac{\partial y^{\text{FT}}}{\partial k}\right|_{k=k^\star} = \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\alpha(1-\varphi)\sigma^{\star\varphi}(1-\sigma^\star)^{(1-\alpha)(1-\varphi)}A^{1-\varphi}k^{\star\alpha(1-\varphi)-1}$$

Let us now recall that by definition of B, we have $B \equiv \sigma^{\star\varphi}(1-\sigma^\star)^{(1-\alpha)(1-\varphi)}A^{1-\varphi}$, such that the last equation rewrites:

$$\left.\frac{\partial y^{\text{FT}}}{\partial k}\right|_{k=k^\star} = \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}Bk^{\star\alpha(1-\varphi)-1} = \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\left.\frac{\partial y^{\text{A}}}{\partial k}\right|_{k=k^\star}$$

Since $\varepsilon \geqslant 0$, it follows that

$$\left.\frac{\partial y^{\text{FT}}}{\partial k}\right|_{k=k^\star} \geqslant \left.\frac{\partial y^{\text{A}}}{\partial k}\right|_{k=k^\star}$$

Q.E.D □

Proposition 4 Let us recall that the steady of any of the small open economies we consider can be characterized by

$$\tau k = sy$$

where $\tau = ((1+\gamma)(1+n)+\delta-1)$ and where y is defined by proposition 3. Therefore, we can define the steady state function $\vartheta(k) = y/k$, which at each τ/s associates one or several steady states. This function takes the form

$$\vartheta(k) = \begin{cases} \frac{\Phi}{k}p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right) & \text{for } k \leqslant \widetilde{k}(p) \\ \frac{\Phi}{k}p^{\varphi}A\overline{\theta}^{1-\alpha}k^{\alpha-1} & \text{for } k > \widetilde{k}(p) \end{cases}$$

Uniqueness of the steady state then corresponds to the fact that this function is bijective over the whole interval. A necessary condition for bijection is that the function is monotonic. As can be seen from the definition of the $\vartheta(k)$, it is monotonically decreasing for $k > \widetilde{k}(p)$ since $\alpha \in (0,1)$. Therefore, monotonicity requires the function to be decreasing for values of the capital labor ratio $k \in (0, \widetilde{k}(p)]$. Over the latter interval,

$$\vartheta'(k) = \frac{\Phi p^{\varphi-1}}{k^2}\left[\frac{\varepsilon(1-\alpha)}{\alpha-\varepsilon(1-\alpha)}\frac{\alpha}{1-\alpha}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}-1\right]$$

such that

$$\vartheta'(k) \gtreqless 0 \Longleftrightarrow k \gtreqless \widehat{k} \equiv \left(\frac{\alpha-\varepsilon(1-\alpha)}{\alpha\varepsilon}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-\frac{1}{\alpha}} = \varsigma\widetilde{k}(p)$$

where $\varsigma \equiv \left(\frac{\alpha-\varepsilon(1-\alpha)}{\alpha\varepsilon}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}$. As long as $0 < \varepsilon < \alpha$, ς is greater than one, such that $\vartheta'(k) < 0$ over $(0, \widetilde{k}(p))$. Therefore the function is monotonic and the economy admits a unique steady state.

Q.E.D □

Proposition 5 First of all, let us recall that the steady state of the model economy is defined by

$$((1+n)(1+\gamma)-(1-\delta))k = sy \text{ or equivalently } \nu = \frac{k}{y}$$

where y is given in proposition 3. From the definition of y, we can straightforwardly express ν as a function of y as

$$\nu = \begin{cases} g_1(y) \equiv \frac{\Psi(p)}{y}\left(y - \Phi p^{\varphi-1}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}} & \text{if } y \leqslant \widetilde{y}(p) \\ g_2(y) \equiv \left[\Phi p^{\varphi} A \overline{\theta}^{1-\alpha}\right]^{-\frac{1}{\alpha}} y^{\frac{1-\alpha}{\alpha}} & \text{if } y < \widetilde{y}(p) \end{cases}$$

where $\Psi(p) \equiv \left(\frac{1-\alpha}{\alpha}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-\frac{1}{\alpha}}$, and $\widetilde{y}(p)$ is obtained by plugging the definition of $\widetilde{k}(p)$ into the production function, implying

$$\widetilde{y}(p) = \frac{\Phi p^{\varphi-1}}{1-\alpha}$$

The standard change of variable formula states that, for any random variable continuous X with density functionμ_X and for any one–to–one function $\Phi : S \subset \mathbb{R} \longrightarrow \mathbb{R}$, such that $\Phi'(.)$ exists and does not vanish on S, the random variable $Z = \Phi(X)$ is continuous with density function given by

$$\mu_Z(Z) = \mu_X(\Phi^{-1}(Z))\,|\Phi^{-1\prime}(Z)|$$

Since, as long as $0 < \varepsilon < \alpha$, the production function is one–to–one (see proposition 4), we can apply the preceding result to our economy. Therefore, let us assume that the saving rate is distributed according to the density function $\mu_s(s)$, the distribution of income can be obtained applying the previous formula. Noting that

$$\left|\frac{\partial \nu}{\partial y}\right| = \begin{cases} g_1(y) \times \left[\frac{\alpha\Phi p^{\varphi-1} - \varepsilon(1-\alpha)y}{\alpha y(y - \Phi p^{\varphi-1})}\right] & \text{if } y \leqslant \widetilde{y}(p) \\ \frac{1-\alpha}{\alpha} \times \frac{g_2(y)}{y} & \text{if } y > \widetilde{y}(p) \end{cases}$$

The result then follows.

Q.E.D □

Proposition 6 In order to prove the first part of the proposition, we refer explicitly to section 1.2. For the density to reduce at the median of the distribution — in the absence of a change in the distribution of ν — it must be the case that the marginal effect of ν on y evaluated at the median ν must increase:

$$\left.\frac{\partial y^{\text{FT}}}{\partial \nu}\right|_{\nu=\nu^m} \geqslant \left.\frac{\partial y^{\text{A}}}{\partial \nu}\right|_{\nu=\nu^m} \text{ if } 0 < \epsilon < \alpha$$

where FT and A denote respectively free trade and autarky. Let us first compute $\left.\frac{\partial y^{\text{A}}}{\partial \nu}\right|_{\nu=\nu^m}$. Since

$$\left.y^{\text{A}}\right|_{\nu=\nu^m} = Bk^{\star\alpha(1-\varphi)}$$

and given that $\nu = k^\star/y^\star$, we have

$$\left.y^{\text{A}}\right|_{\nu=\nu^m} = B^{\frac{1}{1-\alpha(1-\varphi)}}\nu^{m\,\frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}}$$

Therefore

$$\left.\frac{\partial y^{\text{A}}}{\partial \nu^m}\right|_{\nu=\nu^m} = \frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}B^{\frac{1}{1-\alpha(1-\varphi)}}\nu^{m\,\frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}-1} = \left.\frac{\alpha(1-\varphi)}{1-\alpha(1-\varphi)}\frac{y^{\text{A}}}{\nu}\right|_{\nu=\nu^m}$$

We now have to compute the free trade counterpart of this quantity, $\left.\frac{\partial y^{\text{FT}}}{\partial \nu}\right|_{\nu=\nu^m}$. Since, we consider the case where we open the economy in a neighborhood of the steady state, it has to be the case that $k^\star \leqslant \widetilde{k}(p)$ so that

$$y^{\text{FT}}\big|_{\nu=\nu^m} = \Phi p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} k^{\star\,\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right)$$

using the definition of ν, we obtain

$$y^\star = \Phi p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} (\nu^m y^\star)^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right)$$

Taking the total derivative of this expression, we obtain

$$\begin{aligned}\mathrm{d}y &= \frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\Phi p^{\varphi-1}\left(\frac{\alpha}{1-\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} (\nu^m y)^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right)\frac{\mathrm{d}\nu}{\nu}\\ &+\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\Phi p^{\varphi-1}\left(\frac{\alpha}{1-\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}} (\nu^m y)^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right)\frac{\mathrm{d}y}{y}\end{aligned}$$

where it should be clear that y stands for $y^{\text{FT}}\Big|_{\nu=\nu^m}$ and ν stands for ν^m. Remembering that $\Phi = \varphi^\varphi(1-\varphi)^{1-\varphi}$, $\overline{\theta} = (1-\sigma^\star)^{-\varepsilon}$, $\sigma^\star = \frac{\varphi}{\varphi+(1-\alpha)(1-\varphi)}$ and $p = \frac{1-\varphi}{\varphi}\frac{\sigma^\star}{A(1-\sigma)^{1-\alpha}k^{\star\alpha}}$, it is straightforward to see that the latter equation reduces to

$$\mathrm{d}y = \alpha(1-\varphi)\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\frac{y^{\text{A}}}{\nu}\mathrm{d}\nu + \alpha(1-\varphi)\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\mathrm{d}y$$

Hence,

$$\left.\frac{\partial y^{\text{FT}}}{\partial \nu}\right|_{\nu=\nu^m} = \frac{\alpha^2(1-\varphi)}{\alpha(1-\alpha(1-\varphi))-\varepsilon(1-\alpha)}\times\left.\frac{y^{\text{A}}}{\nu}\right|_{\nu=\nu^m} = \frac{\alpha(1-\alpha(1-\varphi))}{\alpha(1-\alpha(1-\varphi))-\varepsilon(1-\alpha)}\times\left.\frac{\partial y^a}{\partial \nu}\right|_{\nu=\nu^m}$$

Since $0\leqslant\alpha\leqslant 1$, $0\leqslant\varphi\leqslant 1$ and $\varepsilon>0$, we have $\frac{\alpha(1-\alpha(1-\varphi))}{\alpha(1-\alpha(1-\varphi))-\varepsilon(1-\alpha)} > 1$ such that

$$\left.\frac{\partial y^{\text{FT}}}{\partial \nu}\right|_{\nu=\nu^m} \geqslant \left.\frac{\partial y^{\text{A}}}{\partial \nu}\right|_{\nu=\nu^m}$$

In order to establish the second part of the proposition, and therefore the multimodal shape of the distribution as ε tends to α from below, let us evaluate the density function in $\widetilde{k}(p)$. First of all note that $k=\widetilde{k}(p)$ for a particular constant value of $\nu=\overline{\nu}$, implying that the first component of the distribution — that stemming from μ_ν — is constant, $\overline{\mu}_s = \mu_s(\overline{\nu})$. However, the second component stemming from the derivative term depends fundamentally on how $\widetilde{k}(p)$ is reached. When k approaches $\widetilde{k}(p)$ from above, the technology is its concave part, such that

$$\lim_{k\downarrow\widetilde{k}(p)} |f'(y(k))| = \frac{1-\alpha}{\alpha}\left[pA(1-\alpha)\overline{\theta}^{1-\alpha}\right]^{-\frac{1}{\alpha}}\left(\frac{1-\alpha}{\Phi p^{\varphi-1}}\right)^{\frac{1-2\alpha}{\alpha}}$$

Conversely, when k approaches $\widetilde{k}(p)$ from below, the technology is locally convex and

$$\lim_{k\uparrow\widetilde{k}(p)} |f'(y(k))| = \frac{\alpha-\varepsilon}{\alpha}\frac{1-\alpha}{\alpha}\left[pA(1-\alpha)\overline{\theta}^{1-\alpha}\right]^{-\frac{1}{\alpha}}\left(\frac{1-\alpha}{\Phi p^{\varphi-1}}\right)^{2}\left(\Phi p^{\varphi-1}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}$$

The latter limit indicates that as ε tends to α, the left hand side of the distribution tends to zero as output per worker approaches $y(\widetilde{k}(p))$ while it remains positive on the right hand side. Therefore, provided the pdf of the marginal propensity to accumulate, ν, is not degenerate, no matter its particular form, the income distribution will exhibit at least 2 modes.

Q.E.D □

Proposition 7 Let us first characterize $\underline{\nu}$ and $\overline{\nu}$.

1. Characterizing $\underline{\nu}$: From Figure 4, it shall be clear that there are 2 candidates to be the lowest possible value for the saving rate. The first one is such that $k = \widetilde{k}(p)$ and the second one satisfies the two restrictions
$$\begin{cases} k = \underline{\nu} y \\ 1 = \underline{\nu}\frac{\partial y}{\partial k} \end{cases}$$
where $\nu = s/((1+\gamma)(1+n)-(1-\delta))$. Each restriction is evaluated on the concave part of the function — that is when $\sigma = 0$. We shall first show that this restriction is not relevant. Indeed this amounts to solve the system
$$\begin{cases} k = \underline{\nu}\Phi p^{\varphi} A\overline{\theta}^{1-\alpha} k^{\alpha} \\ 1 = \underline{\nu}\alpha\Phi p^{\varphi} A\overline{\theta}^{1-\alpha} k^{\alpha-1} \end{cases}$$
for k and $\underline{\nu}$. But, multiplying the second equation by k and subtracting the first one from the result, we get
$$\underline{\nu}(\alpha-1)y(\widetilde{k}) = 0$$
This triggers $\underline{\nu} = 0$ and leads to $k = 0$, which cannot be the case as the economy is supposed to lie on the concave part of the function, *i.e.* above $\widetilde{k}(p)$. Henceforth, this cannot be a solution.
Therefore, $\underline{\nu}$ is such that $k(\underline{\nu}) = \widetilde{k}(p) = \left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1/\alpha}$:
$$\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1/\alpha} = \underline{\nu}\Phi p^{\varphi} A\overline{\theta}^{1-\alpha}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1}$$
implying
$$\underline{\nu} = \frac{1-\alpha}{\Phi p^{\varphi-1}}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1/\alpha}$$

2. Characterizing $\overline{\nu}$: Figure 4 illustrates that if multiplicity occurs, it has to be the case that at least one steady state should lie on the convex part of the curve. The highest possible value for the saving rate that satisfies this restriction is such that the ray intersecting the production function is also a tangent. This implies the 2 restrictions
$$\begin{cases} k = \overline{\nu} y \\ 1 = \overline{\nu}\frac{\partial y}{\partial k} \end{cases}$$
These 2 conditions, when evaluated on the convex part of the production function, lead to
$$k = \overline{\nu}\Phi p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right) \quad (39)$$
$$1 = \overline{\nu}\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\Phi p^{\varphi-1}\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}-1} \quad (40)$$
Solving (40) for k, and denoting $\beta = \varepsilon(1-\alpha)/\alpha$, we get
$$\widehat{k} \equiv k(\overline{\nu}) = \overline{\nu}^{\frac{\beta-1}{\beta}}\left(\frac{(1-\alpha)(1-\beta)\tau}{\alpha\Phi}\right)^{\frac{1-\beta}{\beta}}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1/\alpha\beta}p^{\frac{(1-\varphi)(1-\beta)}{\beta}}$$
Plugging this expression in (39), and solving for $\overline{\nu}$, we get
$$\overline{\nu} = \left(\frac{1-\alpha}{\alpha}\right)^{1-\beta}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-1/\alpha}\frac{\beta^{\beta}(1-\beta)^{1-\beta}}{\Phi p^{\varphi-1}}$$

Note that as ε tends to α, β tends toward $1-\alpha$, implying that $\overline{\nu}$ tends to $\underline{\nu}$. In other word,

$$\lim_{\varepsilon\to\alpha}\overline{\nu}=\underline{\nu}$$

We now show that the economy admits at least 3 steady states. Let us recall that the steady state of any of the small open economies we consider can be characterized by

$$k=\nu y$$

where y is defined by proposition 3. Therefore, we can define the steady state function $\vartheta(k)=y/k$, which at each $1/\nu$ associates one or several steady states. This function takes the form

$$\vartheta(k)=\begin{cases}\frac{\Phi}{k}p^{\varphi-1}\left(1+\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}\right) & \text{for } k\leqslant\widetilde{k}(p)\\ \Phi p^{\varphi}A\overline{\theta}^{1-\alpha}k^{\alpha-1} & \text{for } k>\widetilde{k}(p)\end{cases}$$

Uniqueness of the steady state then corresponds to the fact that this function is one–to–one over the whole interval, for which a necessary condition is monotonicity. As can be seen from the definition of the $\vartheta(k)$, it is monotonically decreasing for $k>\widetilde{k}(p)$ since $\alpha\in(0,1)$. Therefore, monotonicity requires the function to be decreasing for values of the capital labor ratio $k\in(0,\widetilde{k}(p)]$. Over the latter interval,

$$\vartheta'(k)=\frac{\Phi p^{\varphi-1}}{k^2}\left[\frac{\varepsilon(1-\alpha)}{\alpha-\varepsilon(1-\alpha)}\frac{\alpha}{1-\alpha}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}}-1\right]$$

such that

$$\vartheta'(k)\gtreqless 0\Longleftrightarrow k\gtreqless\widehat{k}\equiv\left(\frac{\alpha-\varepsilon(1-\alpha)}{\alpha\varepsilon}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}\left(A(1-\alpha)\overline{\theta}^{1-\alpha}p\right)^{-\frac{1}{\alpha}}=\varsigma\widetilde{k}(p)$$

where $\varsigma\equiv\left(\frac{\alpha-\varepsilon(1-\alpha)}{\alpha\varepsilon}\right)^{\frac{\alpha-\varepsilon(1-\alpha)}{\alpha}}$. Further, ς is lower than one as long as $\varepsilon>\alpha$. Therefore, there exists $\widehat{k}\in(0,\widetilde{k}(p))$ such that the sign of the slope of $\vartheta(k)$ switches from negative to positive. Finally, note that — by definition of $\underline{n}$ and $\overline{n}$

$$\vartheta(\widehat{k})=\frac{1}{\overline{\nu}}\text{ and }\vartheta(\widetilde{k}(p))=\frac{1}{\underline{\nu}}$$

Therefore, to recap, $\vartheta(k)$ is a decreasing function for $k\in(0,\widehat{k}\equiv k(\overline{\nu}))$, (equivalently $\nu\in(\overline{\nu},1)$) and an increasing function for $k\in(\widehat{k},\widetilde{k}(p))$, (equivalently $\nu\in(\underline{\nu},\overline{\nu})$). Then for any value of $k>\widetilde{k}(p)$ (equivalently $\nu\in(0,\underline{\nu})$, $\vartheta(.)$ is a decreasing function, that is depicted in Figure 11. Hence, for any $\nu\in(\underline{\nu},\overline{\nu})$, there exist 3 values of k such that $\vartheta(k)=1/\nu$.
Q.E.D □

Proposition 8 As established in proposition 7, when trade is free $\alpha<\varepsilon<\frac{\alpha}{1-\alpha}$, the open economy admits 3 steady states.

Let us focus first on the upper steady state that arises when the economy fully specializes in the production of good 2. In such a situation, $\sigma=0$ and the production function reduces to

$$y_t=\Phi {p_t}^{\varphi}Ak_t^{\alpha}\overline{\theta}^{1-\alpha}$$

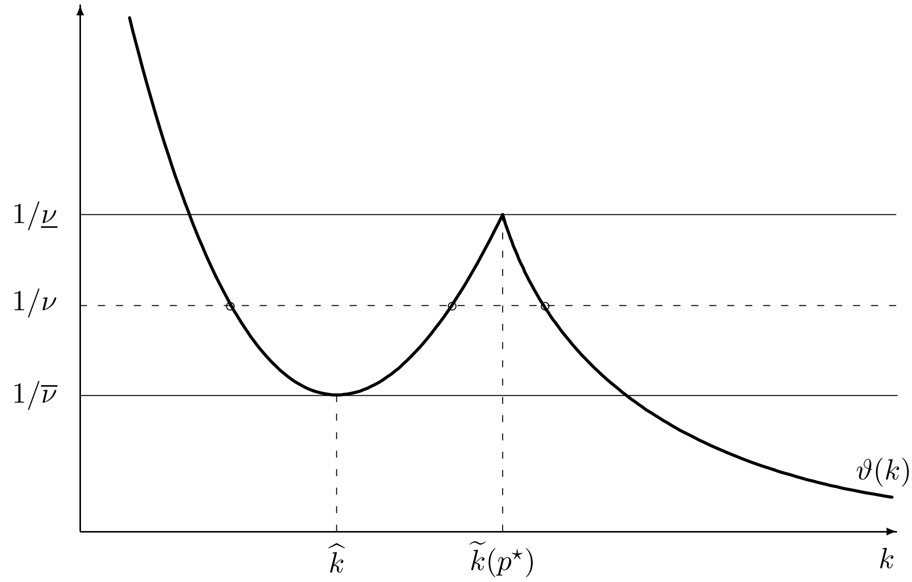
such that the dynamics of capital writes

$$(1+n)(1+\gamma)k_{t+1}=s\Phi {p_t}^{\varphi}Ak_t^{\alpha}\overline{\theta}^{1-\alpha}+(1-\delta)k_t$$

In the neighborhood of the steady state, the log–linearized version of the dynamics writes

$$x_{t+1}=\omega x_t$$

Figure 11: Multiplicity



where $x_t = \log(k_t/k)$ and $\omega = \left(\alpha + \frac{(1-\alpha)(1-\delta)}{(1+\gamma)(1+n)}\right) < 1$, for $n, \gamma, \delta > 0$. Since $\omega < 1$ this particular steady state is locally stable.

Let us now focus on local dynamic properties of the economy, when there is no specialization. In this case, the dynamics are given by the system:

$$k_{t+1} = \frac{s\Phi}{(1+\gamma)(1+n)} p_t^{\varphi-1}\left[1 + \frac{\alpha}{1-\alpha}(A(1-\alpha)p_t)^{\frac{1}{\alpha}}\theta_t^{\frac{1-\alpha}{\alpha}}k_t\right] + \frac{1-\delta}{(1+\gamma)(1+n)}k_t \quad (41)$$

$$\theta_{t+1} = \overline{\theta}(A(1-\alpha)p_t)^{\frac{\varepsilon}{\alpha}}\theta_t^{\frac{\alpha(1-\alpha)}{\alpha}}k_t^{\varepsilon} \quad (42)$$

The log–linear version of the system around a steady state yields the following representation

$$\begin{pmatrix} k_{t+1} \\ \theta_{t+1} \end{pmatrix} = \begin{pmatrix} \zeta(k) + \frac{1-\delta}{(1+n)(1+\gamma)} & \zeta(k)\frac{1-\alpha}{\alpha} \\ \varepsilon & \varepsilon\frac{1-\alpha}{\alpha} \end{pmatrix}\begin{pmatrix} k_t \\ \theta_t \end{pmatrix}$$

where

$$\zeta(k) = \frac{s\Phi}{(1+\gamma)(1+n)}p^{\varphi-1}\frac{\alpha}{1-\alpha}\left((1-\alpha)A\overline{\theta}^{1-\alpha}p\right)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}k^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}-1}$$

Note that $\zeta'(k) > 0$. The local dynamic properties of this economy may then be revealed, studying the properties of the characteristic polynomial, $P(\lambda)$, of the approximated system

$$P(\lambda) = \lambda^2 - Tr\lambda + Det$$

where Tr and Det denote respectively the trace and the determinant of the above defined matrix

$$Tr = \zeta(k) + \frac{1-\delta}{(1+n)(1+\gamma)} + \varepsilon\frac{1-\alpha}{\alpha}$$

$$Det = \varepsilon\frac{1-\alpha}{\alpha}\frac{1-\delta}{(1+n)(1+\gamma)}$$

First of all note that both the trace and the determinant of the matrix are positive implying that both eigenvalues characterizing the local dynamic properties of the economy are positive since both their sum and their product ($P(0) = Det$) are positive.

Denoting $a = \frac{1-\delta}{(1+n)(1+\gamma)}$ and $b = \varepsilon\frac{1-\alpha}{\alpha}$, the discriminant of the polynomial is given by

$$\Delta = \zeta(k)^2 + 2(a+b)\zeta(k) + a^2 + b^2 + 2ab - 4ab = \zeta(k)^2 + 2(a+b)\zeta(k) + (a-b)^2 > 0$$

Therefore, the two roots of the polynomial are real, implying that the polynomial crosses the zero line.

Let us now study $P(1)$

$$P(1) = 1 - Tr + Det = 1 - \left(\zeta(k) + \frac{1-\delta}{(1+n)(1+\gamma)} + \varepsilon\frac{1-\alpha}{\alpha}\right) + \varepsilon\frac{1-\alpha}{\alpha}\frac{1-\delta}{(1+n)(1+\gamma)}$$

which may be rewritten as

$$P(1) = \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right) - \zeta(k)$$

From proposition 7, we now that, in the case of multiplicity, two steady states are distributed on both side of $\widehat{k}$ which is defined by

$$\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\overline{\nu}\Phi p^{\varphi-1}\frac{\alpha}{1-\alpha}(A(1-\alpha)\overline{\theta}^{1-\alpha}p)^{\frac{1}{\alpha-\varepsilon(1-\alpha)}}\widehat{k}^{\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}-1} = 1$$

which rewrites

$$\frac{\alpha}{\alpha-\varepsilon(1-\alpha)}\frac{(1+\gamma)(1+n)}{(1+\gamma)(1+n)-(1-\delta)}\zeta(\widehat{k}) = 1 \Longleftrightarrow \zeta(\widehat{k}) = \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right)$$

Then let us consider a situation where $\nu < \overline{\nu}$, implying that $k < \widehat{k}$. Since $\zeta(.)$ is an increasing function of k, we have

$$\begin{aligned}
P(1) &> \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right) - \zeta(\widehat{k}) \\
&> \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right) - \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right) \\
&> 0
\end{aligned}$$

Further note that the derivative of the polynomial $P'(\lambda) = 2\lambda - Tr$ then equals zero for $\lambda = Tr/2 > 0$. Since

$$Tr = \zeta(k) + \frac{1-\delta}{(1+n)(1+\gamma)} + \varepsilon\frac{1-\alpha}{\alpha}$$

and $\zeta(.)$ is an increasing function of k, we have for$k < \widehat{k}$ $(\nu < \overline{\nu})$

$$\begin{aligned}
0 < Tr &< \zeta(\widehat{k}) + \frac{1-\delta}{(1+\gamma)(1+n)} + \varepsilon\frac{1-\alpha}{\alpha} \\
&< \left(1 - \frac{\varepsilon(1-\alpha)}{\alpha}\right)\left(1 - \frac{1-\delta}{(1+\gamma)(1+n)}\right) + \frac{1-\delta}{(1+\gamma)(1+n)} + \varepsilon\frac{1-\alpha}{\alpha} \\
&< 1 + \frac{\varepsilon(1-\alpha)(1-\delta)}{\alpha(1+\gamma)(1+n)} \\
&< 2
\end{aligned}$$

Since $Tr < 2$, we have $\lambda \in (0,1)$ implying that $P'(\lambda)$ shifts from negative to positive values between zero and 1. Since $P(0) = Det > 0$, $P(1) > 0$ and the two roots are real, and given $P'(\lambda)$ switches from negative to positive in $(0,1)$, we know that the two roots lie within the unit circle. Therefore, the steady state lying below $\widehat{k}$ is stable.

Applying the same reasoning to the steady state above $\widehat{k}$, we find it unstable.

Therefore, for a given long run propensity to accumulate, any economy starting with an initial capital labor ratio above the central steady state will converge to the upper steady state — therefore fully specializing in good 2 — while if it starts just below the central steady state it will converge to the lower steady state.

Q.E.D □

Proposition 9: Before establishing the proposition, we need first to characterize the behavior of firms in a given economy. Under free trade, the firm attempts to solve

$$\max_{\sigma_{k,t},K_t,Q_{jt}(\ell);j=1,2,\ell\in(0,1)} Z_{1t}+p_tZ_{2t}-q_tK_t-\int_0^1\left(p^{\mathrm{Q}}_{1t}(\ell)Q_{1t}(\ell)+p^{\mathrm{Q}}_{2t}(\ell)Q_{2t}(\ell)\right)\mathrm{d}\ell$$

subject to $0\leqslant\sigma_{k,t}\leqslant 1$. This yields the set of conditions:

$$q_t=\alpha_1\frac{Z_{1t}}{K_t}+\alpha_2p\frac{Z_{2t}}{K_t} \tag{43}$$

$$\alpha_1\frac{Y_t}{\sigma_{k,t}}-\alpha_2p\frac{Y_t}{1-\sigma_{k,t}}+\lambda^0_{k,t}-\lambda^1_{k,t}=0 \tag{44}$$

$$p^{\mathrm{Q}}_{1t}(\ell)=(1-\alpha_1)Z_{1t}Q_{1t}(\ell)^{\rho-1}M_{1t}^{-\rho} \tag{45}$$

$$p^{\mathrm{Q}}_{2t}(\ell)=(1-\alpha_2)p_tZ_{2t}\theta^{\mathrm{s}}_t(\ell)^{\rho}Q_{2t}(\ell)^{\rho-1}M_{2t}^{-\rho} \tag{46}$$

$$\lambda^0_{k,t}\sigma_{k,t}=0 \tag{47}$$

$$\lambda^1_{k,t}(1-\sigma_{k,t})=0 \tag{48}$$

The intermediate firm then maximizes

$$\max_{\{L_t\geqslant 0,\sigma_t(\ell)\in[0,1]\}} p^{\mathrm{Q}}_{1t}(\ell)\sigma_t(\ell)\Gamma_tL_t(\ell)+p^{\mathrm{Q}}_{2t}(\ell)(1-\sigma_t(\ell))\Gamma_tL_t(\ell)-W_tL_t(\ell) \tag{49}$$

subject to (45) and (46). The set of first order conditions is then given by

$$W_t=\rho(1-\alpha_1)Z_{1t}\frac{Q_{1t}(\ell)^{\rho}}{L_t(\ell)}M_{1t}^{-\rho}+\rho(1-\alpha_2)\Gamma_tp_tZ_{2t}\theta^{\mathrm{s}}_t(\ell)^{\rho}\frac{Q_{2t}(\ell)^{\rho}}{L_t(\ell)}M_{2t}^{-\rho} \tag{50}$$

$$\rho(1-\alpha_1)Z_{1t}\frac{Q_{1t}(\ell)^{\rho}}{\sigma_t(\ell)}M_{1t}^{-\rho}-\rho(1-\alpha_2)p_tZ_{2t}\theta^{\mathrm{s}}_t(\ell)^{\rho}\frac{Q_{2t}(\ell)^{\rho}}{1-\sigma_t(\ell)}M_{2t}^{-\rho}+\lambda^0_{L,t}-\lambda^1_{L,t}=0 \tag{51}$$

$$\lambda^0_{L,t}\sigma_{L,t}=0 \tag{52}$$

$$\lambda^1_{L,t}(1-\sigma_{L,t})=0 \tag{53}$$

In a symmetric equilibrium, we have

$$\begin{aligned}\sigma_{L,t}(\ell)&=\sigma_t & L_t(\ell)&=L_t\\ Q_{1t}(\ell)&=Q_{1t} & Q_{2t}(\ell)&=Q_{2t}\end{aligned}$$

Furthermore, rational expectations and symmetric equilibrium imply that $\theta^{\mathrm{s}}_t(\ell)=\theta_t(\ell)=\theta_t$, such that $M_{1t}=Q_{1t}=\sigma_{L,t}\Gamma_tL_t$ and $M_{2t}=\theta_tQ_{2t}=\theta_t(1-\sigma_{L,t})\Gamma_tL_t$. Hence, the labor allocation choice becomes

$$\begin{aligned}&\rho(1-\alpha_1)Z_{1t}(\Gamma_tL_t)^{\rho}\sigma_t(\ell)^{\rho-1}(\sigma_{L,t}\Gamma_tL_t)^{-\rho}\\ -\;&\rho(1-\alpha_2)p_tZ_{2t}\theta_t^{\rho}(\Gamma_tL_t)^{\rho}(1-\sigma_{L,t})^{\rho-1}(\theta_t(1-\sigma_{L,t})\Gamma_tL_t)^{-\rho}\\ +\;&\lambda^0_{L,t}-\lambda^1_{L,t}=0\end{aligned} \tag{54}$$

An interior solution for σ_k and σ_L implies, from (44) and (54)

$$\frac{\sigma_{k,t}}{\sigma_{L,t}}=\frac{\alpha_1(1-\alpha_2)}{\alpha_2(1-\alpha_1)}\times\frac{1-\sigma_{k,t}}{1-\sigma_{L,t}} \tag{55}$$

We can now formally prove the proposition. A simple way to check that output–per–worker is locally convex is to show that the user cost of capital is increasing in capital. In a non–specialized symmetric equilibrium, the user cost of capital is given by

$$q_t=\alpha_1\frac{z_{1,t}}{k_t}+\alpha_2p\frac{z_{2,t}}{k_t}$$

where lowercases denote aggregate levels deflated for both technological progress and population growth. Recall that in a symmetric equilibrium, the allocation of capital and labor, in the non specialized area, is determined by

$$\alpha_1 \frac{z_{1t}}{\sigma_{k,t}} \leqslant \alpha_2 p_t \frac{z_{2t}}{1-\sigma_{k,t}} \tag{56}$$

$$(1-\alpha_1)\frac{z_{1t}}{\sigma_{L,t}} \leqslant (1-\alpha_2) p_t \frac{z_{2t}}{1-\sigma_{L,t}} \tag{57}$$

Therefore, the user cost of capital rewrites

$$\begin{aligned} q_t &= \alpha_1 \frac{z_{1,t}}{k_t}\frac{\sigma_{k,t}}{\sigma_{k,t}} + \alpha_2 p \frac{z_{2,t}}{k_t}\frac{1-\sigma_{k,t}}{1-\sigma_{k,t}} \\ &= \alpha_1 \frac{z_{1,t}}{\sigma_{k,t}k_t}\sigma_{k,t} + \alpha_1 \frac{z_{1,t}}{\sigma_{k,t}k_t}(1-\sigma_{k,t}) \\ &= \alpha_1 \frac{z_{1,t}}{\sigma_{k,t}k_t} = \alpha_1 A_1 \left(\frac{\sigma_{k,t}k_t}{\sigma_{L,t}}\right)^{\alpha_1-1} \end{aligned}$$

Using (55) , one can express $\sigma_{k,t}/\sigma_{L,t}$ as a function of $\sigma_{L,t}$ only

$$\frac{\sigma_{k,t}}{\sigma_{L,t}} = \frac{\alpha_1(1-\alpha_2)}{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}} \tag{58}$$

Therefore, q_t simplifies to

$$q_t = \alpha_1 A_1 k_t^{\alpha_1-1}\left(\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{\alpha_1(1-\alpha_2)}\right)^{1-\alpha_1} \tag{59}$$

Then

$$\frac{dq_t}{dk_t} = \frac{\partial q_t}{\partial k_t} + \frac{\partial q_t}{\partial \sigma_{k,t}}\frac{\partial \sigma_{k,t}}{\partial k_t}$$

Using (59), we find

$$\begin{aligned} \frac{\partial q_t}{\partial k_t} &= \alpha_1(\alpha_1-1)\frac{z_{1,t}}{\sigma_{k,t}k_t^2} \\ \frac{\partial q_t}{\partial \sigma_{k,t}} &= -\frac{\alpha_1(1-\alpha_1)(\alpha_2-\alpha_1)}{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}} \times \frac{z_{1,t}}{\sigma_{k,t}k_t} \end{aligned}$$

Therefore

$$\frac{dq_t}{dk_t} = -\alpha_1(1-\alpha_1)\frac{z_{1,t}}{\sigma_{k,t}k_t}\left(\frac{1}{k_t} + \frac{(\alpha_2-\alpha_1)}{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}\frac{\partial \sigma_{k,t}}{\partial k_t}\right)$$

Note that since $0 \leqslant \alpha_1, \alpha_2 \leqslant 1$, $\sigma_{L,t} \in (0,1)$ and $\alpha_2 \geqslant \alpha_1$, we have $\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t} \geqslant 0$ Hence, a necessary condition for the user cost of capital to be an increasing function of capital is

$$\frac{\partial \sigma_{k,t}}{\partial k_t} \leqslant -\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{(\alpha_2-\alpha_1)k_t}$$

We thus have to compute $\partial\sigma_{k,t}/\partial k_t$.

In a symmetric equilibrium, $(\sigma_{L,t}, \sigma_{k,t})$ is obtained from the system (56)–(57). The ratio of the two relations yields (55), that can be plugged into (57) to give

$$(1-\alpha_1)A_1\left(\frac{\sigma_{k,t}}{\sigma_{L,t}}k_t\right)^{\alpha_1} = (1-\alpha_2)A_2\overline{\theta}^{1-\alpha_2}\left(\frac{\alpha_2(1-\alpha_1)}{\alpha_1(1-\alpha_2)}\frac{\sigma_{k,t}}{\sigma_{L,t}}k_t\right)^{\alpha_2}(1-\sigma_{L,t})^{\varepsilon(1-\alpha_2)}$$

Making use of (58), the latter expression rewrites

$$\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t} = \Omega(p)k_t(1-\sigma_{L,t})^{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}} \tag{60}$$

where

$$\Omega(p) = \left[p\frac{\alpha_2^{\alpha_2}(1-\alpha_2)^{1-\alpha_1}}{\alpha_1^{\alpha_1}(1-\alpha_1)^{1-\alpha_2}}\overline{\theta}^{1-\alpha_2}\frac{A_1}{A_2}\right]^{\frac{1}{\alpha_2-\alpha_1}}$$

Differentiating (60) with respect to k_t, we get

$$-(\alpha_2-\alpha_1)\frac{\partial\sigma_{L,t}}{\partial k_t} = \Omega(p)(1-\sigma_{L,t})^{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}} - \frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\Omega(p)k_t(1-\sigma_{L,t})^{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}-1}\frac{\partial\sigma_{L,t}}{\partial k_t}$$

which, from (60), rewrites

$$\frac{\partial\sigma_{L,t}}{\partial k_t} = \frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{k_t} \times \frac{1}{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{1-\sigma_{L,t}} - (\alpha_2-\alpha_1)}$$

Then

$$\frac{\partial\sigma_{k,t}}{\partial k_t} \leqslant -\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{(\alpha_2-\alpha_1)k_t}$$

amounts to

$$\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{k_t} \times \frac{1}{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{1-\sigma_{L,t}} - (\alpha_2-\alpha_1)} \leqslant -\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{(\alpha_2-\alpha_1)k_t}$$

which, since $\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t} \geqslant 0$, simplifies to

$$\frac{\alpha_2-\alpha_1}{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{1-\sigma_{L,t}} - (\alpha_2-\alpha_1)} \leqslant -1$$

or

$$\frac{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{1-\sigma_{L,t}}}{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t}}{1-\sigma_{L,t}} - (\alpha_2-\alpha_1)} \leqslant 0 \tag{61}$$

Note that (58) implies

$$\alpha_2(1-\alpha_1)-(\alpha_2-\alpha_1)\sigma_{L,t} = \alpha_1(1-\alpha_2)\frac{\sigma_{L,t}}{\sigma_{k,t}} = \alpha_2(1-\alpha_1)\frac{1-\sigma_{L,t}}{1-\sigma_{k,t}}$$

such that (61) rewrites

$$\frac{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)}{1-\sigma_{K,t}}}{\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)}{1-\sigma_{K,t}} - (\alpha_2-\alpha_1)} \leqslant 0$$

Since $\varepsilon \geqslant 0$, $0 \leqslant \alpha_1, \alpha_2 \leqslant 1$, $\alpha_2 \geqslant \alpha_1$, the numerator is positive, such that a condition for $\frac{\partial q_t}{\partial k_t} > 0$ is

$$\frac{\varepsilon(1-\alpha_2)}{\alpha_2-\alpha_1}\frac{\alpha_2(1-\alpha_1)}{1-\sigma_{K,t}} - (\alpha_2-\alpha_1) \leqslant 0$$

or

$$\varepsilon \leqslant \frac{(\alpha_2-\alpha_1)^2}{\alpha_2(1-\alpha_1)(1-\alpha_2)}(1-\sigma_{K,t}) \leqslant \frac{(\alpha_2-\alpha_1)^2}{\alpha_2(1-\alpha_1)(1-\alpha_2)}$$

which completes the proof.

□ Q.E.D