

IFN Working Paper No. 850, 2010

The World Distribution of Productivity: Country TFP Choice in a Nelson-Phelps Economy

Erika Färnstrand Damsgaard and Per Krusell

The World Distribution of Productivity: Country TFP Choice in a Nelson-Phelps Economy*

Erika Färnstrand Damsgaard[†] Per Krusell[‡]

September 22, 2010

Abstract

This paper builds a theory of the shape of the distribution of total-factor productivity (TFP) across countries. The data on productivity suggests vast differences across countries, and arguably even has “twin peaks”. The theory proposed here is consistent with vast differences in long-run productivity, and potentially also with a twin-peaks outcome, even under the assumption that all countries are ex-ante identical. It is based on the hypothesis that TFP improvements in a given country follow a Nelson-Phelps specification. Thus, they derive from past investments in the country itself and, through a spillover (or catch-up) term, from past investments in other countries. We then construct a stochastic dynamic general equilibrium model of the world which has externalities: each country invests in TFP and internalizes the dynamic effects of its own investment, while treating other countries’ investments as given. Average world growth is endogenous, as is the distribution of TFP across countries. We find that small idiosyncratic TFP shocks can lead to large long-run differences in TFP levels and that, in the long run, the world distribution of TFP across countries may be asymmetric, i.e., twin-peaked, or bimodal. More specifically, twin-peaked world distributions of TFP arise if the catch-up term in the Nelson-Phelps equation has a sufficiently low weight. If, on the other hand, technological catch-up is important, the world distribution of TFP is unimodal, though it may still have large dispersion.

Keywords: growth, inequality

JEL classification: F43, O1

*Erika Färnstrand Damsgaard gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Research Foundation and the Gustaf Douglas Research Program on Entrepreneurship at IFN. Per Krusell gratefully acknowledges financial support from the European Research Council. We would like to thank Fabrice Collard, John Hassler, and seminar participants at the IIES Brown Bag Seminar, the University of Munich, Queens University, the NHH-UiO Workshop in Economic Dynamics 2007, the 8th SAET Conference 2007, the Swedish Riksbank, the SED Annual Meeting 2008, and the 6th Hydra Workshop on Dynamic Macroeconomics 2008 for valuable comments and suggestions. All remaining errors are our own.

[†]Research Institute of Industrial Economics (IFN). E-mail: erika.farnstrand.damsgaard@ifn.se.

[‡]Institute for International Economic Studies, CAERP, CEPR, and NBER. E-mail: per.krusell@iies.su.se.

1 Introduction

What explains the relative levels of riches across countries? Building on the approach used in Klenow and Rodríguez-Clare (1997) and Hall and Jones (1999), Jones (2008) offers a recent account attributing the bulk of observed income-per-capita differences among countries to “TFP”, or Total-Factor-Productivity, i.e., essentially to residually measured productivity differences. TFP differences are residual in that they capture the productivity differences that remain after differences in physical and human capital (per capita, or per worker) have been measured and taken into account, using the assumption of an aggregate Cobb-Douglas production function. Thus, under the assumptions maintained in these accounting analyses, to understand differences in income across countries it is not sufficient to understand differences in rates of accumulation of physical or human capital, at least if the available measures of physical and human capital are accurate. Therefore, much of the growth theory that analyzes such accumulation may “miss the point”, if the purpose is to explain relative income differences across countries quantitatively. In conclusion, it seems reasonable to concur with the way Prescott (1998) puts it: “Needed: A Theory of Total Factor Productivity”. The present paper proposes one direction in which theories with this aim might be built.

We also take the point of view that a theory of the world distribution of TFP should, ideally, display two features. One is that long-run world TFP growth is endogenous, i.e., “nontrivially” determined within the model. Thus, in this sense we very much take the view of the endogenous growth literature (Romer (1986), Lucas (1988), Rebelo, (1991), and others). Whether the long-run world growth *rate* (or perhaps just the level increments) is endogenous is perhaps not central; for simplicity here though, we assume that it is. The second feature that we view as essential is that whereas each given country can influence its relative TFP level, it can do little to influence its own long-run rate of TFP growth. More precisely, we view countries as small and their long-run growth rates as determined by the rate at which “world TFP” (i.e., some average, or perhaps the frontier) grows. This view is partly based on there only having been rather minor increases in the dispersion among countries during the period over which there is reliable country income data (essentially, Heston, Summers and Aten (2006)). In addition, modern economies are arguably (i) highly dependent on world technological developments, as evidenced by, among others, Benhabib and Spiegel (1994), who find productivity *catch-up* to be statistically significant for growth of TFP; and (ii) open in the sense that even if conventional trade is not entirely free, technology spillovers flow rather freely across countries in the world. In sum, we view the world average growth rate as endogenous but, conditional on this rate, any individual country’s growth rate as exogenous.

Given this perspective, we construct a theory of the world TFP distribution that has at its core the specification of technology advancement proposed in Nelson and Phelps (1966). They hypothesized a form of technology, or human-capital, catch-up: in a country context, the growth rate of technology, or human capital, in a given country can be increased if this country invests, and the further is the distance from the technology frontier, the more

productive is such an investment. That is, if a country is further behind, the potential for rapid growth is higher, since the country can “free-ride” on technologies/human capital accumulated elsewhere. A Nelson-Phelps view of development is indeed taken explicitly in Jones’s (1998) textbook on economic growth: he embeds the Nelson-Phelps formulation into a developing-country growth model, and uses it as a framework for productivity accounting and growth dynamics. In this paper we thus take the Nelson-Phelps-Jones perspective on development but then go on to ask a further question: if countries are subject to this “technology for productivity growth”, and each country operates the technology optimally from the viewpoint of maximizing the utility of its citizens, what is the implied *equilibrium world distribution of country TFP*? Since the catch-up term by definition means that countries benefit from spillovers, there is an obvious force for convergence, but how does this force play out in equilibrium?

Our main findings can be described as follows. Although convergence forces are always present in the model, if the catch-up term is weak enough, the stable long-run world distribution of TFP is not single-peaked but bimodal. There is one group of countries with high TFP in relative terms, with the remainder of the countries operating at a much lower TFP level. All countries grow at the same rate, but the high-TFP countries invest more in technology than do low-TFP countries. The catch-up term thus allows the low-TFP countries to grow at the same rate and not fall further behind in relative terms. The bimodality is first derived using a model without idiosyncratic (country-specific) TFP shocks, and with this setting bimodality means that the long-run TFP distributions have two (and exactly two) groups of countries, within each of which all countries have the same TFP level. In the model with shocks, in contrast, there is a “smooth” long-run distribution, and this distribution can have two visible peaks or just one. However, even when there is a single peak in the TFP distribution, and where the corresponding long-run, no-shocks outcome has a single group of countries with identical TFP, the distribution can have significant dispersion around the mode. Thus, the theory embodies strong forces pulling countries apart.

Formally, in our model, individual countries can invest in a technology-enhancing input, e , in order to increase their TFP, T . We think of this investment input as a traded one (such as educated workers), which can thus be allocated across countries, and there will be an equilibrium world price for this input. The accumulation of TFP, as mentioned, is of the Nelson-Phelps (1966) form: TFP growth in a given country depends positively on investment in TFP and, in addition, on the country’s distance to the world technology frontier, which generates catch-up. The formal specification is that $\frac{T_{i,t+1}}{T_{i,t}} = (\frac{\bar{T}_t}{T_{i,t}})^\gamma H(e_{i,t})$, where i is a country index and t stands for time; \bar{T} is world average (or frontier) TFP, and the function H , which is increasing, describes how e units translate into TFP growth units. In the model, all countries are symmetric; they have identical technologies and only potentially differ by their initial conditions. The distribution of countries over TFP levels is determined by two counteracting forces. First, the technological catch-up of less developed countries generates convergence. Second, the internalization of country-specific dynamic gains from TFP investment generates divergence. Thus, in terms of this specification, we find that the

distribution of TFP can be bimodal, provided that the weight on the catch-up factor in the Nelson-Phelps equation, γ , is sufficiently close to zero.

The intuition why catch-up is a force for convergence is straightforward. Why, though, is there a force for divergence if γ is small? Consider the extreme case, i.e., $\gamma = 0$. Here, we can write the TFP technology as $T_{i,t+1} = T_{i,t}H(e_{i,t})$, and it is clear that, as a dynamic technology for accumulation, the system has increasing returns to scale, as long as H is increasing: if $T_{i,t}$ and $e_{i,t}$ both double, $T_{i,t+1}$ more than doubles, or doubles exactly if H is flat. Thus, when individual country planners operate technologies of this sort, there are dynamic increasing returns to scale, and thus internalizing these gains is important and leads to a force for divergence: those countries with initially (or, by idiosyncratic stochastic events) high T_i levels have an advantage in further growth.

How can our findings be used to interpret the available country data? Aside from providing a theory of relative TFPs in the world (and a potential for bimodality), we use the theory to speculate on the findings in Acemoglu (2009). He plots a sequence of distributions of countries according to GDP per worker, for 1960, 1980, and 2000, and it reveals some visible changes in the shape of the distribution and, in particular, in the “degree of bimodality”. Thus, the 1980 and 2000 distributions are “more bimodal” than is the 1960 distribution. In terms of the model, a move toward bimodality can for example be due to (i) a decrease in effective spillovers (a lower γ); (ii) a decrease in the cost of trading e ; (iii) increases in the extent to which e affects TFP (a change in the shape of H); or (iv) transitional dynamics. We interpret and compare these channels in Section 8 of the paper.

Our present setting is not designed as a full-fledged quantitative assessment of the kind of theory we propose; among other things, we abstract from capital accumulation (thus, TFP and labor productivity are synonymous). Perhaps more importantly, a serious quantitative model of world TFP would require more geographic detail than our setting with (ex-ante) identical countries, since we envisage the critical distinction here—between which TFP investments are internalized and which are not—to coincide with country borders. Thus, for example, γ should in principle be determined based on how large countries are, and it would arguably also be more appropriate with a model which has heterogeneity in γ across countries. Such an extension, and a full quantitative investigation, would be an interesting one to pursue. Thus, our present analysis is to be interpreted as a first pass at how “rational TFP accumulation decisions on the country level”, in combination with technology spillovers as proposed by Nelson-Phelps, deliver a model of relative riches in the world.

The paper is organized as follows. The related literature is presented in Section 2. Section 3 describes the model. The balanced growth equilibrium of the model is defined in Section 4. Section 5 describes the symmetric balanced growth equilibrium, and analyses its stability. Asymmetric balanced growth equilibria and their stability properties are characterized in Section 6. In Section 7, the model is extended to allow for country-specific shocks to TFP. Section 8 entails an analysis of what model parameters might change the distribution of relative TFPs over time, and Section 9 concludes.

2 Related literature

In an influential article, Nelson and Phelps (1966) argued that in an economy with technological change, the more educated the workforce is, the faster new technologies of production will be introduced. The argument was formalized in a model where advancement of technology depends positively upon investment in education and upon the gap between the best-practice, or frontier, technology and the technology currently used. Parente and Prescott (1994) incorporate a similar mechanism into a model of an economy featuring firms, households, and a government. They propose that differences in barriers to technology adoption can account for the observed income disparities across countries. In the model, a firm can invest in adoption of more productive technologies, and the amount of investment needed depends on the barriers to technology adoption in the country where it operates. Parente and Prescott calibrate the model and argue that the differences in barriers required to account for the observed cross-country income differences are not implausibly large. Similarly, Jones (1998) endogenizes technology diffusion in the Romer model, by introducing the model proposed by Nelson and Phelps. He assumes that a country's human capital or skill accumulation depends on investment in education as well as technological spillover from the technology frontier. While Parente and Prescott (1994) and Jones (1998) introduce the Nelson-Phelps framework into growth models, both explore the implications for income levels and growth rates in a partial equilibrium setting. In this paper, we extend the analysis to general equilibrium.

Large differences in per capita income across countries as well as a “twin-peaked” distribution of world income has been documented by for example Quah (1993), Quah (1997), and Kremer, Onatski, and Stock (2001). Several economists have constructed models aimed at explaining these empirical findings. An example is Chari, Kehoe and McGrattan (1997) which uses a neoclassical growth model to determine how much of the variation in incomes across countries that can be explained by distortions to capital accumulation. The distortions are modeled as a stochastic process for the price of investment relative to consumption. Similarly, Acemoglu and Ventura (2002) explain the world income distribution by accumulation of capital in combination with international trade and specialization. However, technology, rather than physical or human capital, appears to be the main determinant of the differences in incomes. For example, Klenow and Rodríguez-Clare (1997) find that about 50 percent of cross-country variation in GDP per worker is explained by differences in productivity. Similarly, Jones (2008) finds that approximately one-third of differences in income across countries can be explained by differences in capital per person, while differences in TFP explain the remaining two-thirds. Therefore, along with Parente and Prescott (1994), a number of models have been developed to explain income differences by modeling differences in productivity rather than in accumulation of capital.

Several models constructed to explain the differences in income across countries have used the Schumpeterian growth model as a point of departure. For example, Howitt (2000) analyzes a multi-country version of the Aghion-Howitt endogenous growth model with technology transfer across countries. Under the assumption that the countries differ in their R&D

productivities, R&D subsidy rates, or investment rates, the model can generate “club convergence”, whereby countries which invest in R&D will converge to parallel growth paths, and countries which do not will stagnate. Howitt and Mayer-Foulkes (2005) use a similar model with technology transfer across countries. However, the extent to which a country benefits from technology transfer depends on its “absorptive capacity”. In that model, countries sort into three convergence groups characterized by R&D, implementation and stagnation, respectively. Aghion, Howitt and Mayer-Foulkes (2005) introduce credit market imperfections in the multi-country Schumpeterian model. The model exhibits technology transfer, where an investment in R&D is a prerequisite for the receiving country to benefit from the transfer. It is assumed that R&D requires access to external finance, at a cost which depends on the level of financial development. The model predicts that countries above a certain threshold of financial development will converge to a high growth rate, whereas other countries converge to strictly lower growth rates. Lucas (2000), finally, also analyzes a model of world income based on technology diffusion. He calibrates the model and finds that the predictions are consistent with “club convergence” over the postwar period.

The present paper also models investments in productivity, but aside from taking an explicit Nelson-Phelps view on technology diffusion, it differs from the models described above in two additional ways. First, rather than assuming that countries have different characteristics, it treats all countries symmetrically. In addition, it models internalization of the dynamic gains from productivity investments at a country level.

3 Model

The world consists of a continuum of countries indexed i . Each country produces output, and invests in TFP accumulation in order to increase future output. The investment in TFP can be R&D, technology adoption, improving institutions etc. Each country is endowed with both low-skilled and high-skilled labor, where low-skilled labor is used in production of output and high-skilled labor in the accumulation of TFP. The number of high-skilled workers employed in country i at time t is denoted $e_{i,t}$. It is assumed that low-skilled labor is immobile while high-skilled labor flows freely across countries. The total amount of high-skilled labor in the world is fixed, and it is equal to e_W . All countries are of the same size. Country i 's endowment of low-skilled labor is normalized to one, and its endowment of high-skilled labor is equal to the world total (and average), e_W .

Country i 's income $Y_{i,t}$ is the output produced net of TFP investment costs:

$$Y_{i,t} = T_{i,t} \cdot 1 - w_t e_{i,t} + w_t e_W. \quad (1)$$

In (1), $T_{i,t}$ is country i 's level of TFP. Since the amount of low-skilled workers is normalized to 1, it is also equal to country i 's output. w_t is the world wage rate for high-skilled workers, and the investment cost constitutes of wage payments to foreign high-skilled workers. This formulation for income is admittedly simple, but it is a starting point for the analysis.

3.1 TFP accumulation

All countries have the same technology for TFP accumulation but possibly different starting levels of TFP. In each country, the TFP investment is chosen by a country planner.

In the model by Nelson and Phelps (1966), advancement of technology depends positively upon investment in education and upon the gap between the best-practice, or frontier, technology and the technology currently used. This paper follows their formulation, but views the investment as a general investment in TFP, and allows the importance of technological catch-up to vary by a parameter, γ . Consequently, TFP in country i is accumulated according to

$$\frac{T_{i,t+1}}{T_{i,t}} = \left(\frac{\bar{T}_t}{T_{i,t}} \right)^\gamma H(e_{i,t}) \quad (2)$$

where the distance to the world TFP frontier is captured by the term $\frac{\bar{T}_t}{T_{i,t}}$, and \bar{T}_t is the world average, or frontier, TFP level. Common access to the frontier TFP generates a faster catch-up the further behind a country is. The parameter $\gamma \in [0, 1]$ measures the strength of the catch-up effect. The investment in TFP is captured by $H(e_{i,t})$, the TFP production function.

It is assumed that $H(e_{i,t})$ is strictly increasing and strictly concave in $e_{i,t}$. In addition, it must satisfy the following conditions: $H(0) = 1$, $H'(0) = \infty$, and $H(e_{i,t})$ is bounded above by b , where b is not too large. The following functional form for $H(e_{i,t})$

$$H(e_{i,t}) = (b - 1) \left(1 + \frac{1}{e_{i,t}} \right)^{-\kappa} + 1 \quad (3)$$

satisfies the conditions while being relatively simple, and this specification of $H(e_{i,t})$ will be used throughout the analysis.

The world TFP average, \bar{T}_t , has the form

$$\bar{T}_t \equiv \left(\int T_{i,t}^\psi di \right)^{1/\psi}$$

The parameter ψ determines the extent to which \bar{T}_t depends on the leading, or frontier, TFP level in the world. For $\psi = 1$, \bar{T}_t is the arithmetic average of all countries' TFP levels, and for $\psi = \infty$, it is equal to the highest TFP level. The world TFP level grows according to

$$\bar{T}_{t+1} = \bar{T}_t(1 + g_t)$$

where g_t is endogenously determined.

A rewriting of the TFP accumulation function in (2) gives

$$T_{i,t+1} = T_{i,t}^{1-\gamma} \bar{T}_t^\gamma H(e_{i,t}). \quad (4)$$

From this expression, it is clear that investments in TFP have dynamic effects, some of which are specific to the country, as captured by the term $T_{i,t}$ and others which are international, as captured by \bar{T}_t . The parameter γ governs the share of these dynamic effects which is country-specific. For example, if $\gamma = 1$, dynamic gains of TFP investment arise only through technological catch-up. If $\gamma = 0$, dynamic gains of TFP investment are completely internalized within the country. The parameter γ shows to be crucial for the model's results, and will be discussed at length below. For comparison, Jones (1998) uses the same formulation as (2) but views $T_{i,t}$ as human capital and $e_{i,t}$ as education investment.

3.2 Consumers

Each country i has a dynastic household which maximizes utility given the utility function

$$U(C_i) = \sum_{t=0}^{\infty} \beta^t \log(C_{i,t}) \quad (5)$$

where β is the discount factor. The consumer is endowed with 1 unit of low-skilled labor and e_W units of high-skilled labor.

3.3 Country planner problem

The aim of this paper is to explain the distribution of TFP across countries. Therefore, the model focuses on inter-country relationships and is solved in general equilibrium, while intra-country relationships are given a cursory representation. We do not consider the aggregation of individual firms' TFP into country TFP in a given country, and output is specified only at the country level, as given by the expression in (1). Hence, it is assumed that technology flows freely within countries, and all dynamic effects are internalized within a country. Therefore, we will characterize the country planner's solution of the individual country's optimization problem. The country planner chooses a sequence of consumption allocations and investments in TFP so as to maximize consumer utility, taking the sequence of world prices and average TFP, $\{p_t, w_t, \bar{T}_t\}_{t=0}^{\infty}$ as given. The problem can be stated as follows.

$$\begin{aligned} & \max_{\{T_{i,t}, C_{i,t}, e_{i,t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log(C_{i,t}) \\ & \text{s.t.} \\ & \sum_{t=0}^{\infty} p_t C_{i,t} = \sum_{t=0}^{\infty} p_t (T_{i,t} \cdot 1 - w_t e_{i,t} + w_t e_W) \\ & T_{i,t+1} = T_{i,t}^{1-\gamma} \bar{T}_t^{\gamma} H(e_{i,t}), \end{aligned} \quad (6)$$

where p_t is the time-0 price of the time t good and $p_0 = 1$. The country planner maximizes utility of consumption given two constraints. The first is the resource constraint, and the second governs the accumulation of TFP. Since each country is assumed to be small, its TFP

choice has no effect on the average TFP level and therefore the country planner takes \bar{T}_t as given.

3.4 World equilibrium

A **world equilibrium** consists of sequences of allocations $\{T_{i,t}, C_{i,t}, e_{i,t}\}_{t=0}^{\infty}$ for all i and prices $\{w_t, p_t\}_{t=0}^{\infty}$, such that

1. $\{T_{i,t}, C_{i,t}, e_{i,t}\}_{t=0}^{\infty}$ solves the problem in (6) for all i ;
2. $\bar{T}_t = \left(\int T_{i,t}^{\psi} di \right)^{1/\psi}$ for all t and
3. $e_W = \int e_{i,t} di$ for all t .

Condition 2 states that average TFP in the world is consistent with individual countries' TFP choices. Condition 3 ensures that there is market clearing in the market for high-skilled labor.

3.5 Initial characterization of a country's investment decision

In the model, it is assumed that there are perfect world capital markets. This implies that the interest rates p_t/p_{t+1} are exogenous from the point of view of an individual country. Given this assumption, each planner's utility maximization problem can be separated into two independent problems: an income-maximization problem, and an intertemporal consumption allocation problem. We state each of them in turn.

In the income-maximization problem, the country planner chooses a sequence of future investments in TFP so as to maximize output net of investment costs,

$$\max_{\{T_{i,t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} p_t \left(T_{i,t} - w_t H^{-1} \left(\frac{T_{i,t+1}}{T_{i,t}^{1-\gamma} \bar{T}_t^{\gamma}} \right) \right), \quad (7)$$

taking the sequence of world prices and average TFP, $\{p_t, w_t, \bar{T}_t\}_{t=0}^{\infty}$, as given. The expression in (7) is obtained by inserting the expression for $e_{i,t}$ from (4) into (1). The term $w_t e_W$ can be dropped since it is constant from the individual country's point of view.

Next, we turn to the intertemporal consumption allocation problem. Given sequences of prices and country income, $\{p_t, Y_{i,t}\}_{t=0}^{\infty}$, the household in country i chooses its intertemporal consumption allocation so as to maximize

$$\begin{aligned} & \max_{\{C_{i,t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \log(C_{i,t}) \\ & s.t. \\ & \sum_{t=0}^{\infty} p_t C_{i,t} = \sum_{t=0}^{\infty} p_t Y_{i,t}. \end{aligned}$$

Optimization yields the following relationship between consumption growth and prices

$$\frac{C_{i,t+1}}{C_{i,t}} = \beta \frac{p_t}{p_{t+1}}. \quad (8)$$

In a world equilibrium, country and world consumption growth between t and $t + 1$ equals $\beta(p_t/p_{t+1})$. Using the separation of the optimization problem, the world equilibrium can thus be redefined as follows.

A **world equilibrium** thus consists of sequences of allocations $\{T_{i,t}, C_{i,t}, e_{i,t}\}_{t=0}^{\infty}$ for all i and prices $\{w_t, p_t\}_{t=0}^{\infty}$, such that

1. $\{T_{i,t}\}_{t=0}^{\infty}$ solves the problem in (7) for all i , and $e_{i,t} = H^{-1} \left(\frac{T_{i,t+1}}{\bar{T}_{i,t}^{1-\gamma} \bar{T}_t^\gamma} \right)$;
2. $\bar{T}_t = \left(\int T_{i,t}^\psi di \right)^{1/\psi}$ for all t ;
3. $e_W = \int e_{i,t} di$ for all t ;
4. $\{C_{i,t}\}_{t=0}^{\infty}$ is given by

$$\sum_{t=0}^{\infty} p_t C_{i,t} = \sum_{t=0}^{\infty} p_t Y_{i,t}$$

and $\frac{C_{i,t+1}}{C_{i,t}} = \beta \frac{p_t}{p_{t+1}}$ for all i, t .

In Section 5, the world equilibrium will also be defined recursively, for the special case of a distribution consisting of two groups of countries.

4 Balanced growth equilibria

This analysis will focus on the long-run world distributions of TFP. Therefore, we restrict our attention to balanced growth equilibria. A balanced growth equilibrium is a world equilibrium, as defined in Section 3.5, in which all variables grow at constant rates. There is a common world growth rate g of \bar{T}_t , $T_{i,t}$, $C_{i,t}$ and w_t . The common growth rate allows us to define a TFP-adjusted wage for high-skilled workers, \hat{w} :

$$\hat{w} \equiv \frac{w_0}{\bar{T}_0} = \frac{w_t}{\bar{T}_t}$$

for all t .

Similarly, the relationship in (8) can be rewritten as

$$\frac{p_t}{p_{t+1}} = \frac{1+g}{\beta}. \quad (9)$$

Before defining the balanced growth equilibrium, we restate the optimization problem in terms of relative TFP levels, $z_{i,t}$:

$$z_{i,t} \equiv \frac{T_{i,t}}{\bar{T}_t}.$$

This implies that (4) can be expressed as

$$z_{i,t+1} = z_{i,t}^{1-\gamma} \frac{H(e_{i,t})}{1+g}. \quad (10)$$

Using (9) and the variables thus defined, the country planner's income-maximization problem in a balanced growth equilibrium can be stated as follows. The country planner chooses a sequence of future relative TFP levels so as to maximize output net of investment costs

$$\max_{\{z_{i,t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left(z_{i,t} - \hat{w} H^{-1} \left(\frac{z_{i,t+1}}{z_{i,t}^{1-\gamma}} (1+g) \right) \right)$$

taking \hat{w} and g as given.

The solution to the optimization problem above results in the following Euler equation (where subscript i is omitted)

$$\frac{\hat{w}}{H'(e_t)} \frac{1}{z_t^{1-\gamma}} = \frac{\beta}{1+g} + \frac{\hat{w}\beta(1-\gamma)}{H'(e_{t+1})} \frac{z_{t+2}}{z_{t+1}^{2-\gamma}}. \quad (11)$$

As can be seen from (11), there are three effects of an increase in the relative TFP level on income. First, there is an increase in investment costs at time t , as captured by the term on the left-hand side. Second, there is an increase in output at time $t+1$, corresponding to the first term on the right-hand side. Third, there is a decrease in investment costs at time $t+1$, as given by the second term on the right-hand side. The last effect depends directly on γ . For low values of γ , a large part of the dynamic gains to TFP are country-specific, and investment in TFP today generates large decreases in future investment costs. As γ increases, the country-specific gains decrease, and catch-up with the frontier becomes relatively more important.

4.1 Definition of a balanced growth equilibrium

Formally, a balanced growth equilibrium, a BGE, is a world equilibrium, as defined in Section 3.5, such that $p_t = \beta^t$, $e_{i,t} = e_i \forall i, t$ and $T_{i,t} = T_i(1+g)^t \forall i, t$ for $g > 0$.

What is the distribution of countries over relative TFP levels in a balanced growth equilibrium? Let this distribution be described by $\Gamma(z)$, a probability measure on (S, β_s) where $S \in [z_{\min}, z_{\max}]$ and β_s is the associated Borel σ -algebra. This measure will be discussed further below. Using $\Gamma(z)$, we can redefine the balanced growth equilibrium in recursive notation.

A **balanced growth equilibrium** consists of a stationary probability measure $\Gamma(z)$, variables \hat{w} and g , and functions $v(z)$ and $E(z)$ such that

1. $\forall z, v(z)$ solves

$$v(z) = \max_e z - \hat{w}e + \beta v \left(z^{1-\gamma} \frac{H(e)}{1+g} \right);$$

2. $\forall z, E(z)$ is

$$E(z) = \arg \max_e v(z);$$

3. $\int_S \left(\frac{z^{1-\gamma}}{1+g} H(E(z)) \right)^{1/\psi} d\Gamma(z) = 1;$

4. $\int_S E(z) d\Gamma(z) = e_W;$ and

5. $\Gamma(z)$ satisfies

$$\Gamma(B) = \int_{z \in S: \frac{z^{1-\gamma}}{1+g} H(E(z)) \in B} d\Gamma(z) \quad \forall B \in \beta_S.$$

The first and second conditions give the value function and policy function, respectively. The third condition states that the integral over all relative TFP levels must equal 1. The fourth condition is the market-clearing condition for high-skilled workers. The last condition ensures that the probability measure $\Gamma(z)$ is stationary. $\Gamma(z)$ is stationary if, for each set $B \in \beta_S$, the distribution of countries over relative TFP levels is time-invariant.

The astute reader has noticed that the consumption allocation is absent from the definition of the balanced growth equilibrium above. The reason is twofold. First, the relationship between consumption growth and prices, as specified in (9), is already embedded in the definition of $v(z)$. Second, given the assumption of perfect capital markets, the distribution of consumption across countries is independent of the relative TFP levels z . In a balanced growth equilibrium, the level of consumption for a country is given by the consumer's intertemporal consumption allocation problem. Although consumption growth is identical in all countries, the level of consumption in a given country depends on initial conditions.

An obvious candidate for a balanced growth equilibrium is one in which all countries behave identically: a symmetric one. In a symmetric BGE, or SBGE, the measure $\Gamma(z)$ is degenerate with all its mass at $z = 1$. For any other shape of $\Gamma(z)$ it must be the case that the function $E(z)$ implies multiple stationary points.¹ We will discuss the different possible outcomes in turn, starting with the SBGE.

¹A stationary point is a stationary solution to (10), or in recursive terms; $z = z^{1-\gamma} \frac{H(E(z))}{1+g}$.

5 Symmetric balanced growth equilibria

A candidate for an SBGE has the following two characteristics. First, all countries choose a level of employment of high-skilled labor equal to the world average, e_W . Second, the world distribution of country relative TFP levels, $\Gamma(z)$, is degenerate at $z_i = 1$ for all i , thus it is trivially unimodal.

In an SBGE, the world growth rate g is determined by (10), which evaluated in $z_{t+1} = z_t = 1$ gives

$$g = H(e_W) - 1.$$

Is this allocation optimal for each country? The TFP-adjusted wage rate for high-skilled workers, \hat{w} , will be such that the first-order condition from the country planner's optimization problem is satisfied. Therefore, the Euler equation, (11), gives the wage rate as

$$\hat{w} = \frac{\beta H'(e_W)}{(1+g)(1-\beta(1-\gamma))}.$$

As a result of the symmetry across countries, both g and \hat{w} are determined by e_W , the average number of high-skilled workers in the world. In addition, \hat{w} is decreasing in γ whereas g is independent of γ .

Through the determination of the wage rate, the necessary condition for optimality is thus satisfied. However, the first-order condition is not automatically sufficient, as the objective function is not necessarily concave. In fact, the concavity of the objective function is determined by the parameter γ . We resort to numerical solutions for $v(z)$ and $E(z)$ which show that below a threshold level of γ , the second-order conditions for optimality are not satisfied at $z = 1$ and an SBGE does not exist. The following sections contain this analysis. In Section 5.3, we then consider stability analysis and transitional dynamics.

5.1 A numerical example

To illustrate the characteristics of the balanced growth equilibrium, we provide a numerical example. The parameter values have been set as follows. The parameter b , which governs the upper bound of the TFP production function function, is set to 4.5. κ , which governs the concavity of the function, is set to 0.4. The consumers' discount factor, β , equals 0.9. The total number of high-skilled workers in the world, e_W , is set to 0.1. Finally, the parameter governing the weight of frontier countries' TFP in average TFP, ψ , is set to 1 which implies that \bar{T} is an arithmetic average of all T_i ; we briefly discuss the effects of varying ψ in Section 6. Note that at this stage, the parameter values are not chosen to match real-world data.

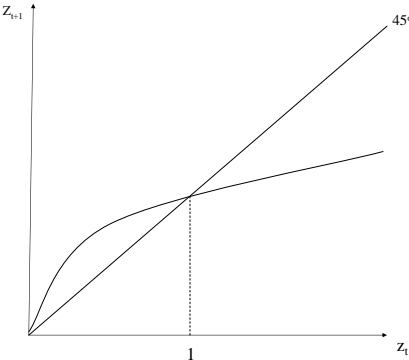
5.2 Existence in an SBGE: a country's policy function

Using the numerical example described above, we calculate an individual country planner's optimal policy function $z_{t+1} = f(z_t)$. Figures 1-4 below show the properties of the policy

function for different values of γ , under the assumption that all other countries are in a symmetric balanced growth equilibrium.² The stationary points depicted are those for which sufficient conditions for optimality are satisfied.

Figure 1 depicts the case when γ is high, i.e., when technological catch-up is important for TFP growth. The policy function has a unique stationary point which is the symmetric equilibrium $z = 1$.

Figure 1: Example of a country’s policy function



As γ decreases, the policy function starts to bend downward to the left of $z = 1$ and upward to the right of $z = 1$. This case is shown in Figure 2.

Figure 2: Example of a country’s policy function

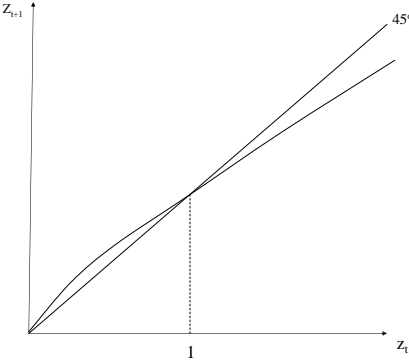
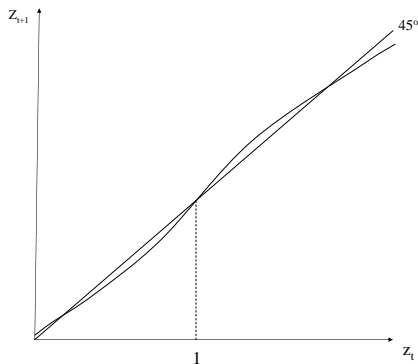


Figure 3 illustrates that as γ decreases further, two new stationary points emerge, one

²We obtained the policy functions through value function iteration. Figures 1-4 are stylized descriptions of the properties of the functions.

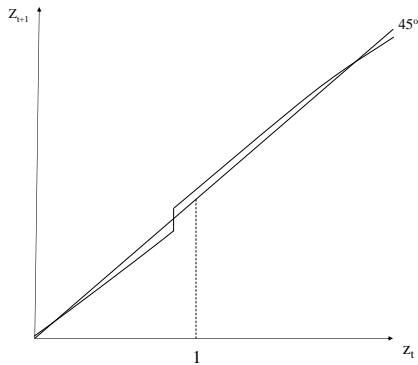
on each side of $z = 1$. The stationary point $z = 1$ itself now becomes unstable, a result that will be discussed in the stability analysis below.

Figure 3: Example of a country's policy function



Finally, as depicted in Figure 4, for sufficiently low values of γ , the symmetric stationary point ceases to exist, while the two asymmetric points remain, albeit further apart.

Figure 4: Example of a country's policy function



A more detailed account of how existence of the SBGE depends on γ in our numerical example is given in Table 1 below.

5.3 Stability and transitional dynamics

The previous section showed that for high values of γ , an SBGE exists. The one-group model does not exhibit any transitional dynamics around this SBGE. That is, if the initial distribution of relative TFP levels, $z_{i,0}$, is identical for all countries i , then all countries will

choose the same investment in TFP in the initial period, e_W , and the resulting growth rate will be identical to that of the SBGE, g , from the beginning of time. If all countries start in a symmetric equilibrium, the consumption levels of all countries are identical in all periods. If a country starts with an initial level z_0 higher (lower) than 1, it will have a higher (lower) level of consumption in every period than a country with $z_0 = 1$.³ Next, we want to ascertain whether the symmetric equilibrium is stable.

There are several kinds of perturbations of countries, or groups of countries, with respect to which the equilibrium could be stable (or unstable). First, one can determine the stability with respect to a perturbation of one single country. Thus, one country is given a relative TFP level z that is slightly different from $z = 1$, while the rest are at $z = 1$. What path will that country then follow? If, and only if, it converges back to $z = 1$, then the SBGE is stable with respect to perturbations of a single country. In addition, one can determine stability with respect to perturbations of groups of countries, where the groups could be of any size. Now, suppose countries are divided into n groups, of arbitrary size; within a group, all countries have the same level of z . Suppose that all groups are given initial relative TFP levels z that are slightly different from $z = 1$. If all groups converge back to $z = 1$, then the SBGE is stable with respect to perturbations of groups of countries. In this analysis, we will determine the stability of the equilibrium with respect to two kinds of perturbations; to a perturbation of one country, which we denote *measure-zero stability*, and to a 2-group perturbation, denoted *2-group stability*.

In order to perform the stability analysis, we define a 2-group recursive world equilibrium in which transitions are possible, i.e., where the equilibrium does not have to exhibit balanced growth. This definition will also be the point of departure for a characterization of asymmetric balanced growth equilibria, which are discussed the next section.

Let the two groups have relative TFP levels z_1 and z_2 . Within a group, all countries are identical. As in the numerical example, ψ is set to 1, which implies that \bar{T}_t is an arithmetic average of all $T_{i,t}$. Let φ be the share of countries belonging to group 1, which we denote the low-TFP group. The group with relative TFP level z_2 is denoted the high-TFP group. The sum of relative TFP levels must equal 1, which implies that $z_2 = \frac{1-\varphi z_1}{1-\varphi}$. Consequently, the programming problem can be defined using only one aggregate state variable; z_1 . A recursive 2-group world equilibrium can be defined as follows.

A recursive **2-group world equilibrium** consists of $v(z, z_1)$, $E(z, z_1)$, $w(z_1)$, $f(z_1)$, and $g(z_1)$ such that (subscripts i omitted for convenience)

1. $\forall(z, z_1)$, $v(z, z_1)$ solves

$$v(z, z_1) = \max_e z - w(z_1)e + \beta v \left(z^{1-\gamma} \frac{H(e)}{f(z_1)}, g(z_1) \right);$$

³If a nontrivial distribution of initial asset positions is allowed, that will also influence the level of consumption.

2. $\forall(z, z_1), E(z, z_1)$ is

$$E(z, z_1) = \arg \max_e v(z, z_1);$$

3. $\forall z_1, g(z_1) = z_1^{1-\gamma} \frac{H(E(z_1, z_1))}{f(z_1)}$;

4. $\forall z_1, f(z_1) = \varphi z_1^{1-\gamma} H(E(z_1, z_1)) + (1 - \varphi) \left(\frac{1-\varphi z_1}{1-\varphi} \right)^{1-\gamma} H \left(E \left(\frac{1-\varphi z_1}{1-\varphi}, z_1 \right) \right)$; and

5. $\forall z_1,$

$$e_W = \varphi E(z_1, z_1) + (1 - \varphi) E \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right). \quad (12)$$

The first and second conditions give the value function and the policy function, respectively, for a country which faces an aggregate relative TFP level z_1 and chooses its individual relative TFP level z . The third condition states the law of motion for the aggregate state variable z_1 . The function $f(z_1)$ determines the gross aggregate growth rate $1 + g_t$. The last condition is the market-clearing condition for high-skilled workers.

5.3.1 Measure-zero stability

To determine whether the symmetric equilibrium is stable with respect to measure-zero perturbations, we analyze the behavior of an individual country whose relative TFP level differs slightly from the symmetric steady state value, $z = 1$. Using the notation introduced in the definition of the 2-group world equilibrium above, (10) can be restated as

$$z' = z^{1-\gamma} \frac{H(E(z, z_1))}{f(z_1)}. \quad (13)$$

Measure-zero stability can then be established based on the derivative of (13):

$$\frac{\partial z'}{\partial z} = (1 - \gamma) z^{-\gamma} \frac{H(E(z, z_1))}{f(z_1)} + z^{1-\gamma} \frac{H'(E(z, z_1)) E_1(z, z_1)}{f(z_1)}, \quad (14)$$

where $E_1(z, z_1)$ is the derivative of the policy function $E(z, z_1)$ with respect to its first argument. It is possible to deduce $E_1(z, z_1)$ from the recursive version of the Euler equation, (11). By taking the derivative with respect to z , one obtains a second-order equation in $E_1(z, z_1)$. The two solutions for $E_1(z, z_1)$ result in a pair of expressions for $\frac{\partial z'}{\partial z}$. Saddle-path stability corresponds to one expression larger than 1 in absolute value and one expression less than 1 in absolute value. Evaluating (14) at $z_1 = z = 1$, $f(z_1) = H(e_W)$, and $E(z, z_1) = e_W$ yields

$$\frac{\partial z'}{\partial z} = 1 - \gamma + \frac{H'(e_W)}{H(e_W)} E_1(1, 1). \quad (15)$$

Here, we see that the stability properties of the SBGE will depend heavily on γ .

Equivalently, linearization of (11) around $z = 1$ yields the characteristic equation

$$x^2 + Kx + \frac{1}{\beta} = 0 \quad (16)$$

where

$$G(x) = \frac{x}{H'(H^{-1}(x(1+g)))}$$

and

$$K = \frac{(1 - (1 - \gamma)\beta)G(1)}{(1 - \gamma)\beta G'(1)} - \frac{1}{(1 - \gamma)\beta} - (1 - \gamma). \quad (17)$$

The roots to the equation in (16) determine the stability of the system with respect to measure-zero perturbations around the value $z = 1$. The product of the roots is equal to $1/\beta$, which implies that at least one of the roots is larger than 1. Again, as seen from (17) the values of the roots depend on γ . We show, using our numerical example, that for high values of γ , one root is less than 1 in absolute value and one root larger than 1. Hence, the SBGE is saddle-path stable. This is the case depicted in Figures 1 and 2. For intermediate values of γ , both roots are larger than 1 in absolute value and, consequently, the SBGE is unstable. Figure 3 corresponds to this case. Any country with a relative TFP level smaller than $z = 1$ will converge to the stationary point to the left of $z = 1$ and any country with a relative TFP level larger than $z = 1$ will converge to the stationary point to the right of $z = 1$. For low values of γ , the roots are complex, which contradicts optimality and there is no SBGE, as shown in Figure 4. Hence, for values of γ below some threshold, a country whose initial level of TFP is slightly different from $z = 1$ will not converge to the SBGE. The measure-zero stability of the numerically calculated examples is listed in Table 1 below.

The stability analysis established that for some values of γ , a single country perturbed away from the SBGE will not converge to it. The next step is to determine to what relative TFP level the country will converge. In terms of figures 1-4, these levels correspond to the stationary points other than $z = 1$ appearing in some cases. For this purpose, we examine the stationary version of (11)

$$z^{1-\gamma} = \frac{H'(e_w)}{H'(H^{-1}(z^\gamma(1+g)))}. \quad (18)$$

First, note that this equation is satisfied for $z = 1$, the symmetric BGE. Moreover, both the left-hand side and the right-hand side of the equation are increasing in z , indicating that the equation can have multiple solutions. In our numerical example, we found that when γ is low, there are at least two additional solutions to (18). This case is depicted in Figures 3 and 4.

Why do the asymmetric stationary points arise? The model exhibits two countervailing forces; the catch-up effect, which generates convergence, and the dynamic increasing returns,

which generates divergence. When technological catch-up is less important, i.e., γ is low, the divergence effect dominates. Countries with a higher relative TFP level will find it optimal to invest more in TFP than countries with a lower relative TFP level.

5.3.2 2-group stability

In this section, we examine whether the SBGE is stable with respect to 2-group perturbations. As in the definition above, all countries are divided into two groups, denoted by their relative TFP levels: z_1 and z_2 . The SBGE is stable if a group of countries, when given initial values of z slightly different from 1, converges back to the SBGE. Suppose that group 1 is given a TFP level slightly different from $z_1 = 1$. Whether the countries in group 1 will converge back to $z_1 = 1$ is determined by $g(z_1)$. The derivative of $g(z_1)$ is

$$g'(z_1) = (1 - \gamma) z_1^{-\gamma} \frac{H(E(z_1, z_1))}{f(z_1)} + z_1^{1-\gamma} \frac{H'(E(z_1, z_1)) (E_1(z_1, z_1) + E_2(z_1, z_1))}{f(z_1)} - z_1^{1-\gamma} H(E(z_1, z_1)) \frac{f'(z_1)}{f(z_1)^2} \quad (19)$$

which, evaluated at $z_1 = 1$, $f(z_1) = H(e_W)$, and $E(z_1, z_1) = e_W$ equals

$$g'(1) = 1 - \gamma + \frac{H'(e_W)}{H(e_W)} (E_1(1, 1) + E_2(1, 1)) - \frac{f'(1)}{H(e_W)}. \quad (20)$$

In Appendix 1, we show that both $f'(1)$ and $E_2(1, 1)$ are equal to zero. Hence, (20) is identical to (15) and, consequently, the 2-group stability analysis yields conclusions identical to those of the measure-zero stability analysis; for values of γ below a certain threshold, a group of countries whose initial levels of TFP are slightly different from $z = 1$ will not converge to the SBGE. The 2-group stability of the numerically calculated examples is shown in Table 1.

Table 1: SBGE for an example economy						
γ	<i>existence</i>	\hat{w}	g	e	<i>measure-zero stability</i>	<i>2-group stability</i>
0.226	yes	6.18	1.34	0.1	stable	stable
0.224	yes	6.22	1.34	0.1	stable	stable
0.222	yes	6.25	1.34	0.1	unstable	unstable
0.220	yes	6.29	1.34	0.1	unstable	unstable
0.218	yes	6.33	1.34	0.1	unstable	unstable
0.216	yes	6.37	1.34	0.1	unstable	unstable
0.214	no	—	—	—	—	—

Explanatory notes:

\hat{w} : TFP-adjusted wage rate

g : TFP growth rate

e : amount of high-skilled workers employed

measure-zero stability: stability with respect to single-country perturbations

2-group stability: stability with respect to two-group perturbations

The table shows computed SBGE for values of γ ranging from 0.216 to 0.226. The growth rate g is equal to 1.34 for all symmetric BGE since it is independent of γ . This rate is very high, but is a result of the choice of parameter values for the function $H(e)$. The TFP-adjusted wage rate \hat{w} is decreasing in γ . Symmetry across countries implies that the amount of high-skilled labor employed in TFP accumulation is 0.1, which is equal to the average amount of high-skilled workers in the world. The two last columns in Table 1 indicate whether the equilibrium is stable or not. For $\gamma = 0.224$ or higher, the equilibrium is stable. However, as γ decreases, it becomes unstable, and finally, for $\gamma = 0.214$ or lower, the equilibrium ceases to exist.

6 Asymmetric balanced growth equilibria

An asymmetric balanced growth equilibrium, an ABGE, is one in which all variables grow at constant rates and there is more than one level of relative TFP chosen by the country planners. We will focus on a particular type of ABGE, namely 2-group BGE. These are “natural” outcomes (as opposed to outcomes with more than two groups), as will be argued below.

6.1 2-group balanced growth equilibria

The 2-group BGE is a specific case of the 2-group world equilibrium defined in Section 5, in which the growth rate is constant: $T_{i,t} = T_i(1+g)^t \forall i, t$ for $g > 0$. Since there are no transitional dynamics by definition in a BGE, we can omit the aggregate state variable z_1 from the optimization problem.

A recursive **2-group balanced growth equilibrium** consists of $v(z)$, $E(z)$, $w(z_1)$, $f(z_1)$, and $g(z_1)$ such that

1. $\forall z$, $v(z)$ solves

$$v(z) = \max_e z - w(z_1)e + \beta v \left(z^{1-\gamma} \frac{H(e)}{f(z_1)} \right);$$

2. $\forall z$, $E(z)$ is

$$E(z) = \arg \max_e v(z);$$

3. $\forall z_1$, $g(z_1) = z_1$ and $g(z_2) = z_2$, where $g(z_1) = z_1^{1-\gamma} \frac{H(E(z_1))}{f(z_1)}$;

4. $\forall z_1$, $f(z_1) = \varphi z_1^{1-\gamma} H(E(z_1)) + (1-\varphi) \left(\frac{1-\varphi z_1}{1-\varphi} \right)^{1-\gamma} H \left(E \left(\frac{1-\varphi z_1}{1-\varphi} \right) \right)$;

and

5. $\forall z_1,$

$$e_W = \varphi E(z_1) + (1 - \varphi) E\left(\frac{1 - \varphi z_1}{1 - \varphi}\right). \quad (21)$$

The first and second conditions give the value function and the policy function, respectively. The third condition ensures that the 2-group distribution is stationary. The function $f(z_1)$ determines the constant gross aggregate growth rate $1 + g$. In the ABGE, the world growth rate is determined by the division of high-skilled labor into the low- and high-TFP groups, and the relative size of the two groups. In addition, it depends on γ , whereas the growth rate in the SBGE is independent of γ . The last condition is the market-clearing condition for high-skilled workers.

Let e_1 and e_2 be the amount of high-skilled labor employed in the low-TFP group and the high-TFP group, respectively. The unknown parameters e_1 , e_2 , and φ are determined by combining the Euler equations in steady state, (18), for both groups

$$H(e_1)^{\frac{1-\gamma}{\gamma}} H'(e_1) = H(e_2)^{\frac{1-\gamma}{\gamma}} H'(e_2) \quad (22)$$

with the market-clearing condition for high-skilled labor (condition 5 in (21)). This system of two equations is underdetermined; it has one more unknown than equations. Consequently, if one solution exists, there are infinitely many solutions, indeed a whole continuum. Each solution has a corresponding distinct world growth rate, and TFP-adjusted wage rate for high-skilled labor.

Equation (22) allows us to identify a key necessary condition for a 2-group ABGE to exist: what is required is non-monotonicity of $H(e)^{\frac{1-\gamma}{\gamma}} H'(e)$. Whether or not there is non-monotonicity depends on the primitives γ and $H(e)$. Since H is increasing and strictly concave, there are opposing forces, as a straightforward derivative of this expression reveals. We do not attempt to provide general conditions on H and γ that satisfy the necessary non-monotonicity here; for the specific functional form for H we consider in this paper, however, non-monotonicity is satisfied in a range $(0, \bar{\gamma})$, with $\bar{\gamma} < 1$.

Also, note that for ABGEs with more than 2 groups, when there is non-monotonicity, for a generic e , there is an odd number of solutions to (22). Here, stability arguments (like those discussed in Section 5.3) make us focus on 2 (as opposed to 3) groups. We have not fully analyzed the case where (22) would admit more than 3 (say, 5) solutions, but such a possibility does at least not seem feasible with the functional form for H used here.

In an ABGE, the ratio of TFP between the low- and high-TFP group, $\frac{T_1}{T_2}$, is given by

$$\frac{T_1}{T_2} = \left(\frac{H(e_1)}{H(e_2)}\right)^{\frac{1}{\gamma}}$$

and the TFP-adjusted wage rate for high-skilled workers is given by

$$\hat{w} = \frac{\beta z_i^{1-\gamma} H'(e_i)}{(1+g)(1-\beta(1-\gamma))}, \quad (23)$$

where $i \in \{1, 2\}$.

The assumption of free movement of high-skilled labor across countries ensures that the wage rate paid to high-skilled workers is identical in the two groups. Therefore, the TFP-adjusted wage rate can be obtained from the Euler equation for either group. In the asymmetric equilibrium, the wage rate depends on the relative TFP levels as well as the division of high-skilled labor across the two groups.

As in the symmetric one, the asymmetric equilibrium has consumption growth that is identical for both groups and thus for all countries, but the level of consumption in a given country depends on its total income and on initial conditions. All countries in the low-TFP group will have a lower total income, and hence a lower level of consumption than countries in the high-TFP group. If all countries start in the asymmetric equilibrium, countries within the same group will have identical consumption levels.

6.2 Numerical example

The numerical example presented in Section 5.1 can be used to characterize the 2-group asymmetric equilibria as well as the symmetric ones. Table 2 below shows the computed symmetric and asymmetric balanced growth equilibria for different values of γ .

Table 2: Balanced growth equilibria for an example economy						
γ	<i>SBGE</i>		<i>Range of ABGE</i>			
	<i>stability</i>	<i>g</i>	<i>gap</i>	\hat{w}	<i>g</i>	φ
0.226	stable	1.34	—	—	—	—
0.220	unstable	1.34	0.190-0.195	6.289-6.161	1.342-1.342	0.151-0.598
0.214	—	—	0.093-0.097	6.400-6.047	1.346-1.345	0.325-0.732
0.208	—	—	0.052-0.055	6.510-5.868	1.353-1.350	0.428-0.800
0.202	—	—	0.031-0.033	6.614-5.632	1.363-1.356	0.502-0.844
0.196	—	—	0.019-0.021	6.709-5.351	1.375-1.363	0.560-0.876

Explanatory notes:

g: TFP growth rate

gap: ratio of TFP between low- and high-TFP group

\hat{w} : TFP-adjusted wage rate

φ : share of countries in low-TFP group

Table 2 displays, for each value of γ , the characteristics of the symmetric and asymmetric BGE. For the symmetric BGE, it shows whether the equilibrium is stable along with its growth rate g . For the asymmetric BGE, it shows the resulting values for the “gap”, i.e., the ratio of TFP between the low- and high-TFP group, as well as \hat{w} , g , and φ . The indeterminacy of the system of equations in the 2-group BGE implies that the numerical solutions entail

ranges of values for the parameters. The table shows that for values of γ of 0.226 or higher, only the symmetric BGE exists, and it is stable. For γ equal to 0.220, the symmetric BGE is unstable, and there exists a continuum of asymmetric BGE.⁴ As γ decreases to 0.214 or less, the symmetric BGE ceases to exist, while the asymmetric BGE remain. Within the group of ABGE, the table shows that the growth rate g decreases in γ .

From Table 2 we can conclude that if technological catch-up is important for TFP growth, i.e., γ is high, the distribution of TFP is symmetric. If, instead, technological catch-up is less important, i.e., γ is low, the distribution of TFP is asymmetric: twin-peaked. As mentioned above, the intuition for the rise of asymmetric BGE is that for sufficiently low values of γ , the dynamic increasing returns effect (which creates divergence) dominates the catch-up effect (which creates convergence). This implies that countries with a higher relative TFP level will find it optimal to invest more in TFP than countries with a lower relative TFP level. The former are technological leaders while the latter are technological laggards that benefit from technology spillovers from the leaders. If the countries have different initial TFP levels, then countries with lower initial relative TFP will invest less, such that they eventually reach z_1 , which constitutes the low-TFP group of the ABGE. Similarly, countries with higher initial relative TFP will invest more, such that they eventually reach z_2 , which constitutes the high-TFP group.⁵

Let us return to the parameterization of ψ , which governs the weight of frontier countries' TFP in the world average TFP level. In our numerical example, $\psi = 1$ and \bar{T} is an arithmetic average of all T_i . It is conceivable that frontier countries' TFP could have a larger weight in world average TFP, which would correspond to $\psi > 1$. How would that affect the equilibria of the model? In an SBGE, all countries have identical TFP levels, and the value of ψ has no effect on the average world TFP. In an ABGE, countries' relative TFP levels differ, and ψ does affect the world average TFP, \bar{T} , and its rate of growth, g . An increase in ψ increases g which, in turn, decreases \hat{w} as seen in (23). However, an individual country takes \bar{T} as given and therefore ψ does not affect the country's optimal choice of e , the amount of skilled workers to employ. The system of equations determining the 2-group ABGE, given by (22) and condition 5 in (21), is unaffected by ψ and so are the range of ABGE calculated in our numerical example. Since an increase in ψ affects g and \hat{w} it also affects the out-of-steady-state dynamics, but we have not analyzed this effect in any more detail.

⁴For a small range of values for γ , symmetric and asymmetric BGE coexist. Within that range, as γ decreases, the symmetric BGE goes from being stable to unstable.

⁵If all countries start at the same initial TFP level, our conjecture is that they will split up into two groups, one which starts to invest less, such that it eventually reaches z_1 , and one which starts to invest more, such that it eventually reaches z_2 . In the initial period, the sequence of wages and relative TFP levels, $\{w_t, z_t\}_{t=0}^{\infty}$, must be such that the countries are indifferent between joining the low- and the high-TFP groups, and that they choose to split up into groups of relative size which is consistent with the ABGE.

6.3 Stability properties of the ABGE

To ascertain whether the asymmetric balanced growth equilibria are stable, we perform the same type of stability analyses as for the symmetric balanced growth equilibria: measure-zero stability and 2-group stability.

6.3.1 Measure-zero stability

An ABGE characterized by the triplet e_1 , e_2 , and φ is stable with respect to measure-zero perturbations if a single country, which is given an initial relative TFP level slightly different from z_1 , converges back to z_1 , and a single country perturbed away from z_2 converges back to z_2 . Whether the country converges back or not is determined by $\frac{\partial z'}{\partial z}$, as given by (14), evaluated at $z = z_1$ and $z = z_2$, respectively. As in the case of the measure-zero stability analysis of the SBGE, the derivative of the recursive version of (11) with respect to z gives a second-order equation in $E_1(z, z_1)$ which results in a pair of expressions for $\frac{\partial z'}{\partial z}$. This pair is then evaluated at both z_1 and z_2 . We compute measure-zero stability for the ABGE in the numerical example characterized above, and Table 3 below displays the results.

6.3.2 2-group stability

An ABGE characterized by the triplet e_1 , e_2 , and φ is stable with respect to 2-group perturbations if the following holds: when the low-TFP group is given an initial relative TFP level slightly different from z_1 , it converges back to z_1 . (Since a perturbation of one group affects the remaining group, it is sufficient to analyze perturbations of one group only.) Suppose that the low-TFP group is perturbed away from z_1 . Whether it will converge back to z_1 is determined by the derivative of $g(z_1)$, as given by (19).

Unlike the in SBGE, the derivatives $f'(z_1)$ and $E_2(z_1, z_1)$ are not equal to zero, and must therefore be solved for in order to evaluate $g'(z_1)$. To that end, we compute the derivative with respect to z_1 of the recursive version of (11), and evaluate it at $z = z_1$ and $z = z_2$, respectively. Combining the resulting two equations with the expressions for $f'(z_1)$, $g'(z_1)$, and the derivative with respect to z_1 of condition 5 in (12), we obtain a system of 5 equations. There are five unknowns; $f'(z_1)$, $g'(z_1)$, $E_2(z_1, z_1)$, $E_2(z_2, z_1)$, and $w'(z_1)$. The system yields a second-order equation in $E_2(z_1, z_1)$ and therefore has two solutions. We solve the system of equations, and for each solution obtain an expression for $g'(z_1)$; see Appendix 1 for details. If one of the expressions is larger than 1 in absolute value, and one is less than 1 in absolute value, there is saddle-path stability. We compute the values of $g'(z_1)$ for the ABGE in the numerical example characterized above, and the results are reported in Table 3.

Table 3: Range of ABGE for an example economy

γ	gap	\hat{w}	g	φ	<i>measure-zero stability</i>	<i>2-group stability</i>
0.220	0.190-0.195	6.289-6.161	1.342-1.342	0.151-0.598	stable	stable
0.214	0.093-0.097	6.400-6.047	1.346-1.345	0.325-0.732	stable	stable
0.208	0.052-0.055	6.510-5.868	1.353-1.350	0.428-0.800	stable	stable
0.202	0.031-0.033	6.614-5.632	1.363-1.356	0.501-0.844	stable	stable
0.196	0.019-0.021	6.709-5.351	1.376-1.363	0.560-0.876	stable	stable

Explanatory notes:*gap*: ratio of TFP between low- and high-TFP group \hat{w} : TFP-adjusted wage rate g : TFP growth rate φ : share of countries in low-TFP group*measure-zero stability*: stability with respect to single-country perturbations*2-group stability*: stability with respect to 2-group perturbations

As Table 3 shows, the ABGE in the numerical example are stable, both with respect to measure-zero and to 2-group perturbations.

7 TFP shocks

In this section, the model is extended to allow for country-specific shocks to TFP. The motivation for this extension is twofold. First, it is to create a more smooth and realistic world TFP distribution where individual countries can move between groups and potentially experience both growth miracles and growth disasters. Second, it is an attempt to eliminate the indeterminacy of the asymmetric steady states obtained in the baseline model.

7.1 Model

The following assumptions are added to the model described in Section 3. Each country is subject to a TFP shock $\varepsilon_{i,t}$, $\varepsilon \in \{\varepsilon_L, \varepsilon_H\}$ where $\varepsilon_L < \varepsilon_H$. The probability of a good shock is denoted $P(\varepsilon_H) = \pi$. The shock ε is *iid* across countries and across time. The country planner sets $e_{i,t}$ before observing the shock $\varepsilon_{i,t}$. With TFP shocks, the accumulation of TFP has the following form

$$T_{i,t+1} = (1 + \varepsilon_{i,t}) T_{i,t}^{1-\gamma} \bar{T}_t^\gamma H(e_{i,t})$$

which is (4) with the addition of the shock $\varepsilon_{i,t}$.

It is also assumed that the world has perfect consumption insurance and frictionless borrowing and lending. This ensures that a separation of the optimization problem into an income-maximization problem and an intertemporal consumption allocation problem is still valid. The intertemporal consumption allocation problem is the same as the one in the model without technology shocks, and the resulting allocation will be the same, given total income. However, the income-maximization problem is different, as will be described below.

7.2 A country's investment decision

As in the model without shocks, we restrict our attention to balanced growth equilibria. In the income-maximization problem, the country planner chooses an investment in TFP so as to maximize output net of investment costs. The recursive formulation of the optimization problem is

$$v(z) = \max_e z - \hat{w}e + \beta(\pi v(z'_H) + (1 - \pi)v(z'_L))$$

subject to, for $j = L, H$

$$z'_j = z^{1-\gamma} \frac{1 + \varepsilon_j}{1 + g} H(e).$$

The country planner employs an amount e of high-skilled workers in a given time period. With probability π the country is hit by a positive shock, and the resulting next period relative TFP level is z'_H and the corresponding value function is $v(z'_H)$. With probability $1 - \pi$ the country is hit by a negative shock, and the resulting next period relative TFP level is z'_L , associated with the value function $v(z'_L)$.

7.3 Definition of a balanced growth equilibrium

A balanced growth equilibrium is a world equilibrium, where all variables grow at rate g on average. Individual countries can have a TFP growth that is faster or slower than this average rate g . The distribution of z is constant. As in the model without shocks, let $\Gamma(z)$ be a probability measure on (S, β_s) where $S \in [z_{\min}, z_{\max}]$ and β_s is the associated Borel σ -algebra. ψ is set to 1, which implies that \bar{T} is an arithmetic average of all T_i .⁶

A **balanced growth equilibrium** consists of a stationary probability measure $\Gamma(z)$, variables \hat{w} and g , and functions $v(z)$ and $E(z)$ such that

1. $\forall z$, $v(z)$ solves

$$v(z) = \max_e z - \hat{w}e + \beta \left(\pi v \left(z^{1-\gamma} \frac{1 + \varepsilon_H}{1 + g} H(e) \right) + (1 - \pi) v \left(z^{1-\gamma} \frac{1 + \varepsilon_L}{1 + g} H(e) \right) \right);$$

2. $\forall z$, $E(z)$ is

$$E(z) = \arg \max_e v(z);$$

⁶Readers familiar with the contributions of Bewley (undated), Huggett (1993), and Aiyagari (1994) will recognize a similarity of the steady-state analyses of consumer inequality there with the balanced growth equilibrium with country inequality here. A technical difference between these is that there are two unknown long-run quantities here—the real wage and the endogenous growth rate—as opposed to only one—the real wage (or, alternatively, the real interest rate)—in those settings.

3. g solves

$$\pi \int_S z^{1-\gamma} \frac{1 + \varepsilon_H}{1 + g} H(E(z)) d\Gamma(z) + (1 - \pi) \int_S z^{1-\gamma} \frac{1 + \varepsilon_L}{1 + g} H(E(z)) d\Gamma(z) = 1;$$

4. \hat{w} solves $\int_S E(z) d\Gamma(z) = e_W$; and

5. $\Gamma(z)$ satisfies

$$\Gamma(B) = \pi \int_{z \in S: z^{1-\gamma} \frac{1 + \varepsilon_H}{1 + g} H(E(z)) \in B} d\Gamma(z) + (1 - \pi) \int_{z \in S: z^{1-\gamma} \frac{1 + \varepsilon_L}{1 + g} H(E(z)) \in B} d\Gamma(z) \quad \forall B \in \beta_S. \quad (24)$$

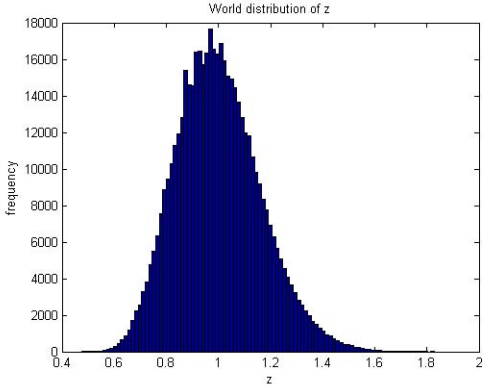
The first and second conditions give the value function and policy function, respectively. The third condition states that the growth rate g must be such that the integral over all relative TFP levels equals 1. The fourth condition states that the wage \hat{w} must be such that there is market clearing in the market for high-skilled workers. The last condition ensures that the probability measure $\Gamma(z)$ is stationary.

7.4 A numerical example

The model with TFP shocks is solved numerically using the parameterization in the example applied to the baseline model. There are three additional parameters values to be set; π , ε_L , and ε_H . π is set to 0.5, implying that the TFP shock is high and low with equal probability. The size of the shock is chosen such that the shock is symmetric; $\varepsilon_L = -0.05$ and $\varepsilon_H = 0.05$. Numerical solutions are then obtained for different values of γ ; the solution method is described in Appendix A2. Figures 5-7 below depict the world distribution of relative TFP levels z_i corresponding to $\gamma=0.24$, 0.225, and 0.22, respectively. From the results obtained in this numerical example, the equilibrium appears to be unique for a given value of γ .

Figure 5 shows the world distribution of z_i for $\gamma=0.24$. The distribution is single-peaked. As a result of the TFP shocks, there is significant dispersion around the center value $z = 1$, creating the smooth symmetric shape of the distribution.

Figure 5: $\gamma = 0.24$



When γ is decreased to 0.225, as displayed in Figure 6, the distribution is still single-peaked but the dispersion has increased, some countries lag behind, while others are at much higher relative TFP levels.

Figure 6: $\gamma = 0.225$

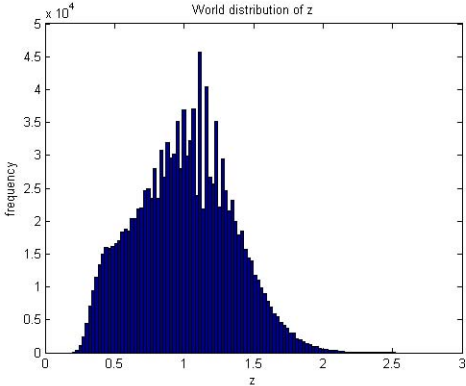
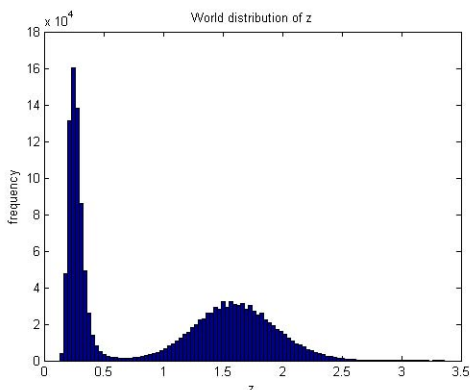


Figure 7 shows that for γ as low as 0.22, the distribution of TFP is asymmetric; it is twin-peaked. One group of countries has settled at a low TFP level, and another at a high TFP level. As countries are hit by shocks, it is possible for a given country to move from one of the groups to the other. However, the distribution of countries over TFP levels remains constant. We can conclude that for high values of γ , the distribution is single-peaked, while for sufficiently low values of γ the distribution of countries is twin-peaked.

Figure 7: $\gamma = 0.22$



Even though an individual country's TFP grows at rate g on average, its consumption grows at rate g for all t . The relative levels of consumption are determined by initial conditions. All countries which start at the same initial relative TFP level z_0 have the same consumption level. A country which starts at a higher (lower) initial level has a higher (lower) level of consumption in each time period.

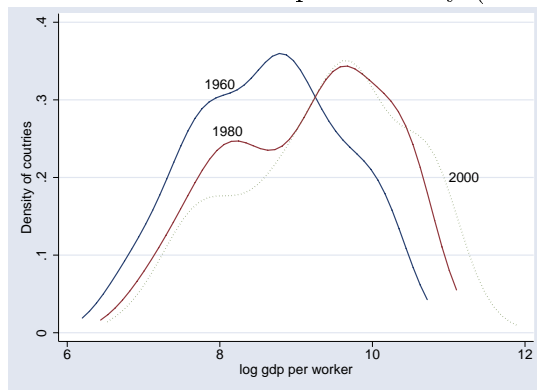
How do the results from the model with TFP shocks compare to the model without shocks? The numerical results show that with shocks, the balanced growth equilibrium appears to be unique for a given value of γ , whereas without shocks, the model produced a continuum of asymmetric equilibria. However, the main results from the baseline model remain; if technological catch-up is important for TFP growth, (γ is high), the distribution of TFP is symmetric. If, instead, technological catch-up is less important, (γ is low), the distribution of TFP is asymmetric: twin-peaked.

8 Changes in the distribution of world TFP

The distribution of world TFP is not constant over time. Three snapshots of the distribution of labor productivity, measured as log (PPP-adjusted) output per worker, from Acemoglu (2009) are depicted in Figure 8 below.⁷

⁷The distributions are based on kernel density estimates.

Figure 8: Distributions of world labor productivity (from Acemoglu, 2009)



We observed marked twin peaks in 1980 and more dispersion in 2000, though here the peak on the left is less clear, whereas the 1960 distribution shows only weak signs of bimodality. Whether there are two peaks or not, these distributions are different, and the general question we ask here is what fundamental determinants lie behind the shape of the distribution of TFP generated by our model. We will, in turn, and very briefly, discuss some of the determinants: changes in γ , changes in trading costs for e , changes in H , and transitional dynamics.⁸

8.1 The value of γ

First, we saw above that γ is a key determinant of whether bimodality is an outcome in our model; low values are required, *ceteris paribus*, for bimodality, and generally speaking a lower γ tends to make the distribution more dispersed. Thus, a move toward bimodality/more dispersion might be interpreted as a decrease in γ . What, then, is γ exactly? We have remained deliberately vague on this, and the Nelson-Phelps setting indeed is “ad hoc”: it is not derived from first (technological) principles. As specified, γ captures two features at once. First, it captures something that we might label a technology (either physical technology or information technology), i.e., how easily *transferable* TFP knowledge is from the world to a given economy. With this interpretation, a fall in γ means weaker transferability. Given the rise of IT technology and, more generally, globalization, this interpretation would suggest that γ has risen, as opposed to fallen. The other feature γ captures, however, is how much of the TFP investments are *internalized*, as opposed to how much they are treated as exogenous. Here, in contrast, at least to the extent that countries are able to better coordinate action across countries, and perhaps also better internalize TFP externalities that occur within a given country, one would argue that a fall in γ over this period is a reasonable assumption. A conclusion from this discussion is that one would, perhaps, want a model that goes beyond

⁸Other primitives, such as β and e_W or the nature of the shocks to TFP, can also influence TFP distributions, but we include no analysis of them here.

the simple Nelson-Phelps formulation by making the distinction between transferability (a technology/information concern) and the degree of internalization of the dynamics of TFP (a concern about how well countries are able to do this locally and how well they are able to coordinate between them).

8.2 Trading costs

Globalization was argued above to at least potentially be a determinant of γ . However, since γ also captures other important model features, a cleaner experiment is to consider a decrease in the costs of trading. Our theory so far assumes that e is an input that is traded in world markets at a competitively determined price. Consider instead the polar opposite case, namely that where e cannot be traded at all. Thus, with symmetric endowments, all countries would invest the same amount, e_W , in TFP accumulation. The TFP accumulation equation thus reads

$$T_{i,t+1} = T_{i,t}^{1-\gamma} \bar{T}_t^\gamma H(e_W),$$

which can be rewritten as

$$\frac{T_{i,t+1}}{\bar{T}_{t+1}} = \left(\frac{T_{i,t}}{\bar{T}_t} \right)^{1-\gamma} \frac{H(e_W) \bar{T}_t}{\bar{T}_{t+1}} = \left(\frac{T_{i,t}}{\bar{T}_t} \right)^{1-\gamma},$$

since \bar{T} grows at rate $H(e_W)$. Thus, relative TFPs converge to one: there is global convergence. More generally, based on the comparison of the extreme cases, with trade costs in the fundamental input into the accumulation of TFP, we would thus expect to see a force toward convergence. A fall in trade costs would then produce a force away from convergence and toward the bimodal distribution of world TFP.

8.3 Other model parameters and transitional dynamics

We saw in Section 6.1 that the function H , which translates TFP investment input, e , into TFP growth, is an important determinant behind the distribution of world TFP; recall, e.g., that a necessary condition for bimodality in an ABGE based on interior solutions is that $H(e)^{\frac{\gamma}{1-\gamma}} H'(e)$ be non-monotonic. First, notice that if H is flat, i.e., $H(e)$ does not depend on e , asymmetric outcomes are never possible. Thus, if “R&D has become more productive”, perhaps due to basic exogenous technical change, we would see a force toward bimodality. The shape of H would matter too, and using our functional form for H , it appears that higher curvature, i.e., more rapidly decreasing returns, also is a force toward bimodality.

Finally, transitional dynamics can of course explain changes in the TFP distribution. That is, suppose that there are no changes in primitives but that these primitives imply a bimodal long-run outcome, and suppose that the initial TFP distribution is a unimodal one with less inequality in TFP than observed currently. Then we would see something like the observed changes play out over time. However, this requires additional understanding of the origins of the initial distribution. These origins, in turn, are likely due to technology, or

information, or “country organization” being different in the present than the past. Again take the trade costs as an example: if no trade is possible, we would see an outcome of equal TFP in all countries, and if trade were made possible, there would then be gradual movement from an equal to an unequal TFP distribution. A similar history would likely be generated if TFP shocks were very minor in the past but have become significant now, or if H was flat in the past (R&D was unproductive) but is upward-sloping in modern economies.⁹

9 Concluding comments

This paper tries to answer the question: if there is technological catch-up, as proposed by Nelson and Phelps (1966), and each country takes this into account while maximizing the utility of its citizens, what is the resulting equilibrium distribution of TFP? To this end, the paper presents a general equilibrium model in which individual countries invest in a technology-enhancing input that is traded in world markets, and the accumulation of TFP is modeled according to the Nelson-Phelps specification. Even though all countries are treated symmetrically, the model can generate a nontrivial long-run world distribution of TFP. The model predicts that if technological catch-up is important for TFP growth, the distribution of countries over relative TFP is symmetric. If, instead, technological catch-up is less important for TFP growth, the distribution is asymmetric: twin-peaked. There is one group of countries with high TFP in relative terms, with the remainder of the countries operating at a much lower relative TFP level. All countries grow at the same rate, but the high-TFP countries invest more in technology than do low-TFP countries. The catch-up term thus allows the low-TFP countries to grow at the same rate and not fall further behind in relative terms. More generally, independently of whether bimodality is an important feature of the data, the present model does produce equilibrium world distributions of TFP that display more or less dispersion depending on the fundamental parameters of the model, and the same parameters that tend to generate bimodality tend also to produce more dispersion.

The analysis we present is only a first attempt at a theory of the world distribution of country TFP. A serious quantitative analysis would require moving toward a model with (i) capital accumulation; (ii) some extent of incomplete insurance against TFP shocks; and, last but not least, (iii) geographical detail, so that large countries are different than small countries (and, say, there is heterogeneity in γ s). Such extensions would be very interesting to pursue in future work.

⁹Looking at a recent period, and based on a model similar in spirit to that in Greenwood and Yorukoglu (1997), Beaudry, Collard and Green (2005) conclude that demographic change accounts for a large part of the transitional changes in the distribution of world income. This is likely to be an important factor also in an extended version of the present model.

Bibliography

- Acemoglu, D. (2009), *Introduction to Modern Economic Growth*, Princeton: Princeton University Press.
- Acemoglu, D. and Ventura, J. (2002), “The World Income Distribution”, *The Quarterly Journal of Economics*, Vol. 117, No. 2, pp. 659–694.
- Aghion, P. Howitt, P. and Mayer-Foulkes, D. (2005), “The Effect of Financial Development on Convergence: Theory and Evidence”, *The Quarterly Journal of Economics*, Vol. 120, No. 1, pp. 173–222.
- Aiyagari, R. (1994), “Uninsured Idiosyncratic Risk and Aggregate Saving”, *The Quarterly Journal of Economics*, Vol. 109, No. 3, pp. 659–684.
- Beaudry, P. Collard, F. and Green, D. (2005), “Demographics and recent productivity performance: insights from cross-country comparisons”, *Canadian Journal of Economics*, Vol. 38, No. 2, pp. 309–344, May.
- Benhabib, J. and Spiegel, M. (1994), “The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data,” *Journal of Monetary Economics*, Vol. 34, pp. 143–173.
- Bewley, T. (undated), “Interest Bearing Money and the Equilibrium Stock of Capital, mimeo.
- Chari, V. Kehoe, P. and McGrattan, E. (1997), “The Poverty of Nations: A Quantitative Investigation”, Federal Reserve Bank of Minneapolis Research Department Staff Report 204/JV.
- Greenwood, J. and Yorukoglu, M. (1997), “1974” *CarnegieRochester Conference Series on Public Policy* Vol. 46, pp. 49-95.
- Hall, R. and Jones, C. (1999), “Why Do Some Countries Produce So Much More Output Per Worker Than Others?”, *The Quarterly Journal of Economics*, Vol. 114, No. 1, pp. 83–116.
- Heston, A., Summers R. and Aten B. (2006), Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.
- Howitt, P. (2000), “Endogenous Growth and Cross-Country Income Differences”, *The American Economic Review*, Vol. 90, No. 4, pp. 829–846.
- Howitt, P. and Mayer-Foulkes, D. (2005), “R&D, Implementation, and Stagnation: A Schumpeterian Theory of Convergence Clubs”, *Journal of Money, Credit, and Banking*, Vol. 37, No. 1, pp. 147–177.
- Huggett, M. (1993), “The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies,” *Journal of Economic Dynamics and Control*, Vol. 17, pp. 953–969.
- Jones, C. (2008), *Macroeconomics*, New York: W.W. Norton.

- Jones, C. (1998), *Introduction to Economic Growth*, New York: W.W. Norton.
- Klenow, P. and Rodríguez-Clare, A. (1997), “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” *NBER Macroeconomics Annual*, B. Bernanke and J. Rotemberg, eds. Cambridge: MIT Press.
- Kremer, M. Onatski, A. and Stock, J. (2001), “Searching for Prosperity”, NBER Working Paper no 8250.
- Lucas, R. (1988), “On the Mechanics of Economic Development”, *Journal of Monetary Economics*, Vol. 22, No. 1, pp. 3–42.
- Lucas, R. (2000), “Some Macroeconomics for the 21st Century”, *Journal of Economic Perspectives*, Vol. 14, No. 1, pp. 159–168.
- Nelson, R. and Phelps, E. (1966), “Investment in Humans, Technological Diffusion, and Economic Growth”, *The American Economic Review*, Vol. 56, No. 1/2, pp. 69–75.
- Parente, S. and Prescott, E. (1994), “Barriers to Technology Adoption and Development”, *The Journal of Political Economy*, Vol. 102, No. 2, pp. 298–321.
- Prescott, E. (1998), “Needed: A Theory of Total Factor Productivity”, *International Economic Review*, Vol. 39, No. 3, pp. 525–551.
- Quah, D. (1993), “Empirical cross-section dynamics in economic growth”, *European Economic Review*, Vol. 37, No. 2/3, pp. 426–434.
- Quah, D. (1997), “Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs”, *Journal of Economic Growth*, Vol. 2, No. 1, pp. 27–59.
- Rebelo, S. (1991), “Long-Run Policy Analysis and Long-Run Growth”, *The Journal of Political Economy*, Vol. 99, No. 3, pp. 500–521.
- Romer, P. (1986), “Increasing Returns and Long-Run Growth”, *The Journal of Political Economy*, Vol. 94, No. 5, pp. 1002–1037.

Appendix A1

2-group stability of the SBGE

2-group stability is determined by

$$g'(z_1) = (1 - \gamma) z_1^{-\gamma} \frac{H(E(z_1, z_1))}{f(z_1)} + z_1^{1-\gamma} \frac{H'(E(z_1, z_1))(E_1(z_1, z_1) + E_2(z_1, z_1))}{f(z_1)} - z_1^{1-\gamma} H(E(z_1, z_1)) \frac{f'(z_1)}{f(z_1)^2}, \quad (25)$$

which evaluated at $z_1 = 1$, $f(z_1) = H(e_W)$, and $E(z_1, z_1) = e_W$ equals

$$g'(1) = 1 - \gamma + \frac{H'(e_W)}{H(e_W)} (E_1(1, 1) + E_2(1, 1)) - \frac{f'(1)}{H(e_W)}.$$

The derivative $E_2(1, 1)$ can be obtained from the market-clearing condition for high-skilled labor:

$$e_W = \varphi E(z_1, z_1) + (1 - \varphi) E\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right).$$

The derivative of this expression with respect to z_1 is

$$\frac{\partial e_W}{\partial z_1} = \varphi (E_1(z_1, z_1) + E_2(z_1, z_1)) + (1 - \varphi) \left(E_2\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right) + E_1\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right) \frac{-\varphi}{1 - \varphi} \right). \quad (26)$$

Evaluated at $z_1 = z_2 = 1$ and set to equal zero, this equation gives

$$\varphi (E_1(1, 1) + E_2(1, 1)) + (1 - \varphi) \left(E_2(1, 1) + E_1(1, 1) \frac{-\varphi}{1 - \varphi} \right) = 0,$$

or equivalently

$$\varphi E_1(1, 1) + \varphi E_2(1, 1) + (1 - \varphi) E_2(1, 1) - \varphi E_1(1, 1) = 0,$$

which implies that $E_2(1, 1) = 0$.

The derivative of $f(z_1)$ is

$$\begin{aligned} f'(z_1) &= \varphi (1 - \gamma) z_1^{-\gamma} H(E(z_1, z_1)) + \varphi z_1^{1-\gamma} H'(E(z_1, z_1))(E_1(z_1, z_1) + E_2(z_1, z_1)) \\ &\quad - \varphi (1 - \gamma) \left(\frac{1 - \varphi z_1}{1 - \varphi} \right)^{-\gamma} H\left(E\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right)\right) \\ &\quad + (1 - \varphi) H'\left(E\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right)\right) E_2\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right) \left(\frac{1 - \varphi z_1}{1 - \varphi}\right)^{1-\gamma} \\ &\quad + (1 - \varphi) H'\left(E\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right)\right) E_1\left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1\right) \left(\frac{-\varphi}{1 - \varphi}\right) \left(\frac{1 - \varphi z_1}{1 - \varphi}\right)^{1-\gamma}. \end{aligned}$$

Evaluated at $z_1 = z_2 = 1$ this expression becomes

$$\begin{aligned} f'(1) &= \varphi H'(e_W)(E_1(1, 1) + E_2(1, 1)) \\ &\quad + (1 - \varphi)H'(e_W) \left(E_2(1, 1) + E_1(1, 1) \left(\frac{-\varphi}{1 - \varphi} \right) \right). \end{aligned}$$

Using the previous result that $E_2(1, 1) = 0$, the expression reads

$$\begin{aligned} f'(1) &= \varphi H'(e_W)E_1(1, 1) + (1 - \varphi)H'(e_W) \left(E_1(1, 1) \left(\frac{-\varphi}{1 - \varphi} \right) \right) \\ &= 0. \end{aligned}$$

Inserting $E_2(1, 1) = 0$ and $f'(1) = 0$ into $g'(1)$ gives

$$g'(1) = 1 - \gamma + \frac{H'(e_W)}{H(e_W)}E_1(1, 1),$$

which is equivalent to

$$\frac{\partial z'}{\partial z} = 1 - \gamma + \frac{H'(e_W)}{H(e_W)}E_1(1, 1).$$

Hence, in the SBGE, 2-group stability is equivalent to measure-zero stability.

We obtain $E_1(1, 1)$ as follows. Let the Euler equation in recursive notation be denoted F , where

$$F = \frac{\beta}{\hat{w}} z^{1-\gamma} \frac{H(E(z, z_1))}{f(z_1)} - \frac{H(E(z, z_1))}{H'(E(z, z_1))} + (1 - \gamma)\beta \frac{H \left(E \left(z^{1-\gamma} \frac{H(E(z, z_1))}{f(z_1)}, g(z_1) \right) \right)}{H' \left(E \left(z^{1-\gamma} \frac{H(E(z, z_1))}{f(z_1)}, g(z_1) \right) \right)}. \quad (27)$$

Taking the first-order condition gives

$$(E_1(z, z_1))^2 + pE_1(z, z_1) + q = 0,$$

where

$$\begin{aligned} p &= \frac{\left(\frac{\beta}{\hat{w}} \frac{z^{1-\gamma}}{f(z_1)} H'(E(z, z_1)) - 1 + \frac{H''(E(z, z_1))H(E(z, z_1))}{(H'(E(z, z_1)))^2} + (1 - \gamma)^2 \beta z^{-\gamma} \frac{H(E(z, z_1))}{f(z_1)} \right)}{(1 - \gamma)\beta \frac{z^{1-\gamma}}{f(z_1)} \left(H'(E(z, z_1)) - \frac{H'(E(z, z_1))H(E(z, z_1))H''(E(z, z_1))}{(H'(E(z, z_1)))^2} \right)} \\ &\quad - \frac{(1 - \gamma)^2 \beta z^{-\gamma} \frac{(H(E(z, z_1)))^2 H''(E(z, z_1))}{f(z_1)(H'(E(z, z_1)))^2}}{(1 - \gamma)\beta \frac{z^{1-\gamma}}{f(z_1)} \left(H'(E(z, z_1)) - \frac{H'(E(z, z_1))H(E(z, z_1))H''(E(z, z_1))}{(H'(E(z, z_1)))^2} \right)} \end{aligned}$$

and

$$q = \frac{H(E(z, z_1))}{\hat{w}z \left(H'(E(z, z_1)) - \frac{H'(E(z, z_1))H(E(z, z_1))H''(E(z, z_1))}{(H'(E(z, z_1)))^2} \right)}.$$

Hence

$$E_1(z, z_1) = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}, \quad (28)$$

where p and q are given above. This expression evaluated at $z = z_1 = 1$ gives $E_1(1, 1)$.

2-group stability of the ABGE

As in the case of the SBGE, 2-group stability is given by (25). In order to evaluate $g'(z_1)$ in a given ABGE, the derivatives $E_1(z_1, z_1)$, $E_2(z_1, z_1)$, and $f'(z_1)$ must be computed. We solve for $E_1(z_1, z_1)$ and $E_2(z_1, z_1)$ by evaluating (28) at $z = 1$ and $z = 2$, respectively. The expression for $E_1(z_1, z_1)$ corresponding to the stable root is inserted into $g'(z_1)$. Inspection of (26) reveals that when evaluated at $z = z_1$ it does not imply that $E_2(z_1, z_1) = 0$. Consequently, $f'(z_1)$ is not zero, and both $E_2(z_1, z_1)$ and $f'(z_1)$ must be computed. In order to solve for $g'(z_1)$, we construct the following system of equations. The first two equations are the derivative with respect to z_1 of the recursive version of the Euler equation, (27), evaluated at $z = z_1$ and $z = z_2$, respectively. The third equation states that $\frac{\partial e_W}{\partial z_1} = 0$. Combining these equations with $f'(z_1)$ and $g'(z_1)$ yields the following system of five equations:

1. $\left. \frac{\partial F}{\partial z_1} \right|_{z=z_1} = 0;$

2. $\left. \frac{\partial F}{\partial z_1} \right|_{z=z_2} = 0;$

- 3.

$$\varphi (E_1(z_1, z_1) + E_2(z_1, z_1)) + (1 - \varphi) \left(E_2 \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) + E_1 \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \frac{-\varphi}{1 - \varphi} \right) = 0;$$

- 4.

$$\begin{aligned} g'(z_1) &= (1 - \gamma) z_1^{-\gamma} \frac{H(E(z_1, z_1))}{f(z_1)} + z_1^{1-\gamma} \frac{H'(E(z_1, z_1)) (E_1(z_1, z_1) + E_2(z_1, z_1))}{f(z_1)} \\ &\quad - z_1^{1-\gamma} H(E(z_1, z_1)) \frac{f'(z_1)}{f(z_1)^2}; \end{aligned}$$

and

- 5.

$$\begin{aligned} f'(z_1) &= \varphi (1 - \gamma) z_1^{-\gamma} H(E(z_1, z_1)) + \varphi z_1^{1-\gamma} H'(E(z_1, z_1)) (E_1(z_1, z_1) + E_2(z_1, z_1)) \\ &\quad - \varphi (1 - \gamma) \left(\frac{1 - \varphi z_1}{1 - \varphi} \right)^{-\gamma} H \left(E \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \right) \\ &\quad + (1 - \varphi) H' \left(E \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \right) E_2 \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \left(\frac{1 - \varphi z_1}{1 - \varphi} \right)^{1-\gamma} \\ &\quad + (1 - \varphi) H' \left(E \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \right) E_1 \left(\frac{1 - \varphi z_1}{1 - \varphi}, z_1 \right) \left(\frac{-\varphi}{1 - \varphi} \right) \left(\frac{1 - \varphi z_1}{1 - \varphi} \right)^{1-\gamma}. \end{aligned}$$

The system has five unknown variables: $f'(z_1)$, $g'(z_1)$, $E_2(z_1, z_1)$, $E_2(z_2, z_1)$, and $w'(z_1)$. This system yields a second-order equation in $E_2(z_1, z_1)$. Therefore, it has two solutions, and two corresponding expressions for $g'(z_1)$. If, for a given triplet e_1 , e_2 , and φ , one of the expressions is larger than 1 in absolute value and the other one smaller than 1 in absolute value, then the ABGE exhibits saddle-path stability.

Appendix A2

Solution method for the model with TFP shocks

The solution method used is similar to the one in Aiyagari (1994) albeit with two unknown variables, g and \hat{w} . The value function is obtained using global search over a discrete grid. Between grid points, the functions are interpolated with a cubic spline. The method involves the following steps. Start with an initial guess for g and \hat{w} . Solve the dynamic programming problem using the guess and obtain the policy function $E(z)$. Simulate an individual country's choice of e and the resulting z for T time periods, where T is very large. Use the data generated to check whether condition 4 in (24) is satisfied. If not, update the guess for \hat{w} using the bisection method and repeat the procedure until condition 4 holds. Then, check whether condition 3 in (24) is satisfied. If not, update the guess for g using the bisection method. Given the new guess for g , find the \hat{w} for which condition 4 is satisfied. Given this combination of \hat{w} and g , check whether condition 3 holds. If not, update the guess for g . Repeat the procedure until both conditions 3 and 4 in (24) are satisfied.