

Knowledge Spillovers and the World Income Distribution*

Stefano Gerosa[†]
Università di Roma - "Tor Vergata"

Abstract

This paper studies the implications for the long-run world income distribution (WID) of two possible knowledge spillovers structures. In an appropriate technology framework, where total factor intensity is an index of the technological level of an economy, countries will grab useful knowledge only from their technological neighborhood, the subset of countries which are "technologically near". An equilibrium WID generally clusters the world economy in distinct technological neighbourhoods, providing a mechanism for the endogenous emergence of convergence clubs. If spillovers are backward directed, either because free knowledge flows can only improve existing technologies or because of barriers to technology adoption, we show that an increase of the strength of spillovers raises inequality, while growth and inequality are negatively related.

JEL Classifications: O40, O33, D31

Keywords: Knowledge Spillovers, World Income Distribution, Inequality

*This paper is a revised version of a chapter from my doctoral dissertation at the University of Roma - "Tor Vergata". I would like to thank Pasquale Scaramozzino for constant help and advice. The usual disclaimer applies.

[†]E-mail: stefgerosa@tiscali.it

1 Introduction

The main goal of growth theory is to explain the existing pattern of cross-country income differences and possibly to predict its evolution over time. What are the forces that shape the actual world income distribution (WID) and what will be the long-run WID? Do observed income differences are a transitory phenomenon, so that in the long run the WID will converge to a degenerate point mass at a unique per capita income level, or are they symptoms of the emergence of convergence clubs, so that the long run WID will show the formation of clusters of rich and poor countries and a permanent degree of cross-country inequality? Is the growth process of every country independent of each other or are there diffusion forces that link them together?

The main insight of the Solow growth model, which is true even in the optimizing framework of Ramsey-Cass-Koopmans, is that, assuming a common world technology path $A(t)$, if preferences over consumption are equal, every country will converge to the same long-run income level whatever its initial capital stock $k(0)$. Differences in saving rates or in the parameters that shape the intertemporal profile of consumption translate into level effects on per capita income: in the absence of productivity differences, the growth path of an economy is completely determined by the factor accumulation process, which in turn is supposed independent of technological change.

The conclusion of a decade of empirical research started with the contribution of Mankiw, Romer and Weil (1992) and summarized by Caselli (2005) is that the actual dispersion of cross-country incomes cannot be explained by the dispersion of observables, physical and human capital: there are large productivity differences, computed as a residual with respect to measured factor differences, that account for a large part of the observed income dispersion, so that the assumption of a common world technology should be rejected.

While cross-country growth regressions and growth accounting take a snapshot of the WID to study its static properties, another direction of empirical research started by Quah (1997) focuses on the dynamic properties of the WID. The evolution of the WID seems to show the emergence of bimodality, which can be interpreted as a "convergence club" behaviour of the world economy: countries relatively close to the leading ones are converging to similar per capita incomes, while poor countries are diverging from the upper club.

This paper is part of ongoing research that tries to explain both the existence of productivity differences and the dynamic features of the WID using a specification for the traditional productivity index A that includes knowledge spillovers, based on Eeckhout and Jovanovic (2002): physical-human capital accumulation and technological efficiency are linked in a specific and testable way.

Most of the growth literature that tackles the issue of technology diffusion assumes that knowledge spillovers among countries are an increasing function of the distance from the technological frontier: the farther an economy is from the leader, the greater the knowledge it can freely use, the fastest its productivity growth. This "advantage of backwardness" hypothesis, originally formulated by Gerschenkron (1962), seems to be challenged by the persistence of cross-country productivity differences: even growth miracles such as those of some East Asian countries (Taiwan, Hong Kong, South Korea, Singapore) or of

China have been shown by Young (1995)(2003), using careful growth accounting analysis, to be cases of rapid factor accumulation rather than of exceptional productivity growth.

The main insight of the appropriate technology literature, starting with the seminal paper by Atkinson and Stiglitz (1969), is that a neoclassical production function, which maps factor intensity into output levels, is just the continuous limit of an increasing number of production processes, each one expressed by a unique capital-labor ratio; as Atkinson and Stiglitz point out:

"different points on the curve still represent different processes of production, and associated with each of these processes there will be certain technical knowledge specific to that technique [...] if one brings about a technological improvement in one of this blue-prints this may have little or no effect on the other blueprints"

so that it is realistic to assume a relative independence of each technique: technological change should be modeled not as a general shift of the production function, but as a localized shift which affects only a neighborhood of the improved technology and consequently only a part of the production function, and knowledge spillovers between interacting economies are likely to be local rather than global.

This argument can be applied by assuming that a country whose technological state is measured, as in Basu and Weil (1998), by the factor intensity k , which must be interpreted as an aggregate of human and physical capital, can only have access to the knowledge of countries that are in a neighborhood of k , simply because this is the only useful knowledge with respect to its technology level. As a result, the factor accumulation process and the dynamics of productivity are linked, and the growth experience of every country depends on the entire cross-section distribution of factor intensity and not on a predetermined target, such as its average as in Lucas (1988) or the maximum of the support as in Aghion-Howitt (1998).

Using this framework, it is possible to show that the set of the equilibria is very different from that of the Solow-Ramsey model and from that of Lucas (1988), (2001). Contrary to the former, even if all countries share the same preferences over consumption the equilibrium WID can be different from a degenerate point mass. Contrary to the latter, not any income distribution is an equilibrium and the long run WID is in general clustered, its density mass concentrated over disconnected intervals of the support, delivering an explanation from the convergence club phenomenon.

We also examine another possible structure of technological spillovers, that reverts conventional wisdom about cross-country knowledge diffusion, analyzing the case in which spillovers are an increasing function of a country factor intensity: an economy can spill free knowledge only about technologies it has already developed, so that the more advanced a country the greater the information flow it can intercept. As Boldrin and Levine (2003) suggest, if the non rival character of ideas is a property of their immaterial nature, there exist also a cost associated with learning and implementing ideas into an actual production

process: it is more likely that an advanced economy can integrate, freely or at near zero-cost, knowledge about inferior technologies into its framework rather than the contrary. Technological breakthroughs in nanotechnologies are unlikely to affect productivity in a developing country that lacks physical and human capital resources needed to adopt them, while often a technological advanced economy is able to extract useful knowledge even from traditional and low-tech sectors. A recent case is the one reported by the Economist (2006) on the conflict between Starbucks, the world's largest multinational chain of coffee shops, and Ethiopia, the poorest country in the world, over the recognition of intellectual property rights on three varieties of coffee-beans (Sidamo, Yirgacheffe and Harrar) originally developed in some Ethiopian regions: Starbucks coffee shops sell Sidamo and Harrar coffees up to 27 \$ a pound because of their exotic origin and because of the beans'speciality status, while "Ethiopian coffee farmers only earn between 0.58 and 1.16 \$ for their crop, barely enough to cover the cost of production"¹.

Another channel through which knowledge spillovers might flow from low to high factor intensity countries is international migration: Docquier and Marfouk (2006) show that, for a given source country, emigrants are almost universally more skilled than non emigrants, while Beine, Docquier and Rapoport (2001) show that the net effect of this brain-drain phenomenon is negative for the majority of developing countries.

It is also possible to interpret this knowledge spillovers structure in a more conventional way, as representing the existence of barriers to technology adoption, as in Parente and Prescott (1994): low- k countries can search a limited portion of the distribution of existing technologies, since they lack the physical and human capital infrastructures needed to adopt them, the barrier being aggregate capital intensity *relative* to the maximum K_t , the technological leader at time t .

We show that *backward spillovers* of this kind generate, if copying is undirected and the same preferences over consumption are shared by every country, an equilibrium world factor distribution (WFD) which depends on the strength of spillovers: when spillovers are high a large part of its density mass is concentrated near the lower end of the WFD, since factor-scarce countries are drawn near each other to enlarge the quantity of accessible knowledge, while when spillovers are low the WFD density mass concentrates toward the technological frontier. An increase of the intensity of spillovers always enlarge the support of the WFD and WID, raising inequality.

This results shed a new light on those of Eeckhout and Jovanovic (2002): knowledge spillovers raise inequality not because of their direction, forward or backward along the distribution of factor intensity, but because of their mere existence. Moreover, we find that growth and inequality are negatively correlated: the more concentrated the distribution of relative factor intensity, the closer the technologies operated by countries interacting in the world economy, the higher the degree of increasing returns and the associated common growth rate of the world economy, consistently with empirical evidence documented by

¹Tadesse Meskela, head of the Oromia Coffee Farmers Cooperative Union in Ethiopia reported in Seager (2006). Starbucks attempt to block Ethiopia's application to the US patent and trademark office could cost Ethiopia up to a 92\$ millions reduction in potential earnings, according to the independent development agency OxFam, a 25%increase of Ethiopian annual coffee's export earnings.

Sala-i-Martin (2006).

The rest of the paper is organized as follows. Section 2 presents the model, introduces the two possible structures of knowledge spillovers and characterizes optimal growth of the world economy. Section 3 studies the equilibrium WFD and WID with appropriate and backward knowledge spillovers and analyses the relationship between spillovers, growth and inequality. Section 4 concludes.

2 The Model

2.1 Technology and the Structure of Knowledge Spillovers

Suppose a world economy consisting of a unit mass of countries whose total capital stock varies in the time-varying support $[k_{\min}, K_t]$, and in which there is no trade and no international capital flow. Each country produce the final consumption good using an aggregate of human and physical capital k , which should be considered also as an index of the technological level reached by the economy: the higher the factor intensity, the more advanced the technology operated by the economy.

The production function in per capita terms is given by

$$y_t = F_t(k) = A_t(k) \cdot k \quad (1)$$

where $A_t(k)$ is a productivity parameter which represents the amount of technological knowledge a country can dispose of.

In particular we specify this technology index as

$$A_t(k) = S_t(k)^\beta G_t^{1-\beta} \equiv \left[1 + \int_D \alpha \left(\frac{k}{K_t} \right) h_t \left(\frac{k}{K_t} \right) dz \right]^\beta G_t^{1-\beta} \quad \text{for all } k \in [k_{\min}, K_t] \quad (2)$$

where G_t is an efficiency index which grows at the constant and exogenous rate g shared by each and every country, $z = k/K_t$ is the country factor intensity relative to the supremum K_t of the distribution of k among countries, $H_t(z)$ is the distribution function of z defined over $[z_{\min}, 1]$, $h_t(z) = H_t'(z)$ is the density of z , β is a parameter which measures the intensity of the spillover force, $\alpha(z)$ is a positive bounded function that expresses the direction of copying (if $\alpha' > 0$ copying is directed toward the high- k countries, if $\alpha' < 0$ it is directed toward the low k ones) and D is the domain over which spillovers act.

Equation (2) specifies country technology level as a Cobb-Douglas aggregate of two technical knowledge components: a common general part G_t which is not country-specific and that can be thought as general knowledge, and a knowledge spillovers component S_t that comes from cross-country interactions. The knowledge spillovers part of (2) is totally deterministic, but it can be interpreted as an expectation of the amount of knowledge a firm can copy drawing from the subset $D \subseteq [z_{\min}, 1]$ of the support of $h_t(z)$: in fact, interpreting $h_t(z)$ as a

probability density, the integral in S_t represent the mean of the copying function $\alpha(z)$ conditional on the fact that $z \in D$. Note that the knowledge spillovers component of technology is bounded from below by 1.

The crucial step is the choice of the subset D over which spillovers flow:

- if $D = [z_{\min}, 1]$ then $A_t(k)$ depends on the average level of the copying function $\alpha(z)$ over the entire relative cross-country factor distribution, as in Lucas (1988), in which there is an externality based on the average level of human capital, and in Romer (1986). In this case

$$A'_t(k) = 0 \quad k \in [k_{\min}, K_t] \quad (3)$$

Since the spillover force is global and acts over the whole support, the position of the single country is irrelevant and the marginal effect of an increase of the relative factor intensity is null.

- if $D = [\frac{k}{K_t}, 1] \equiv [z, 1]$, as in Eeckhout and Jovanovic (2002), then each country is supposed to freely extract useful technical knowledge from all countries that operate superior technologies. In this case the spillover force is negatively related to factor intensity: low k -countries have access to the knowledge of a large part of the cross-country distribution of techniques, while high k -countries can't copy much and have to rely more on investment.

In this framework the knowledge access function $A_t(k)$ is decreasing in k

$$\begin{aligned} A'_t(k) &= -\frac{\beta}{K_t} S_t(k)^{\beta-1} G_t^{1-\beta} \alpha\left(\frac{k}{K_t}\right) h_t\left(\frac{k}{K_t}\right) \\ &= -\frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)}\right)^{1-\beta} \alpha(z) h_t(z) \leq 0 \end{aligned} \quad (4)$$

for all $k \in [k_{\min}, K_t]$ or for all $z \in [z_{\min}, 1]$.

With this kind of *forward knowledge spillovers* the marginal effect on productivity of an increase in relative capital intensity is always negative: increasing its own stock of factors, advancing its technological level, and gaining rank along the cross-country factor distribution, reduces the amount of knowledge a country can copy. This is why the Eeckhout-Jovanovic model cannot explain the WID: growth accounting exercises, as Hall and Jones (1999, Table 1), finds a positive correlation between income levels, levels of physical and human capital and TFP. High k -countries display at the same time higher levels of efficiency, a feature that a representation of technology in terms of forward knowledge spillovers cannot explain.

The appropriate technology assumption can be introduced in the model by modifying the specification of the technology index $A_t(k)$: a country total factor intensity k is also an index of its technological level, so that it can

grab useful knowledge only from neighbouring countries along the world factor distribution, the ones which are "technologically close". One can let $D = [\underline{\delta}(k/K_t), \bar{\delta}(k/K_t)]$, where $\bar{\delta} = (1 + \delta)$ and $\underline{\delta} = (1 - \delta)$ so that

$$A_t(k) = \begin{cases} G_t^{1-\beta} \left[1 + \int_{z_{\min}}^{\bar{\delta}(k/K_t)} \alpha(z) h_t(z) dz \right]^\beta & k \in \left[k_{\min}, \frac{k_{\min}}{\underline{\delta}} \right] \\ G_t^{1-\beta} \left[1 + \int_{\underline{\delta}(k/K_t)}^{\bar{\delta}(k/K_t)} \alpha(z) h_t(z) dz \right]^\beta & k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\delta} \right] \\ G_t^{1-\beta} \left[1 + \int_{\underline{\delta}(k/K_t)}^1 \alpha(z) h_t(z) dz \right]^\beta & k \in \left[\frac{K_t}{\delta}, K_t \right] \end{cases} \quad (5)$$

Now we have two parameters which control the spillover's amplitude: β that identifies the general strength of the spillover force and δ that measures the extension of the appropriate technology spillovers. Notice that as $\delta \rightarrow 0$ and the appropriate technology spillovers are null (i.e. every capital intensity k corresponds to a different production process whose technology is independent from each other), the model converges to a traditional growth model with exogenous technological change, while for $\delta \rightarrow \frac{1-z_{\min}}{z_{\min}}$ then $D \rightarrow [z_{\min}, 1]$ and the model converges to the Lucas-Romer setting.

Notice that in (5) the amplitude of the technological neighborhood is proportional to the relative factor intensity of the country, that indexes its technological level: the more advanced the technology a country does operate, the larger the set of neighbouring technologies from which it can get useful knowledge (Figure 1). This assumption may seem strange at first: why a developed economy should be able to freely obtain useful knowledge over an inferior technology from a less developed one, while the latter cannot access the more advanced technology of the former?

The general idea is that an economy using a backward technology cannot simply spill knowledge from a frontier technology without a costly investment, while an advanced technology often can integrate in its framework knowledge from an inferior technology (one can think of the case of the genetic improvement of crop varieties, often using traditional agricultural knowledge of local communities). This specification seems to be consistent with the empirical evidence discussed by Caselli (2005) and Feyrer (2003): there is a significant positive correlation of the productivity residual A and measures of physical and human capital stock, so that the cross-country factor and TFP distributions seem to depend on each other. Equation (5) has the potential to explain cross-country variation in both per capita income and TFP *through* cross-country variation in factor intensity: in a companion paper, Gerosa (2007), we provide evidence supporting this specification.

The immediate consequence of the localization of knowledge spillovers, coherent with the appropriate technology assumption, is that the marginal effect of an increase of a country capital intensity k now depends on the shape of the WFD distribution: by moving up along the distribution a country passes into a different technology neighborhood and the variation in the knowledge flow depends on the WFD densities at the extremes of its current neighborhood

$$A'_t(k) = \begin{cases} \frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)} \right)^{1-\beta} \bar{\delta} \alpha(z_+) h_t(z_+) \geq 0 & k \in \left[k_{\min}, \frac{k_{\min}}{\delta} \right] \\ \frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)} \right)^{1-\beta} [\bar{\delta} \alpha(z_+) h_t(z_+) - \underline{\delta} \alpha(z_-) h_t(z_-)] & k \in \left[\frac{k_{\min}}{\delta}, \frac{K_t}{\delta} \right] \\ -\frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)} \right)^{1-\beta} \underline{\delta} \alpha(z_-) h_t(z_-) \leq 0 & k \in \left[\frac{K_t}{\delta}, K_t \right] \end{cases} \quad (6)$$

where $z_+ = \bar{\delta}z$ and $z_- = \underline{\delta}z$.

Note that for countries that are sufficiently far from the boundary values of the support $[k_{\min}, K_t]$, i.e. for $k \in \left[\frac{k_{\min}}{\delta}, \frac{K_t}{\delta} \right]$, the marginal effect of a factor intensity increase now depends on the shape of the WFD in the technological neighborhood in which the economy is located, i.e. on the values $h_t(z_+)$ and $h_t(z_-)$, and on the size of the technological neighborhood, as measured by the parameter δ : for certain WFD this effect and the elasticity can be positive, so that an increase of its capital intensity and a gain in rank along the WFD can improve the productivity of a country. On the other hand, for very low- k (high- k) countries the marginal effect on productivity of an increase in k is always positive (negative), since their technological neighborhood is left-truncated (right-truncated).

Backward knowledge spillovers can be introduced letting $D = \left[\frac{k_{\min}}{K_t}, \frac{k}{K_t} \right] \equiv [z_{\min}, z]$ so that the technology index becomes:

$$A_t(k) = S_t(k)^\beta G_t^{1-\beta} \equiv \left[1 + \int_{z_{\min}}^{k/K_t} \alpha(z) h_t(z) dz \right]^\beta G_t^{1-\beta} \quad \text{for all } k \in [k_{\min}, K_t] \quad (7)$$

Here knowledge spillovers come from the sampling of the portion of the distribution of capital intensities indexing technologies already reached by the economy through costly investment: free informational leakage can only improve or upgrade existing production processes.

It is possible to interpret such a specification of the efficiency index also in terms of barriers to technology adoption, in the spirit of Parente and Prescott (1994): here low- k countries face a barrier that preclude them the access to the upper portion of the distribution of the world technologies, reducing their aggregate technological efficiency.

In this case the marginal effect of an increase in capital intensity is always positive

$$A'_t(k) = \frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)} \right)^{1-\beta} \alpha(z) h_t(z) \geq 0 \quad \text{for all } k \in [k_{\min}, K_t] \quad (8)$$

By raising its own aggregate capital intensity a country also increases accessed knowledge since it can extract useful knowledge from a greater part of

$h_t(z)$. In Gerosa (2007) we show that also (7) is consistent with empirical evidence.

Given (1), an increase in k through investment has a positive direct effect on a country per capita output but also an indirect effect, because it changes the component of a country's resources it can freely borrow from other countries; in particular

$$F'_t(k) = A'_t \cdot k + A_t(k) = A_t(k) [1 + \varepsilon_t(k)] \quad (9)$$

where $\varepsilon_t(k) = \frac{A'_t(k) \cdot k}{A_t(k)}$ is the elasticity of the technology access function $A_t(k)$, which measures the percentage variation in productivity associated with a marginal increase in k : this elasticity is a crucial object in the following general equilibrium analysis, because an equilibrium WFD should equalize $\varepsilon_t(k)$ among countries (from (9) is also evident that $\varepsilon_t(k)$ must be greater than -1, ensuring that the marginal productivity of capital is not negative).

With forward spillovers à la Eeckhout-Jovanovic one has

$$\varepsilon_t(k) = -\beta z \frac{\alpha(z) h_t(z)}{S_t(k)} \quad (10)$$

which measure the percentage loss associated with a one per cent increase in k : in this case (9) reveals that the production technology always display *decreasing returns to scale* for each country in the world economy, their degree varying with the shape of the WFD.

With appropriate technology spillovers one has

$$\varepsilon_t(k) = \begin{cases} \frac{\beta z}{S_t(k)} \bar{\delta} \alpha(z_+) h_t(z_+) \geq 0 & k \in \left[k_{\min}, \frac{k_{\min}}{\delta} \right] \\ \frac{\beta z}{S_t(k)} [\bar{\delta} \alpha(z_+) h_t(z_+) - \underline{\delta} \alpha(z_-) h_t(z_-)] & k \in \left[\frac{k_{\min}}{\delta}, \frac{K_t}{\delta} \right] \\ -\frac{\beta z}{S_t(k)} \underline{\delta} \alpha(z_-) h_t(z_-) \leq 0 & k \in \left[\frac{K_t}{\delta}, K_t \right] \end{cases} \quad (11)$$

so that very low- k (high- k) countries always have increasing (decreasing) returns to scale, while for countries with complete technological neighborhoods returns to scale depend on the shape of the WFD and on the parameter δ which controls the size of the neighborhoods.

Finally with backward spillovers one has

$$\varepsilon_t(k) = \beta z \frac{\alpha(z) h_t(z)}{S_t(k)} \quad (12)$$

and the percentage gain associate to a marginal increase in k also measures the degree of increasing returns of each economy.

The presence of increasing returns to scale with appropriate technology and backward spillovers implies that a competitive equilibrium in general does not

exist: if k is paid its marginal productivity then $F'_t(k)k = A_t(k)k[1 + \varepsilon_t(k)] > A_t(k)k = F_t(k)$ and factor payments exceed the value of production. A competitive equilibrium exists if, as in Romer (1986), in every country there exist a large number N of small firms so that aggregate capital intensity is given by the sum $k = \sum_{i=1}^N k_i$ and the effect of an increase of the capital intensity of a single firm is irrelevant: in this case every firm takes A_t as given, while in the aggregate A_t is an increasing function of aggregate capital intensity k . Here we focus on optimal growth, that is on a world economy composed by countries whose factor accumulation trajectories are chosen by social planners that maximize total utility taking care of the externality associated with the presence of knowledge spillovers.

2.2 Optimal Growth of the World Economy

The economy faces convex adjustment costs that are proportional to its output level, so that passing from k to \tilde{k} in a unit time interval costs $y \cdot C\left(\frac{\tilde{k}}{k}\right)$ units of foregone consumption, where $C' > 0, C'' > 0$ and the total capital stock depreciates at the constant rate η so that $C(1 - \eta) = 0$.

The social planner problem is that of maximizing total utility over an infinite horizon given the production technology (1): this is equivalent to maximizing net production coherently with the representative agent's preferences over consumption. Period t production net of adjustment costs is given by $\left[1 - C\left(\frac{\tilde{k}}{k}\right)\right]y$, so that the PDV of national production of over the infinite horizon $v_t(k)$ solves

$$v_t(k) = \max_{\tilde{k}} \left\{ \left[1 - C\left(\frac{\tilde{k}}{k}\right)\right] A_t(k) \cdot k + \frac{1}{1+r} v_{t+1}(\tilde{k}) \right\} \quad (13)$$

where r is the exogenous world interest rate.

Using the first-order and envelope conditions as in Eeckhout-Jovanovic (2002), it is possible to obtain a second-order difference equation in k which characterizes the optimal factor accumulation path

$$(1+r)C'\left(\frac{k_{t+1}}{k_t}\right)A_t(k_t) = A_{t+1}(k_{t+1})\left\{\frac{k_{t+2}}{k_{t+1}}C'\left(\frac{k_{t+2}}{k_{t+1}}\right) + \left[1 - C\left(\frac{k_{t+2}}{k_{t+1}}\right)\right][1 + \varepsilon_{t+1}(k_{t+1})]\right\} \quad (14)$$

then one can proceed in finding a solution in which every economy accumulates k at the same rate x_k , so that $h_t(z) = h(z)$ holds and the cross-country factor distribution is time-invariant, while total factor productivity grows only through the increase of common knowledge, $x_A \equiv \frac{A_{t+1}}{A_t} = (1+g)^{1-\beta}$ from (2).

Now suppose that every economy shares the same preferences over consumption, that are characterized by a common CRRA utility function for the representative consumer

$$U = \sum_{t=0}^{\infty} \rho^t \frac{c_t^{1-\gamma} - 1}{1-\gamma} \quad \rho < 1 \quad \gamma > 0 \quad (15)$$

where ρ is the subjective discount rate and γ controls the elasticity of intertemporal substitution of consumption. Now it is possible to eliminate r from (14) using the FOC for consumption and the fact that the growth rate of consumption must equal the growth rate of potential output, or $x_c = x_k x_A$ from (1).

The crucial assumption of this general equilibrium analysis is

$$\varepsilon_t(k) = \varepsilon \quad \text{for all } k \text{ and for all } t \quad (16)$$

An equilibrium WFD is such that the elasticity of the access function $A_t(k)$ to k is equalized among countries: if the percentage change associated with a marginal shift along the support of the WFD is completely independent of the position of the country along the WFD, then there is no incentive to change position and the factor distribution, and by consequence the world income distribution, is the equilibrium one.

If (16) holds, (14) reduces to an equation in the single variable $x_k = x$

$$\Psi(x) \equiv \frac{\left[\frac{(\xi x)^\gamma}{\xi^\rho} - x \right] C'(x)}{1 - C(x)} = 1 + \varepsilon \quad (17)$$

where $\xi = (1 + g)^{1-\beta}$.

We show in the Appendix that Ψ is a strictly increasing function, so that, given a value for ε , the common equilibrium growth rate of every economy is uniquely determined, provided some restrictions on the cost function $C(x)$ and the elasticity of marginal utility γ hold (in the Appendix we present also a detailed derivation of the optimization procedure).

Proposition 1

If the cost function verifies $\left[\frac{\gamma[\xi(1-\eta)]^{\gamma-1}}{\rho} - 1 \right] \frac{(1-\eta)C'(1-\eta)}{1-C(1-\eta)} < 1 + \varepsilon$ and if $\gamma > \frac{\xi x}{1+r}$ then, for any $\varepsilon > -1$, there exist a unique growth rate x which solves (17), where $x \in \left(1 - \eta, \min \left[C^{-1}(1), \frac{(1+r)}{\xi} \right] \right)$. Moreover, the equilibrium growth rate x is an increasing function of ε .

Proof. Appendix A. ■

As an example, if adjustment costs are defined by the convex function $C(x) = B[x - (1 - \eta)]^\theta$, with $B > 0$ and $\theta > 1$, and $\gamma > 1$ then an equilibrium always exists for any given $\varepsilon > -1$. The lower bound on the growth rate is given by the situation in which investments are null and the capital stock depreciates at the constant rate η , while the upper bound is given by the two restrictions of the No-Ponzi game condition and of the non-negativity of net production.

Notice that:

- Since with forward spillovers model $\varepsilon_t(k)$ is always negative, the equilibrium growth rate x_k^* is decreasing in the absolute value $|\varepsilon|$ of this elasticity: the greater the accessed knowledge loss associated with an increase in firm size, the larger the incentives to stay back and not accumulate capital for firms and the lower the steady-state growth rate.
- The no-externality case with $\beta = 0$ (which is just the standard Ak model) and the k -independent externality case of Romer (1986) or Lucas (1988) where A_t is equal to the mean of the distribution and the position of the single country is irrelevant ($\partial A_t / \partial k = 0$), are characterized by the condition $\varepsilon = 0$ and share the same growth rate, which is greater than with forward spillovers: moreover in both cases any initial distribution $H_0(z)$ simply replicates itself and can be sustained in a BGP equilibrium, so that there is no restriction the theory can impose on the equilibrium WFD and WID.
- With appropriate spillovers ε can in principle be either positive or negative, and the WFD and WID can be such that the world economy displays either decreasing or increasing returns to scale: in the next section we show that an equilibrium WFD with appropriate technology knowledge spillovers always induce constant returns to scale for the world economy.
- With backward spillovers ε is always positive and the world economy always displays increasing returns to scale: moreover the growth rate of the world economy is an increasing function of the degree of increasing returns, which in turn depend on the equilibrium WFD.

In any case, the fundamental issue is the cross-country equalization of the elasticity $\varepsilon_t(k)$ of the technology index which depend on the cross-country distribution of factor intensities: an equilibrium WFD equalizes $\varepsilon_t(k) = \varepsilon$ among countries for all possible type of knowledge spillovers, but the shape of the equilibrium WFD depends on the spillovers structure.

3 The Equilibrium World Income Distribution

3.1 The Equilibrium WID with Appropriate Technology Knowledge Spillovers

In the general framework given by (1) and (2) there are three elements which determine the shape of the equilibrium WFD: the copying function $\alpha(z)$, the strength of the spillovers as measured by the parameter β and the elasticity ε of the knowledge function $A_t(k)$. The crucial element for a distribution to be an equilibrium one is the equalization of $\varepsilon_t(k)$ among countries, which could be conjectured to be the outcome of the evolution of the world economy starting with an initial distribution of relative factor intensities/technological processes $H_0(z)$. With forward spillovers the *direction* of the force represented by this

elasticity is the same for every country in the support, even if its magnitude can vary with its position along the WFD: accumulating capital and gaining rank along the WFD reduces the quantity of accessed knowledge for every country in the support, so that the existence of spillovers always acts as an incentive to stay back (it is the "free-riding case", as Eeckhout and Jovanovic (1998) defined it in the working paper version of their paper).

In the appropriate technology framework, as it is specified in (5), the marginal effect of an increase in capital intensity and the sign of the elasticity $\varepsilon_t(k)$ is ambiguous for countries sufficiently far from the boundary values of the support and opposite in direction for very low- k and very high- k countries as (11) shows: it is a *boundary effect* that pushes backward economies up along the WFD, since they always have a net positive effect by moving forward along the distribution, and pulls the advanced ones back because, being the technological leaders, they have no larger technological neighborhood to reach. The presence in the appropriate technology framework of these two opposite forces at the edges of the WFD support makes it difficult to find an equilibrium which is defined as

Definition 1

An equilibrium for the world economy described by (1), (5) and (15) consists of:

- an elasticity $\varepsilon > -1$
- a growth rate $x \in \left(1 - \eta, \min \left[C^{-1}(1), \frac{(1+r)}{\xi} \right] \right)$
- a lower bound $z_{\min} \in [0, 1]$ of the support of the world factor distribution
- a world factor distribution, i.e. a function $h(z) : [z_{\min}, 1] \rightarrow \mathbb{R}$ that verifies

$$\int_{z_{\min}}^1 h(z) dz = 1$$

such that (16) and (17) hold.

Since $\beta > 0$, $\alpha(z)$ and $h_t(z)$ are positive functions so that the integrand in $A_t(k)$ is always positive, equation (11) shows that a solution with $\varepsilon > 0$ cannot exist because for $k \in \left[\frac{K_t}{\delta}, K_t \right]$ the elasticity $\varepsilon_t(k)$ is always negative: unless we have equality and the WFD coincide with a degenerate point mass centered in $z = 1$ and every country has the same factor intensity, the upper tail of the distribution always feels a force that pulls it back. It follows that in equilibrium $\varepsilon \leq 0$.

Can an equilibrium with $\varepsilon < 0$ exist? It seems that at the lower tail of the WFD there is a symmetric problem, because for $k \in \left[k_{\min}, \frac{k_{\min}}{\delta} \right]$ the elasticity is always positive. But as k_{\min} gets smaller and smaller this part of the support shrinks and the elasticity, as a measure of the force that pushes up backward countries, also becomes smaller and smaller. As $k_{\min} \rightarrow 0$ (11) is reduced to a system of two equations

$$\frac{\beta z}{S_t(k)} [\bar{\delta}\alpha(z_+) h_t(z_+) - \underline{\delta}\alpha(z_-) h_t(z_-)] = \varepsilon \quad k \in \left(0, \frac{K_t}{\delta} \right] \quad (18)$$

$$-\frac{\beta z}{S_t(k)} \underline{\delta} \alpha(z_-) h_t(z_-) = \varepsilon \quad k \in \left[\frac{K_t}{\delta}, K_t \right] \quad (19)$$

and an equilibrium WFD with $\varepsilon < 0$ may exist.

In order to characterize the equilibrium WFD one can proceed in the same way as Eeckhout and Jovanovic (2002). Since (16) holds we have:

$$\frac{A'_t(k)}{A_t(k)} = \frac{\varepsilon}{k} \rightarrow A_t(k) = \lambda_t k^\varepsilon \quad (20)$$

for some sequence of constants of integration $\{\lambda_t\}$.
Substituting from (5) gives:

$$G_t^{1-\beta} \left[1 + \int_{\underline{\delta}(k/K_t)}^{\bar{\delta}(k/K_t)} \alpha(z) h(z) dz \right]^\beta = \lambda_t k^\varepsilon \quad k \in \left(0, \frac{K_t}{\delta} \right] \quad (21)$$

and

$$G_t^{1-\beta} \left[1 + \int_{\underline{\delta}(k/K_t)}^1 \alpha(z) h(z) dz \right]^\beta = \lambda_t k^\varepsilon \quad k \in \left[\frac{K_t}{\delta}, K_t \right] \quad (22)$$

and calculating the two sides of (21) and (22) for $k = \frac{K_t}{\delta}$

$$\lambda_t = \left(\frac{K_t}{\delta} \right)^{-\varepsilon} G_t^{1-\beta} \left[1 + \int_{\underline{\delta}/\delta}^1 \alpha(z) h(z) dz \right]^\beta \quad k \in (0, K_t] \quad (23)$$

and eliminating λ_t from (21) and (22) leaves with an implicit characterization of $h(z)$ for the two parts of the support

$$1 + \int_{z_-}^{z_+} \alpha(z) h(z) dz = (\bar{\delta} z)^{\frac{\varepsilon}{\beta}} \left[1 + \int_{\underline{\delta}/\delta}^1 \alpha(z) h(z) dz \right] \quad z \in \left(0, \frac{1}{1+\delta} \right] \quad (24)$$

$$1 + \int_{z_-}^1 \alpha(z) h(z) dz = (\bar{\delta} z)^{\frac{\varepsilon}{\beta}} \left[1 + \int_{\underline{\delta}/\delta}^1 \alpha(z) h(z) dz \right] \quad z \in \left[\frac{1}{1+\delta}, 1 \right] \quad (25)$$

which can be rewritten as

$$[(1 + \delta)z]^{\frac{\varepsilon}{\beta}} = \begin{cases} \frac{1 + \int_{z_-}^{z_+} \alpha(z)h(z)dz}{1 + \int_{\frac{\delta}{1+\delta}}^1 \alpha(z)h(z)dz} & z \in \left(0, \frac{1}{1+\delta}\right] \\ \frac{1 + \int_{z_-}^1 \alpha(z)h(z)dz}{1 + \int_{\frac{\delta}{1+\delta}}^z \alpha(z)h(z)dz} & z \in \left[\frac{1}{1+\delta}, 1\right] \end{cases} \quad (26)$$

But, while the left hand side of this equilibrium equation diverges as $z \rightarrow 0$ since $\varepsilon < 0$, the right hand side of (26) is bounded. In the undirected copying hypothesis ($\alpha(z) = \alpha$) since $h(z)$ is a density function its integral over a subset of the support cannot be greater than 1 and the right hand side must be in $\left[\frac{1}{1+\alpha}, 1 + \alpha\right]$. If copying is directed the boundedness of the RHS of (26) follows from the fact that $\alpha(z)$ is bounded in $[z_{\min}, 1]$.

It follows that an equilibrium WFD with $\varepsilon < 0$ is impossible, and the only possible equilibria in the appropriate technology framework are those with $\varepsilon = 0$: appropriate technology knowledge spillovers always entail an equilibrium world economy characterized by constant returns to scale.

Imposing the equilibrium condition $\varepsilon = 0$ greatly restricts the shape and properties of the equilibrium distribution. Suppose that $h(z)$ is an equilibrium WFD defined over $[z_{\min}, 1] \subseteq [0, 1]$: once again, since $\beta > 0$, $\alpha(z)$ and $h(z)$ are positive functions and $A_t(k)$ is always positive, by (11) the equilibrium condition $\varepsilon = 0$ is verified at the boundaries of the support if and only if the density function vanishes over the parts of the support from which low- z and high- z countries receive knowledge spillovers

$$h(z_+) = 0 \quad z \in [z_{\min}, \frac{z_{\min}}{1-\delta}] \quad (27)$$

$$h(z_-) = 0 \quad z \in [\frac{1}{1+\delta}, 1] \quad (28)$$

so that

$$h(z) = 0 \quad z \in [(1+\delta)z_{\min}, \frac{(1+\delta)}{(1-\delta)}z_{\min}] \quad (29)$$

$$h(z) = 0 \quad z \in [\frac{(1-\delta)}{(1+\delta)}, 1-\delta] \quad (30)$$

For countries that have a complete technological neighborhood the equilibrium condition is in general

$$(1 + \delta) \alpha(z_+) h(z_+) = (1 - \delta) \alpha(z_-) h(z_-) \quad z \in \left[\frac{z_{\min}}{\delta}, \frac{1}{\delta} \right] \quad (31)$$

so that in equilibrium

$$h_t(z_+) = 0 \Rightarrow h_t(z_-) = 0 \quad z \in \left[\frac{z_{\min}}{\delta}, \frac{1}{\delta} \right] \quad (32)$$

Since by definition $z_+ = \frac{(1+\delta)}{(1-\delta)} z_-$ (27) and (28) imply also²

$$h(z) = 0 \quad z \in \left[\frac{(1+\delta)^2}{1-\delta} z_{\min}, \left(\frac{1+\delta}{1-\delta} \right)^2 z_{\min} \right] \quad (33)$$

$$h(z) = 0 \quad z \in \left[\left(\frac{1-\delta}{1+\delta} \right)^2, \frac{(1-\delta)^2}{1+\delta} \right] \quad (34)$$

Iterating this passage one obtains

$$h(z) = 0 \text{ for } \begin{cases} z \in P_n = \left[\frac{(1+\delta)^n}{(1-\delta)^{n-1}} z_{\min}, \left(\frac{1+\delta}{1-\delta} \right)^n z_{\min} \right] \\ z \in Q_n = \left[\left(\frac{1-\delta}{1+\delta} \right)^n, \frac{(1-\delta)^n}{(1+\delta)^{n-1}} \right] \end{cases} \quad n \in \mathbb{N} \quad (35)$$

and denoting with $Z = \left[\left(\bigcup_{n \in \mathbb{N}} P_n \right) \cup \left(\bigcup_{n \in \mathbb{N}} Q_n \right) \right] \cap [z_{\min}, 1]$ the subset of the WFD support over which the density must be zero according to (35), we obtain

Proposition 2

The set of the equilibrium world factor distributions for the world economy described by (1), (5) and (15) consists of all distributions $h(z)$ defined over $[z_{\min}, 1]$, for some $z_{\min} \in [0, 1]$, such that

- $h(z) = 0$ for $z \in Z$
- $(1 + \delta) \alpha(z_+) h_t(z_+) = (1 - \delta) \alpha(z_-) h_t(z_-)$ for $z \in \left[\frac{z_{\min}}{\delta}, \frac{1}{\delta} \right]$ and $z \notin Z$

The associated equilibrium world income distribution is then obtained through (1).

²One can reason along this line: $(1 + \delta) z_{\min} = z_-$ for $z = \frac{(1+\delta)}{(1-\delta)} z_{\min}$, so that $z_+ = \frac{(1+\delta)^2}{1-\delta} z_{\min}$ and since $h[(1 + \delta) z_{\min}] = 0$ then (32) imposes also $h \left[\frac{(1+\delta)^2}{1-\delta} z_{\min} \right] = 0$. On the other hand $1 - \delta = z_+$ for $z = \frac{(1-\delta)}{(1+\delta)}$, and $z_- = \frac{(1-\delta)^2}{1+\delta}$. Since (30) imposes $h[(1 - \delta)] = 0$, then also $h \left[\frac{(1-\delta)^2}{1+\delta} \right] = 0$.

Thus, the set of equilibria for a world economy characterized by the presence of local spillovers or appropriate technology is very different from that of a world economy in which knowledge spillovers are global, either because of a common world technology assumption à la Solow or because the technology index depends on the average factor intensity *à la* Romer-Lucas: even if preferences over consumption are identical for each country, absolute convergence in relative factor intensity and income levels is not the only equilibrium outcome, as in a Solow setting, but also not every cross-country factor distribution and world income distribution are equilibria, as in the Romer-Lucas framework.

Here convergence of the WFD and WID to a degenerate point mass and absolute convergence in relative income levels is indeed an equilibrium, but in general an equilibrium WFD, and the associated equilibrium WID, will not be continuous over its support and its density mass will be concentrated over disconnected intervals of $[z_{\min}, 1]$: this can be interpreted as the equilibrium endogenous formation of convergence clubs, where countries cluster toward different long run relative factor intensity, or relative technological levels, and the world economy is characterized by a permanent degree of inequality (Figure 2 shows a possible configuration of the equilibrium WID). In an appropriate technology framework where knowledge spillovers diffuse over a localized portion of the world capital intensity-technology distribution, an equilibrium distribution is attained by splitting the world economy in separate technological clusters between which there are no knowledge spillovers: in this sense, the observed bimodality of the WID could be interpreted as a sign of a long run tendency toward a split of the world economy in two distinct technological neighbourhoods. In each technological cluster every economy shares the same level of efficiency even if capital intensities differ: every technological cluster behaves as a Romer-Lucas economy in which every distribution of factor intensity is an equilibrium.

3.2 The Equilibrium WID with Backward Knowledge Spillovers

The definition of an equilibrium for a world economy characterized by the presence of backward spillovers/barriers to technology adoption changes slightly, since in this case ε is always positive

Definition 2

An equilibrium for the world economy described by (1), (7) and (15) consists of:

- an elasticity $\varepsilon > 0$
- a growth rate $x \in \left(1 - \eta, \min \left[C^{-1}(1), \frac{(1+r)}{\xi} \right] \right)$
- a lower bound $z_{\min} \in [0, 1]$ of the support of the world factor distribution
- a world factor distribution, i.e. a function $h(z) : [z_{\min}, 1] \rightarrow \mathbb{R}$ that verifies

$$\int_{z_{\min}}^1 h(z) dz = 1$$

such that (16) and (17) hold.

Since (16) and (20) hold, for some sequence of constants of integration $\{\lambda_t\}$, substituting from (7) gives

$$G_t^{1-\beta} \left[1 + \int_{z_{\min}}^{k/K_t} \alpha(z) h_t(z) dz \right]^\beta = \lambda_t k^\varepsilon \quad \text{for all } k \in [k_{\min}, K_t] \quad (36)$$

and evaluating this identity at $k = k_{\min}$ one obtains

$$\lambda_t = G_t^{1-\beta} k_{\min}^{-\varepsilon} \quad (37)$$

and it is now possible to eliminate λ_t from (36)

$$1 + \int_{z_{\min}}^{k/K_t} \alpha(z) h_t(z) dz = \left(\frac{k}{k_{\min}} \right)^{\frac{\varepsilon}{\beta}} \quad \text{for all } k \in [k_{\min}, K_t] \quad (38)$$

Finally, noting that $\frac{k}{k_{\min}} = \frac{z}{z_{\min}}$ and differentiating both sides with respect to z one obtains

$$h(z) = \frac{\varepsilon}{\beta \alpha(z) z_{\min}^{\frac{\varepsilon}{\beta}}} z^{-(1-\frac{\varepsilon}{\beta})} \quad (39)$$

that, together with the restriction that $\int_{z_{\min}}^1 h(z) dz = 1$, gives the equilibrium

WFD for any ε and any copying function $\alpha(z)$.

Suppose copying is undirected and $\alpha(z) = \alpha$ (in the following we maintain this assumption). Then it is possible to obtain a closed form solution for the equilibrium WFD $h(z)$ imposing the normalization:

$$\int_{z_{\min}}^1 h(z) dz = \frac{\varepsilon}{\beta \alpha z_{\min}^{\frac{\varepsilon}{\beta}}} \int_{z_{\min}}^1 z^{-(1-\frac{\varepsilon}{\beta})} = 1 \quad (40)$$

from which

$$z_{\min} = \left(\frac{1}{1+\alpha} \right)^{\frac{\beta}{\varepsilon}} \quad (41)$$

and

$$h(z) = \frac{\varepsilon(1+\alpha)}{\beta \alpha} z^{-(1-\frac{\varepsilon}{\beta})} \quad \text{for } z \in \left[\left(\frac{1}{1+\alpha} \right)^{\frac{\beta}{\varepsilon}}, 1 \right] \quad (42)$$

In this case the shape of the equilibrium WFD is controlled by the ratio $\frac{\varepsilon}{\beta}$, that is by the ratio between the degree of increasing returns of the world economy and the strength of knowledge spillovers: if spillovers are weak ($\frac{\varepsilon}{\beta} > 1$) then the WFD is increasing and a large mass of the world economy is concentrated near the technological frontier, if spillovers are strong ($\frac{\varepsilon}{\beta} < 1$) then the WFD is decreasing and the density mass of the world economy shifts toward the minimum z_{\min} of the support. In fact, imposing the equilibrium condition $\varepsilon_t(k) = \varepsilon$

$$\varepsilon = \beta z \frac{\alpha h(z)}{[1 + \alpha H(z)]} \quad (43)$$

then, given ε and for any relative factor intensity z , an increase of the strength β of spillovers must be offset by a decrease of the $\frac{\alpha h(z)}{[1 + \alpha H(z)]}$ ratio: for any z either its density $h(z)$ decreases or $H(z)$, the density mass concentrated below z , increases.

Intuitively, for a given ε , a high β amplifies the volume of accessed knowledge $S(z)$, which is bigger near the technological frontier since high- z countries have access to a larger portion of the WFD: this effect is countered by a shift of the density mass of the equilibrium WFD toward the lower bound of the support z_{\min} , which in turn, as (41) shows, shifts back following an increase in β .

Proceeding as in Eeckhout and Jovanovic (2002), we measure the impact of an increase in β on the dispersion of the equilibrium WFD as measured by the position of its percentile. Denoting with z_p the value of z that identifies the percentile p of the equilibrium distribution $H(z)$, i.e. that solves

$$1 - p = \int_{z_p}^1 h(z) dz \equiv \Phi\left(\frac{\varepsilon}{\beta}, z_p\right) \quad (44)$$

by the implicit function theorem

$$\frac{\partial z_p}{\partial(\varepsilon/\beta)} = -\frac{\Phi_1}{\Phi_2} \quad \text{for all } p \in [0, 1] \quad (45)$$

Since $\Phi_2 \equiv \frac{\partial \Phi}{\partial z_p} < 0$, the sign of (45) depends on $\Phi_1 \equiv \frac{\partial \Phi}{\partial(\varepsilon/\beta)}$. When copying is undirected this is given by

$$\begin{aligned} \Phi_1 &\equiv \frac{\partial \Phi}{\partial w} = \frac{\partial}{\partial w} \left(\frac{(1 + \alpha)}{\alpha} \int_{z_p}^1 w z^{w-1} dz \right) = \\ &= \frac{(1 + \alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1 + \alpha)}{\alpha} \int_{z_p}^1 w z^{w-1} \ln z dz = \end{aligned}$$

$$\begin{aligned}
 &= \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1+\alpha)}{\alpha} \left(\int_{z_p}^1 w z^{w-1} \ln z dz \right) = \\
 &= \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1+\alpha)}{\alpha} \left([z^w \ln z]_{z_p}^1 - \int_{z_p}^1 z^{w-1} \ln z dz \right) = \\
 &= -\frac{(1+\alpha)}{\alpha} z_p^w \ln z_p > 0 \tag{46}
 \end{aligned}$$

since $z_p < 1$.

It follows that, when copying is undirected, $\frac{\partial z_p}{\partial(\varepsilon/\beta)} > 0$ and cross-country inequality in factor intensity raises with the strength β of spillovers and lowers with the degree ε of increasing returns of the world economy. Notice also that as spillovers disappear the WFD converges to a degenerate point mass centered in $z = 1$, since $\lim_{\beta \rightarrow 0} z_m = 1$. Since, by proposition 1, the higher ε the higher the equilibrium growth rate of the world economy, there is also a clear relationship between growth and inequality over the WFD.

Proposition 3

When copying is undirected, an increase in the intensity of knowledge spillovers entails a rise in the dispersion the equilibrium WFD, since every percentile of the distribution is shifted back ($\frac{\partial z_p}{\partial \beta} < 0$) and its support enlarges ($\frac{\partial z_m}{\partial \beta} < 0$). As spillovers disappear, convergence to a unique world factor intensity is obtained. An increase in the degree of the increasing returns and in the equilibrium growth rate of the world economy entails a reduction in the dispersion of the equilibrium WFD ($\frac{\partial z_p}{\partial \varepsilon} > 0$) and shrinks its support ($\frac{\partial z_m}{\partial \varepsilon} > 0$). Thus, growth and inequality over the WFD are negatively related: the higher the equilibrium growth rate of the world economy the lesser the dispersion of the cross-country distribution of factor intensity.

This results shed a new light on those obtained by Eeckhout and Jovanovic (2002) for the forward spillovers version of the model and show that the inequality-augmenting effect of knowledge spillovers does not depend on their direction but on their mere existence: even an increase in the intensity of backward spillovers, that raises the incentive to gain rank along the relative factor intensity distribution and that could have been supposed to act as an equalizing force, entails a rise of the dispersion of the equilibrium WFD, needed in order to equalize $\varepsilon_t(k)$ across countries. Moreover with backward spillovers growth and inequality over the WFD are negatively related, while the reverse is true with forward spillovers since ε is always negative and a rise of its absolute value lowers the equilibrium growth rate of the world economy.

The final step consists in calculating the equilibrium distribution of productivity and per capita incomes. In an equilibrium with undirected copying

$$A_t(k) = S(k)^\beta G_t^{1-\beta} \equiv \left[1 + \alpha \int_{z_{\min}}^{k/K_t} h(z) dz \right]^\beta G_t^{1-\beta} = [1 + \alpha H(z)]^\beta G_t^{1-\beta} \quad (47)$$

and using (41) and (42)

$$H(z) = \frac{\varepsilon(1+\alpha)}{\beta\alpha} \int_{[1/(1+\alpha)]^{\frac{\beta}{\varepsilon}}}^z s^{-(1-\frac{\varepsilon}{\beta})} ds = \frac{(1+\alpha)z^{\frac{\varepsilon}{\beta}} - 1}{\alpha} \quad (48)$$

so that

$$A_t(k) = G_t^{1-\beta} (1+\alpha)^\beta z^\varepsilon \quad (49)$$

The world productivity distribution (WPD) shifts over time due to exogenous common technological change G_t , but the distribution of relative productivity $a(z) = A_t(z)/A_t(1)$ is time-invariant since

$$a(z) = z^\varepsilon \quad (50)$$

where a is defined over $[a_{\min}, 1]$ and $a_{\min} = z_{\min}^\varepsilon = \left(\frac{1}{1+\alpha}\right)^\beta$.

Abusing notation and interpreting $H(z)$ as the distribution function of a random variable, it is possible to calculate the distribution function $H_A(a)$ of relative productivity in equilibrium noting that

$$\begin{aligned} H_A(a) &= \Pr\{a(z) \leq a\} = \Pr\{z^\varepsilon \leq a\} = \Pr\left\{z \leq a^{\frac{1}{\varepsilon}}\right\} = \\ &= H\left(a^{\frac{1}{\varepsilon}}\right) = \frac{(1+\alpha)a^{\frac{1}{\beta}} - 1}{\alpha} \end{aligned} \quad (51)$$

that implies the density

$$h_A(a) = H'_A(a) = \frac{(1+\alpha)a^{\frac{1}{\beta}-1}}{\beta\alpha} \quad a \in \left[\left(\frac{1}{1+\alpha}\right)^\beta, 1 \right] \quad (52)$$

Notice that, when copying is undirected and total knowledge spillovers are proportional to $H(z)$, the equilibrium distribution of relative productivity does not depend on the elasticity ε but only on the constant α and on the intensity β of spillovers: since an equilibrium WFD equalizes ε across countries in order to

eliminate incentives for a country to change its own relative factor intensity to raise the amount of accessed knowledge, a variation in ε entails a direct effect on the WFD, as Proposition 3 shows, which neutralizes its effect on the world productivity distribution (WPD).

Proceeding as in the case of the equilibrium WFD, the dispersion of the WPD can be shown to be increasing in the intensity of knowledge spillovers.

Proposition 4

When copying is undirected, an increase in the intensity of knowledge spillovers entails a rise in the dispersion the equilibrium WPD, since every percentile of the distribution is shifted back ($\frac{\partial a_p}{\partial \beta} < 0$) and its support enlarges ($\frac{\partial a_m}{\partial \beta} < 0$). An increase in the degree of the increasing returns and in the equilibrium growth rate of the world economy has no effect on the equilibrium WPD.

Per capita income relative to the technological leader, $y_r \equiv \frac{A_t(k) \cdot k}{A_t(K_t) \cdot K_t} = a(z)z$ is given in equilibrium by:

$$y_r = z^{1+\varepsilon} \tag{53}$$

where $y_r \in [y_r^{\min}, 1]$ and $y_r^{\min} = z_{\min}^{1+\varepsilon} = \left(\frac{1}{1+\alpha}\right)^{\frac{\beta(1+\varepsilon)}{\varepsilon}}$.

The equilibrium world income distribution is a transformation of the WFD that can be explicitly derived proceeding as in the case of the WPD, obtaining the density

$$h_Y(y_r) = \frac{\varepsilon \alpha}{\beta(1+\varepsilon)(1+\alpha)} y_r^{\frac{\varepsilon}{\beta(1+\varepsilon)} - 1} \quad y_r \in \left[\left(\frac{1}{1+\alpha}\right)^{\frac{\beta(1+\varepsilon)}{\varepsilon}}, 1 \right] \tag{54}$$

We measure inequality over the WID with the percentile approach applied above. Denoting with y_r^p the p -percentile of the equilibrium distribution of relative income and proceeding as in the WFD case

$$\frac{\partial y_r^p}{\partial \left(\frac{\varepsilon}{\beta(1+\varepsilon)}\right)} > 0 \quad \text{for all } p \in [0, 1) \tag{55}$$

and noting that $\frac{d}{d\varepsilon} \left(\frac{\varepsilon}{\beta(1+\varepsilon)}\right) > 0$, one obtain the following relationship between spillovers, growth and dispersion of the long run WID.

Proposition 5

When copying is undirected, an increase in the intensity of knowledge spillovers entails a rise in the dispersion the equilibrium WID, since every percentile of the distribution is shifted back ($\frac{\partial y_r^p}{\partial \beta} < 0$) and its support enlarges ($\frac{\partial y_r^{\min}}{\partial \beta} < 0$). As spillovers disappear, absolute convergence to a unique relative per capita income is obtained. An increase in the degree of the increasing returns and in the

equilibrium growth rate of the world economy entails a reduction in the dispersion of the equilibrium WFD ($\frac{\partial y^p}{\partial \varepsilon} > 0$) and shrinks its support ($\frac{\partial y^{\min}}{\partial \varepsilon} > 0$). Thus, growth and inequality over the WID are negatively related: the higher the equilibrium growth rate of the world economy the lesser the dispersion of the cross-country distribution of per capita income.

Thus, the existence of spillovers at the same time accelerates growth of the world economy, inducing increasing returns to scale, and creates a dispersion of long-run per capita income that otherwise would have not existed. On the other hand, the rate of growth of the world economy in equilibrium is negatively correlated with the dispersion of the WFD and the WID: the more concentrated the distribution of factor intensity, intuitively representing a situation in which many countries of the world economy operate similar technologies, the higher the elasticity of accessed knowledge, the degree of increasing returns and the growth rate. This seems to be consistent with empirical evidence documented by Sala-i-Martin (2006) on the reduction of various measures of inequality for the world income distribution over the period 1970-2000: in the same period the growth rate of the world economy has continuously increased.

4 Conclusion

We introduced a specification of the technology index in which knowledge spillovers can be either localized, coherently with the appropriate technology hypothesis, or backward directed, coherently with either the upgrading technology assumption or with the existence of barriers to technology adoption linked to relative factor intensity. With appropriate technology, knowledge spillovers between countries are likely to be localized, since only neighbouring technologies provide useful information that can be shared. As a consequence, the spillovers force depends on the local shape of the cross-country capital intensity distribution that, since total factor intensity is also an index of the technological state of an economy, can also be interpreted as the cross-country technology distribution, and the dynamics of the world income distribution now depends on the strength of these spillovers. The resulting set of equilibria for a world economy characterized by appropriate technology and local spillovers is very different from that of the Solow model, with a common world technology assumption, and that of the Lucas-Romer type, in which spillovers are measured by average total capital intensity: even if every country of the world economy shares the same preferences over the intertemporal consumption allocation, absolute convergence is not the only equilibrium outcome as in Solow, nor any WID is an equilibrium as in the Lucas-Romer specification. In general, even if convergence to a degenerate point mass distribution is an equilibrium, any cross-country technology distribution that clusters the world economy into long-run relative technology levels between which there are no spillovers is an equilibrium: thus the model provides an explanation for the emergence of convergence clubs, as a result of the localization of the spillovers force.

With backward knowledge spillovers the equilibrium WFD and WID depend on the strength of spillovers and on the degree of increasing returns attained in equilibrium by the world economy, which in turn depends on the dispersion of the WFD. Given the degree of increasing returns, an increase in the intensity of the spillover force raises inequality over the WFD and the WID to equalize returns to scale in the world economy. Given the intensity of spillovers, a less dispersed WFD in which many countries are near the technological frontier induce an acceleration in the growth rate of the world economy: in equilibrium growth and inequality are negatively related.

References

- Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, MIT Press, Cambridge (MA).
- Atkinson, A.B. and Stiglitz, J. (1969), "A New View of Technological Change", *Economic Journal*, LXXIX, 573-578.
- Basu, S. and Weil, D. (1998), "Appropriate Technology and Growth", *Quarterly Journal of Economics*, 113(4), 1025-54.
- Beine, M., Docquier, F. and Rapoport, H. (2001), "Brain Drain and Economic Growth: Theory and Evidence", *Journal of Development Economics*, 64(1), 275-289.
- Boldrin, M. and Levine, D. (2005), "The Economics of Ideas and Intellectual Property", *Proceedings of the National Academy of Sciences*, 102, 1252-1256.
- Caselli, F. (2005), "Accounting for Cross-Country Income Differences", *Handbook of Economic Growth vol. 1* (Aghion, P. and Durlauf, S. eds.), North-Holland.
- Docquier, F. and Marfouk, A. (2006), "International Migration by Educational Attainment, 1990-2000", *International Migration, Remittances, and the Brain Drain* (Ozden, C. and Schiff, M. eds.), Washington, DC: The World Bank and Palgrave MacMillan, pp. 151-200.
- Eeckhout, J. and Jovanovic, B. (1998), "Inequality", NBER Working Paper n. 6841.
- Eeckhout, J. and Jovanovic, B. (2002), "Knowledge Spillovers and Inequality", *American Economic Review*, 92(5), 1290-1307.
- Feyrer, J. (2003), "Convergence by Parts", mimeo, Dartmouth College.
- The Economist (2006), "Starbucks vs Ethiopia. Storm in a coffee cup", November 16.
- Gerosa, S. (2007), "Knowledge Spillovers and Productivity Differences", mimeo, Università di Roma-Tor Vergata.
- Gerschenkron, A. (1962), *Economic Backwardness in Historical Perspective*, Harvard University Press, Cambridge (MA).
- Hall, R. and Jones, C. (1999), "Why Do Some Countries Produce So Much More Output per Worker Than Others?", *Quarterly Journal of Economics*, 114(1), 83-116.
- Lucas, R. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.
- Lucas, R. (2001) "Some Macroeconomics for the 21st century", *Journal of Economic Perspectives*, 14(1), 159-168.
- Mankiw, N., Romer, D. and Weil, D. (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal Of Economics* 107, 407-437.
- Parente, S. and Prescott, E. (1994), "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 102 (2), 298-321.
- Quah, D. (1997), "Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs", *Journal of Economic Growth* 2(1) 27-59.
- Romer, P. (1986), "Increasing Returns and Long Time Growth", *Journal of Political Economy*, 94(5), 1002-1037.
- Sala-i-Martin, X. (2006), "The World Distribution of Income: Falling Poverty and... Convergence, Period", *Quarterly Journal of Economics*, 121 (2), 351-397.

Seager, A. (2006), "Starbucks, the coffee beans and the copyright row that cost Ethiopia £47m", *The Guardian*, October 26.

Young, A. (1995), "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience", *Quarterly Journal of Economics*, 110, 641-680.

Young, A. (2003), "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period", *Journal of Political Economy*, 111, 1220-1261.

Appendix A - Existence and Uniqueness of the Equilibrium

We have shown in the text that the economy optimal capital accumulation path chosen by the social planner over the infinite horizon must solve the functional equation

$$v_t(k) = \max_k \left\{ \left[1 - C \left(\frac{\tilde{k}}{k} \right) \right] A_t(k) \cdot k + \frac{1}{1+r} v_{t+1}(\tilde{k}) \right\} \quad (\text{A1})$$

We proceed as in Eeckhout and Jovanovic (2002), starting with the first-order and envelope conditions:

$$C' \left(\frac{\tilde{k}}{k} \right) A_t(k) = \frac{1}{1+r} v'_{t+1}(\tilde{k}) \quad (\text{A2})$$

$$\begin{aligned} v'_t(k) &= \left[1 - C \left(\frac{\tilde{k}}{k} \right) \right] A_t(k) + \left[1 - C \left(\frac{\tilde{k}}{k} \right) \right] A'_t(k) \cdot k + C' \left(\frac{\tilde{k}}{k} \right) \frac{\tilde{k}}{k} A_t(k) \\ &= A_t(k) \left[\frac{\tilde{k}}{k} C' + (1 - C) [1 + \varepsilon_t(k)] \right] \end{aligned} \quad (\text{A3})$$

where $\varepsilon_t(k) = \frac{A'_t(k) \cdot k}{A_t(k)}$ is the elasticity of the technology access function $A_t(k)$.

Updating one period the envelope condition (A3) and substituting in (A2) gives the second order difference equation:

$$\begin{aligned} (1+r) C' \left(\frac{k_{t+1}}{k_t} \right) A_t(k_t) &= A_{t+1}(k_{t+1}) \left\{ \frac{k_{t+2}}{k_{t+1}} C' \left(\frac{k_{t+2}}{k_{t+1}} \right) + \right. \\ &\left. + \left[1 - C \left(\frac{k_{t+2}}{k_{t+1}} \right) \right] [1 + \varepsilon_{t+1}(k_{t+1})] \right\} \end{aligned} \quad (\text{A4})$$

that characterizes optimal investment.

In a world equilibrium every country accumulates total capital at the same gross rate $\frac{k_{t+1}}{k_t} = x$ for each t , so that the cross-country factor distribution is time-invariant, i.e. $H_t(z) = H(z)$, and total factor productivity grows only through the exogenous increase in common knowledge

$$x_A \equiv \frac{A_{t+1}}{A_t} = (1+g)^{1-\beta} = \xi \quad (\text{A5})$$

Suppose also that $H(z)$ is such that the elasticity of efficiency to capital is a constant:

$$\varepsilon_t(k) = \varepsilon \quad \text{for all } k \text{ and for all } t \quad (\text{A6})$$

Then (A4) becomes

$$\left[\frac{(1+r)}{\xi} - x \right] C'(x) = [1 - C(x)](1 + \varepsilon) \quad (\text{A7})$$

Denoting with $x_c \equiv \frac{c_{t+1}}{c_t}$ the gross growth rate of consumption, the FOC for optimal consumption, using the CRRA utility function (11), is

$$(1+r) = \frac{x_c^\gamma}{\rho} \quad (\text{A8})$$

Since consumption grows at the same rate of potential output, in such an equilibrium we have $x_c = x_A x = \xi x$ and it is possible to eliminate the interest rate factor $(1+r)$ from (A6) to get an expression in the two unknowns x and ε

$$\Psi(x) \equiv \frac{\left[\frac{(\xi x)^\gamma}{\xi \rho} - x \right] C'(x)}{1 - C(x)} = 1 + \varepsilon \quad (\text{A9})$$

Given ε , which is not determined explicitly in the model but that should be thought as the outcome of the dynamics of the WFD starting from an initial distribution $H_0(z)$ of relative factor intensities, an equilibrium is reached when

$$\Psi(x) = 1 + \varepsilon \quad (\text{A10})$$

holds.

We now show that, provided some restrictions on the cost function C and on the elasticity of marginal utility to consumption γ are verified, the equilibrium exist and is unique.

The first derivative of $\Psi(x)$ with respect to x is

$$\begin{aligned} \Psi'(x) &= \left[\frac{\gamma (\xi x)^{\gamma-1}}{\rho} - 1 \right] \frac{C'(x)}{1 - C(x)} + \\ &+ \left[\frac{(\xi x)^{\gamma-1}}{\rho} - 1 \right] x \frac{C''(x)[1 - C(x)] + [C'(x)]^2}{[1 - C(x)]^2} \end{aligned} \quad (\text{A11})$$

Equation (A1) shows that $1 - C(x) > 0$, otherwise the firm would sustain net losses. In equilibrium the consumption growth rate must not exceed the

interest rate, $x_c < (1+r)$, so that $\frac{(\xi x)^{\gamma-1}}{\rho} = \frac{(\xi x)^\gamma}{\rho \xi x} = \frac{x_c^\gamma}{\rho x_c} = \frac{(1+r)}{x_c} > 1$ and the second term of the RHS of (A11) is positive.

It follows that if

$$\gamma > \frac{x_c}{1+r} = \frac{\xi x}{1+r} \quad (\text{A12})$$

then also the first term of the RHS of (A11) is positive and $\Psi'(x) > 0$.

The lowest possible growth rate of total capital x is that associated with a null investment: since the depreciation rate is η , x must be greater than $1-\eta$ (notice that in this case the capital stock shrinks to zero in the long run). Then, for any given value of ε , for a solution to (A10) to exist the cost function $C(x)$ must verify

$$\left[\frac{\gamma [\xi(1-\eta)]^{\gamma-1}}{\rho} - 1 \right] \frac{(1-\eta)C'(1-\eta)}{1-C(1-\eta)} < 1 + \varepsilon \quad (\text{A13})$$

If (A12) and (A13) hold, the equation $\Psi(x) = 1 + \varepsilon$ has a unique solution for any given value of ε (Figure 3).

Since $C(x)$ is an increasing and convex function, there will be also a growth rate x^* for which $C(x^*) = 1$: then, since C is invertible, $x^* = C^{-1}(1)$ is the maximum sustainable equilibrium growth rate of total capital stock k , since for $x > x^*$ the firm sustains net losses all over its infinite horizon.

Notice that

$$\lim_{x \rightarrow x^*} \Psi(x) = \infty \quad (\text{A14})$$

so that x^* is reached only if $\varepsilon \rightarrow \infty$ (Figure 3).

The other upper bound for the equilibrium growth rate is given by the No-Ponzi game condition $x_c < (1+r)$, that restrict x to be lower than $\frac{(1+r)}{\xi}$ so that $x \in \left(1-\eta, \min \left[x^*, \frac{(1+r)}{\xi} \right] \right)$.

A possible specification for the cost function $C(x)$ is

$$C(x) = B [x - (1-\eta)]^\theta \quad B > 0, \quad \theta > 1 \quad (\text{A15})$$

where $C(1-\eta) = 0$, i.e. there is no cost associated with a null investment, and $C'(1-\eta) = 0$ so that (A13) holds for any $\varepsilon > -1$. In this case the maximum sustainable equilibrium growth rate of total capital stock k is given by $x^* = B^{-\frac{1}{\theta}} + 1 - \eta$.

Figure 1: Appropriate Technology Spillovers

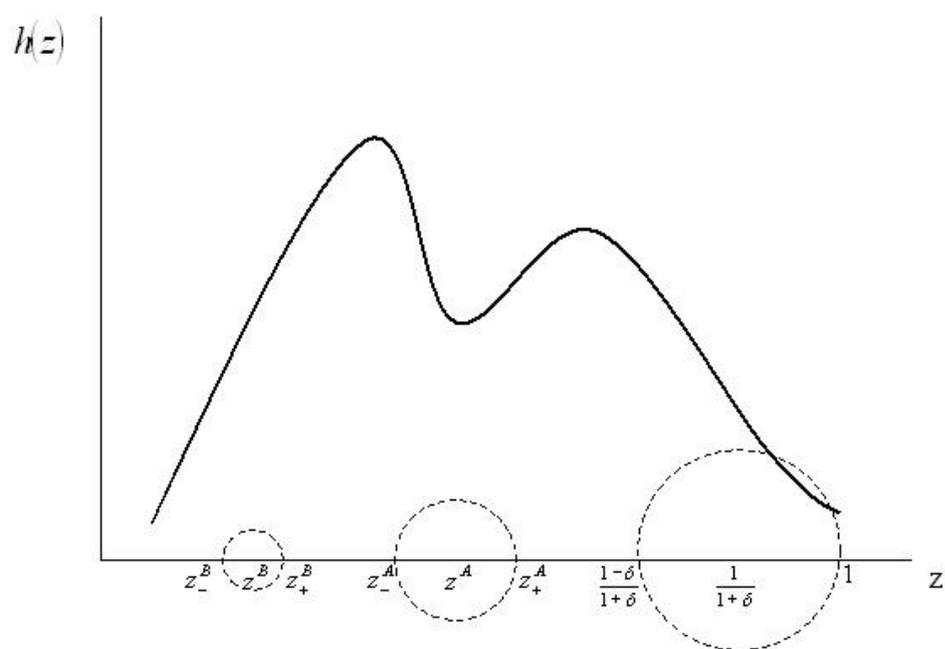


Figure 2: Equilibrium World Income Distribution with Appropriate Technology Spillovers

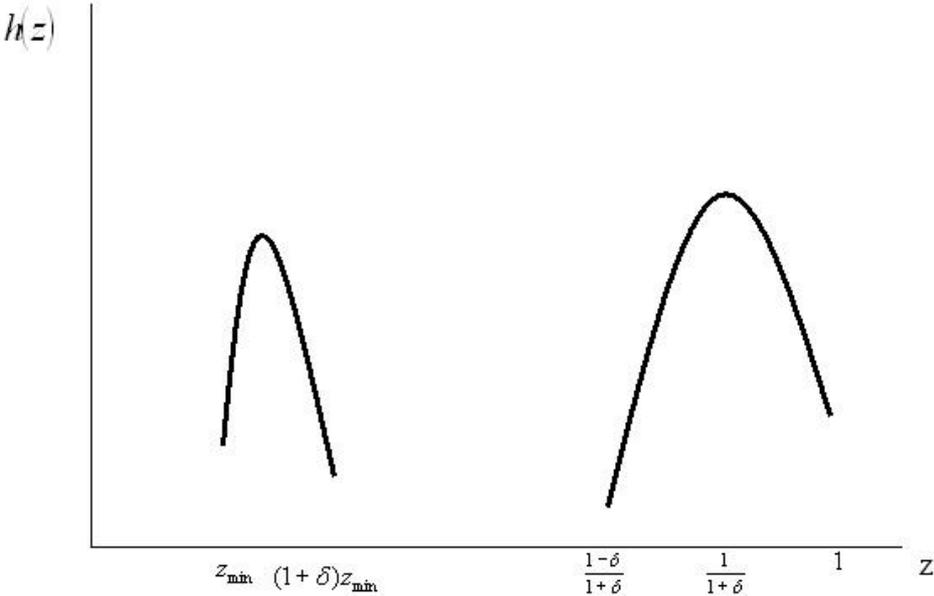


Figure 3: Equilibrium Growth of the World Economy

