

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: appropriate technology and backward knowledge spillovers. 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". In general, an equilibrium WID 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.

We find that both spillovers structures can explain over a half of observed cross-country productivity differences, but the backward spillovers one is more consistent with the observed world income distribution.

JEL Classifications: O40, O33, D31

Keywords: Appropriate Technology, Knowledge Spillovers, World Income Distribution, Inequality, Productivity Differences.

*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.

[†]Department of Economics, University of Rome "Tor Vergata", Via Columbia 2 00133 Roma. E-mail: stefano.gerosa@uniroma2.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.

We build a general equilibrium model of the world economy with cross-country knowledge flows, and we characterize the equilibrium or long-run WID for two possible spillovers structures: appropriate technology and backward knowledge spillovers.

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 empirical evidence documented by Hall and Jones(1999) and more recently by Caselli (2005) shows that low- k countries display also low TFP levels.

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 then examine another possible structure of technological spillovers, that reverts conventional wisdom about cross-country knowledge diffusion, analysing

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, 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 observed 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

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.

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 Sala-i-Martin (2006).

Finally, we perform a first test of the consistency of the knowledge spillovers structures with the observed pattern of cross-country productivity differences using cross-country data on income and factor endowments (physical and human capital) along two dimensions: their ability to explain static observed cross-country productivity differences at a point in time and their consistency with the shape of the observed world income distributions. We find that both knowledge spillovers structures are able to explain over an half of observed static productivity differences but we also show that backward spillovers are more successful in generating a theoretical WID close to the observed one.

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 presents the observed pattern of productivity differences, derives the static relationship between relative productivity and relative factor intensity at a point in time and calibrates the fundamental parameters of the model through regression analysis. Section 5 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

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

(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 backward 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.

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, while activating a new technology always requires an investment in physical-human capital.

It is also 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 portion of the actual distribution of world technologies $h_t(z)$.

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

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

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 $\varepsilon_t(k)$ 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

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

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

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:

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

and

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

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

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

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

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

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

which can be rewritten as

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

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

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

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 $h\left[\frac{(1-\delta)^2}{1+\delta}\right] = 0$.

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 $z \in \left[\frac{z_{\min}}{2}, \frac{1}{\delta} \right]$ and $z \notin Z$

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

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

$$\begin{aligned} \Phi_1 &\equiv \frac{\partial \Phi}{\partial w} = \frac{\partial}{\partial w} \left(\frac{(1+\alpha)}{\alpha} \int_{z_p}^1 w z^{w-1} dz \right) = \\ &= \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 w z^{w-1} \ln z dz = \\ &= \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1+\alpha)}{\alpha} \left(\int_{z_p}^1 w z^{w-1} \ln z dz \right) = \\ &= \frac{(1+\alpha)}{\alpha} \int_{z_p}^1 z^{w-1} dz + \frac{(1+\alpha)}{\alpha} \left([z^w \ln z]_{z_p}^1 - \int_{z_p}^1 z^{w-1} \ln z dz \right) = \\ &= -\frac{(1+\alpha)}{\alpha} z_p^w \ln z_p > 0 \end{aligned} \quad (46)$$

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

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

that implies the density

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

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

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

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 ($\frac{\partial y_r^p}{\partial \varepsilon} > 0$) and shrinks its support ($\frac{\partial y_r^{\min}}{\partial \varepsilon} > 0$). 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 Knowledge Spillovers and Productivity Differences

4.1 The WID and Productivity Differences: Empirical Evidence

The more direct way to assess the relative contribution of observable input and unobservable efficiency in shaping cross-country income differences is the development accounting technique, of which Klenow and Rodriguez (1997) and Hall and Jones (1999) are early examples and Caselli (2005) is a recent summary: by choosing a functional form for the production function and measuring inputs for a cross-section of country at a point in time, it is possible to extract total

factor productivity (TFP) as the residual part of income left unexplained by factors of production.

We follow Caselli (2005) and specify per-worker production as

$$y = Ak^\alpha h^{1-\alpha} \tag{56}$$

where k is the per worker capital stock, h is average human capital and A is efficiency or, in a wide sense, technology. With data on y , k and h , and given a value for the capital share α , efficiency or total factor productivity can be backed out and the structure of cross-country productivity differences can be analysed.

Data are taken from a variety of sources:

- y is real GDP per worker in PPP adjusted international dollars, and it is taken from version 6.1 of the Penn World Tables (PWT 6.1 - Heston, Summers, and Aten (2002)).
- aggregate capital stock K is taken from Caselli (2005) who calculate it using the perpetual inventory equation $K_t = I_t + (1 - \eta)K_{t-1}$, where I_t is investment and η is the depreciation rate of physical capital. I_t is measured from PWT 6.1 as real aggregate investment in PPP and η is set equal to 0.06. K is then divided by the number of workers, taken again from PWT 6.1, to finally obtain per worker capital stock k .
- h is average per worker human capital and is constructed as in Hall and Jones (1999). Since with competitive markets for factors (56) implies that the wage ω of a worker is such that $\ln \omega \propto \ln h$ and since the wage-schooling relationship is widely thought to be log-linear, then it is natural to specify $h = \exp\{\phi(s)\}$ where $\phi'(s)$ is the return to schooling estimated in a Mincerian regression of $\ln \omega$ on years of schooling s . Finally, since international data on education-wage profile documented in Psacharopoulos (1994) show a cross-country convexity across countries, with the return to an extra-year of schooling being higher in low-average schooling countries, $\phi(s)$ is specified as piecewise linear in s with slope 0.13 for $s \leq 4$, 0.10 for $4 < s \leq 8$, and 0.07 for $8 < s$, consistently with Psacharopoulos estimates for sub-saharan Africa, the world average and OECD countries. Data on average years of schooling for each country are taken from the Barro and Lee (2001) dataset.
- the capital share of GDP α is set equal to 0.3, roughly consistent with cross-country evidence documented in Gollin (2002): the mean labor share for a sample of 31 countries oscillates between 0.65 and 0.75, depending on the type of correction for the inclusion of the income of self-employed into the labor share of GDP.

There are 82 countries in our sample for which all relevant data are available for the year 1996. Table 1 list all the countries in the sample and presents across countries differences in income, factor intensity and productivity (TFP) resulting from a decomposition of (56) as in Hall and Jones (1999). We normalize everything with respect to the values of the country which displays the maximum

observed $k^\alpha h^{1-\alpha}$ (Norway): since factor intensity is as an index of the technological state of an economy and the volume of accessed knowledge depends on a country factor intensity *relative* to the technological leader, it is useful to observe the pattern of cross-country income and TFP from the point of view of relative factor intensity $z = k^\alpha h^{1-\alpha} / K_t$, where $K_t = \max \{ \text{observed } k^\alpha h^{1-\alpha} \text{ in } t = 1996 \}$.

Denoting, as in Caselli(2005), with $y_{kh} \equiv k^\alpha h^{1-\alpha}$ the component of income explained by observables (physical-human capital), one can evaluate the ability of the neoclassical-common world technology approach in explaining observed income differences. Extracting the TFP level A for each country using y_{kh} , it is possible to pursue a variance decomposition approach since

$$Var [\ln (y)] = Var [\ln (y_{kh})] + Var [\ln (A)] + 2Cov [\ln (A), \ln (y_{kh})] \quad (57)$$

and if technology does not differ across countries, then $Var [\ln (A)] = Cov [\ln (A), \ln (y_{kh})] = 0$, so that income dispersion should be fully explained by the dispersion of factors endowments.

One can judge the explanatory power of the neoclassical-common world technology approach by evaluating the ratio $Var [\ln (y_{kh})] / Var [\ln (y)]$, the fraction of observed income dispersion explained by differences in observable factor stocks. In our sample $Var [\ln (y_{kh})] / Var [\ln (y)] = 0.36$, meaning that less than 40% of observed static dispersion in per worker income is explained by factor endowments: cross-country productivity differences are large and they systematically amplify income differences produced by factor differences. Indeed the correlation between the observed TFP residual and y_{kh} is very high ($Corr [\ln (A), \ln (y_{kh})] = 0.7852$), meaning that high- y_{kh} countries display at the same time higher levels of technological efficiency.

This simple approach relates a moment, the variance, of two observed distribution, the world income distribution and the world factor distribution, but it is consistent with various possible shapes and local properties of both the WFD and the WID. As Quah (2007) convincingly argues, focusing on single moments of observed distributions of income, factor intensity or efficiency, or on conditional moment as in panel or cross-section regressions, might miss important static or dynamic features of those distribution. Collapsing in a single measure an entire distribution may be useful and in some cases appropriate, but it is uninformative of what is taking place in different subsets of the distribution support and may eventually conceal the existence of theoretically significant properties as multimodality or different degrees of polarization.

As an example, Caselli (2005) shows that the ratio $Var [\ln (y_{kh})] / Var [\ln (y)]$, which measures the success of growth models that rely exclusively on factor accumulation, does vary significantly across different subsets of the distribution: factor accumulation seems to play a major role in OECD countries where $Var [\ln (y_{kh})] / Var [\ln (y)] \simeq 0.6$, while in non-OECD countries $Var [\ln (y_{kh})] / Var [\ln (y)] \simeq 0.3$ and in general the ratio is higher for above the median income countries than for below the median income ones.

A simple and intuitive way to consider the global relationship between income and factor accumulation consists in comparing directly the two observed distribution, appropriately normalized.

We define

$$y_r = \frac{y}{y_{\max}} \quad (58)$$

and

$$z = \frac{y_{kh}}{y_{kh}^{\max}} \quad (59)$$

as respectively per capita income and relative factor intensity relative to the observed maximum, so that both variables take values in $[0, 1]$.

We then estimate the observed densities $h_y(y_r)$ and $h(z)$ using a Gaussian kernel and an "optimal" bandwidth following Silverman (1986) rule of thumb incorporated in Stata 8.2: Figure 1 shows the relationship between the observed WFD and WID³. Several features of this figure deserve comments. First, the 1996 WID displays the "twin peaks" property originally described by Quah (1993): an emergent bimodality that points to the possibility of the clustering of the world economy into a group of low-income and a group of high-income countries. Second, in this static representation of cross-country income differences, this bimodality seems to be driven by factor endowments since the observed WFD, the normalized distribution of y_{kh} across countries, shows the same twin peaks property of the WID: the low- z peak, around which is concentrated a large density mass of the world economy, corresponds to the low-income cluster, while the high- z peak corresponds to the (smaller) high-income cluster. Third, technological efficiency acts in a systematic way over the WFD to transform it into the WID: cross-country productivity differences give rise to the WID by stretching the WFD to the left, widening its support and shifting density mass back along it. This is just another representation of the failure of the common world technology assumption: if the world economy was characterized by a common technology, then the WID would simply mirror the WFD or, with z -independent technology shocks, would not systematically act on the WFD to deform it into the WID.

4.2 Testing Knowledge Spillovers Structures

We test both the appropriate technology and the technology improving knowledge spillovers structures, along two dimensions: their ability to account for observed static dispersion of productivity and their ability to reproduce the observed WID. We restrict the analysis to the case, extensively studied in Section 3, of undirected copying with $\alpha(z) = \alpha$: this restriction assigns exclusively to the WFD and to the free parameter β controlling the strength of spillovers the task of explaining observed productivity differences.

With appropriate technology spillovers, TFP at time t of a country i with relative factor intensity z is given by

³The unit of observation is the single country: we don't weight each data point by its population size and we neglect within country inequality. We see this as a natural choice if one thinks of the world economy as composed of different realizations of a model economy: population weighting assigns a disproportionate role to a few large economies in shaping the distribution of all relevant quantities.

$$A_i(z, t, \delta) = S(z, t, \delta)^\beta G_i(t)^{1-\beta} \equiv \left[\alpha \int_{\underline{\delta}z}^{\bar{\delta}z} h(s, t) ds \right]^\beta T(0) e^{\gamma t + \xi_i} \quad (60)$$

where $T(0) = G(0)^{1-\beta}$ and $G(0)$ being the initial level of the common part of technology, $\gamma = (1 - \beta)g$ being the common TFP growth factor and ξ_i a country specific technology shock reflecting differences in initial conditions due to other factors (e.g. geography or institutions). We also changed slightly the definition of TFP by removing the additive constant 1 from $S(z, t)$, so that we let the lower bound of accessed knowledge being freely determined.

Normalizing with respect to the technological leader, the country $i = L$ with $z = 1$, one obtains the relative TFP:

$$a_i(z, t) \equiv \frac{A_i(z, t, \delta)}{A_L(1, t, \delta)} = \frac{\int_{\underline{\delta}z}^{\bar{\delta}z} h(s, t) ds}{\int_{\underline{\delta}}^1 h(s, t) ds} e^{\xi_i - \xi_L} \quad (61)$$

and passing to logs

$$\ln a_i(z, t) = C + \beta \ln D_\delta(z) + \nu \quad (62)$$

where $C = -\beta \ln \left[\int_{\underline{\delta}}^1 h(s, t) ds \right]$, $D_\delta(z) = \left[\int_{\underline{\delta}z}^{\bar{\delta}z} h(s, t) ds \right]$ and $\nu = \xi_i - \xi_L$.

Then, for each choice of the amplitude δ of the neighbourhood over which spillovers flow, one can obtain an estimate of β by regressing observed relative TFP levels, extracted as residuals from (56), on the observed quantity $D_\delta(z) = H(\bar{\delta}z) - H(\underline{\delta}z)$, which is simply the density mass included in each technological neighbourhood given by the observed cumulative distribution of relative factor intensity $H(z)$. It is then possible to evaluate the ability of appropriate technology spillovers in explaining productivity differences by looking at, for each choice of δ , the explanatory power of the regression, that measures the fraction of observed dispersion in $a(z, t)$ accounted by a country position along the WFD, and by the implied predicted shape of the WID.

For each choice of δ we estimate the associated $\hat{\beta}$ and we calculate the theoretical or counterfactual income level relative to the technological leader for each country as

$$y_{R,\delta}^{th} = \frac{y_i^{th}(z, t, \delta)}{y_L^{th}(1, t, \delta)} = \left[\frac{\frac{\bar{\delta}z}{\hat{\delta}z}}{\frac{1}{\hat{\delta}}} \right]^{\hat{\beta}} z \quad (63)$$

and normalizing with respect to the maximum of $y_{R,\delta}^{th}$ (that may not coincide with $y_L^{th}(1, t, \delta)$), we obtain for each choice of the free parameter δ the theoretical world distribution of relative income levels $h_y^{th}(y_{r,\delta}^{th})$ that can be compared to the observed one $h_y(y_r)$. Theoretical or counterfactual relative income is the relative income generated by knowledge spillovers *in the absence of shocks*, that in our interpretation represent differences in initial levels of technological efficiency: that is, theoretical relative income is the level of income predicted by observed factor differences and cross-country knowledge diffusion as described by our specification of knowledge spillovers.

With backward spillovers, or with barriers to technology adoption measured by factor intensity according to the preferred interpretation, TFP of a country i with relative factor intensity z is given by

$$A_i(z, t) = S(z, t)^\beta G_i(t)^{1-\beta} \equiv \left[1 + \alpha \int_{z_m}^z h(s, t) ds \right]^\beta T(0) e^{\gamma t + \eta_i} \quad (64)$$

and TFP relative to the technological leader is

$$a_i(z, t) \equiv \frac{A_i(z, t)}{A_L(1, t)} = \left[\frac{1 + \alpha \int_{z_m}^z h(s, t) ds}{1 + \alpha} \right]^\beta e^{\xi_i - \xi_L} \quad (65)$$

since $\int_{z_m}^1 h(s, t) ds = 1$.

Supposing that $\alpha \gg 1$ and passing to logs, it is possible to approximate (65) by

$$\ln a_i(z, t) \simeq \beta \ln H(z) + \epsilon \quad (66)$$

where $H(z) = \int_{z_m}^z h(s, t) ds$ is simply the observed density mass of the world economy that lies behind country i along the technology ladder represented by the WFD, while $\epsilon = \xi_i - \xi_L$ is an error term.

It is possible to check the consistency of the assumption $\alpha \gg 1$, needed to identify β separately from α while keeping a specification of $A(z, t)$ in which the lower bound of $S(z, t)$ is non-zero, noting that the theoretical productivity relative to the technological leader of the less advanced country l with $z = z_m$ is

$$a(z_m, t) = \frac{A_l(z_m, t)}{A_L(1, t)} = \left[\frac{1}{1 + \alpha} \right]^\beta \quad (67)$$

In our sample the technological leader with $z = 1$ is Norway, while the country with the lowest total capital stock is Mozambique with $z_m = 0.091$ with an observed relative productivity $a_{obs}(z_m, t) = 0.26$. With an estimated $\hat{\beta}$ obtained from (66), it is possible to recover the normalizing constant $\hat{\alpha}$ by matching theoretical and observed $a(z_m, t)$, so that $\hat{\alpha}$ should be given by $\hat{\alpha} = \left(\frac{1}{0.26}\right)^{1/\hat{\beta}} - 1$, so that we can check the mutual consistency of $\hat{\beta}$ and of the assumption $\alpha \gg 1$.

With the estimates $\hat{\alpha}$ and $\hat{\beta}$ we then calculate the theoretical or counterfactual income level relative to the technological leader for each country as

$$y_r^{th} = \frac{y_i^{th}(z, t)}{y_L^{th}(1, t)} = \left[\frac{1 + \hat{\alpha} \int_0^z h(s, t) ds}{\frac{z_m}{1 + \hat{\alpha}}} \right]^{\hat{\beta}} z \quad (68)$$

which in this case simply coincide with income relative to the theoretical maximum. We can finally compare the theoretical distribution of relative income with backward spillovers $h_y^{th}(y_r^{th})$ with the observed one $h_y(y_r)$.

4.3 Estimations and Results

The two equations we estimate are equation (62) and (66). Data on productivity levels A (TFP) and relative factor intensity z are obtained as described above: absolute TFP levels are normalized with respect to the country with the maximum observed value of total factor intensity (Norway), to obtain relative productivity levels $a_i(z, t)$ for all 82 countries in the sample. $H(z)$ is simply taken to be the empirical distribution function of z , calculated by Stata 8.2 using the 82 observed data points, while $D_\delta(z) = H_\delta(\bar{\delta}z) - H_\delta(\underline{\delta}z)$ is constructed as follows:

- We let δ vary in $[0, 1]$ taking the possible values $\{0.1, 0.2, \dots, 0.9\}$.
- For each possible value of δ we add to the 82 observed values of z , the values $\bar{\delta}z$ and $\underline{\delta}z$ (obviously if $\underline{\delta}z < z_m$ or $\bar{\delta}z > 1$ we don't include those values).
- Finally we compute for each δ a new empirical distribution function $H_\delta(z)$, different from $H(z)$ since it is obtained through the inclusion of unobserved

data points, from which we compute $D_\delta(z) = H_\delta(\bar{\delta}z) - H_\delta(\underline{\delta}z)$ (obviously if $\underline{\delta}z < z_m$ then $H_\delta(\underline{\delta}z) = H_\delta(z_m)$ and if $\bar{\delta}z > 1$ then $H_\delta(\bar{\delta}z) = 1$).

The last assumption, formulated in order to justify OLS estimates of (62) and (66), is that the error terms ν and ϵ are uncorrelated with $H(z)$ and $D_\delta(z)$: initial differences in technology should not be correlated with the actual shape of the world distribution of factor intensity. Controlling for fixed country effects would require a panel data analysis, but here we are focusing on a first check of the consistency of our specification.

Table 2 presents the OLS regressions of (62) for each of the 9 possible values taken by δ :

- The intercept is always positive as predicted by (8), with the single exception for $\delta = 0.9$, but it is statistically significant only in 5 cases. The estimated $\hat{\beta}$ is always significant at the 1% level, but its magnitude varies between 0.47 and 0.97 with the choice of δ . The interpretation of $\hat{\beta}$ within the appropriate technology framework is relatively straight forward: given an amplitude δ of the technological neighbourhood over which spillovers flow, a 1% increase in the density mass of the world economy contained in $D_\delta(z)$ generates a $\hat{\beta}\%$ increase in the relative productivity of a country with relative factor intensity z : it follows that the estimated impact of a 1% increase in $D_\delta(z)$ varies between 0.47% and 0.97% with the choice of δ .
- The R^2 of the regressions, that measure the fraction of observed productivity differences explained by appropriate knowledge spillovers, varies with δ : in particular the R^2 seems to be U-shaped in δ , starting from 0.39 for $\delta = 0.1$, then decreasing monotonically and reaching a minimum of 0.21 for $\delta = 0.5$ and finally increasing monotonically toward its maximum value of 0.54 for $\delta = 0.9$.

Table 3 presents the OLS regression of the backward/technology improving specification of knowledge spillovers (66)

- The intercept is near zero as predicted by (66), even if not statistically significant. The estimated $\hat{\beta}$ is equal to 0.39 and it is highly significant: a 1% increase in the fraction of the world economies with relative capital intensity lower than z generates a 0.39% increase in the relative productivity of a country with relative factor intensity z . Here the linkage between factor accumulation and productivity is more direct than with appropriate technology spillovers, where the circular shape of the domain over which spillovers act introduces some ambiguity: by raising its own z and advancing its own technological state relative to the frontier through accumulation of physical and human capital, a country also raises the quantity of accessed knowledge and its own relative productivity. Finally the R^2 of the regression is 0.55: backward/technology improving spillovers (or knowledge spillovers with barriers measured by relative capital intensity) can explain more than half of the observed cross-country dispersion in TFP levels.

- The estimated $\hat{\beta}$ entails a value of $\hat{\alpha}$ consistent with the assumption $\alpha \gg 1$ used to obtain the regression equation (66): in fact $\hat{\alpha} = \left[\left(\frac{1}{0.26} \right)^{1/0.39} - 1 \right] \simeq 61$, so that $\hat{\beta}$ and the assumption $\alpha \gg 1$ are mutually consistent. We will use this value $\hat{\alpha}$, together with the estimated $\hat{\beta}$, in the computation of theoretical relative income levels given by (68).

An evaluation of the two knowledge spillovers structures based on their ability to explain static observed productivity differences is unable to discriminate between them: both the appropriate technology (with large enough technological neighborhoods, $\delta = 0.9$) and the backward spillovers frameworks are able to account for slightly more than an half of the observed dispersion in relative TFP, hence a significative fraction.

The second dimension over which we evaluate these representation of knowledge spillovers is their consistency with the actual shape of the WID: a conditional mean approach, like the one pursued above in the OLS regressions, identifies a single moment of the distribution of relative productivities and an identical predicted dispersion may translate in different implied shape of the distribution itself.

We use the OLS estimates $\hat{\beta}$ and observed values for $D_\delta(z)$ and $H(z)$ to calculate relative income levels with appropriate technology and backward spillovers, using respectively (63) and (68): Figures 5 and 6 display observed and theoretical kernel density estimations of the world income distribution, respectively for appropriate technology (for each value of δ) and backward spillovers.

It is evident that the appropriate technology framework fails in the replication of the observed WID for almost every choice of δ : predicted relative productivity differences are clearly too high and act on the observed WFD shifting density mass inconsistently with the actual shape of the WID. Even the common world technology hypothesis, that predicts that the WID should simply mirror the WFD, performs better, as Figure 1 shows. Only for $\delta = 0.9$ the predicted WID is similar to the observed one: the WFD is deformed by shifting mass backwards while keeping the original bimodality and its support is enlarged consistently with the observed WID.

backward knowledge spillovers seem to perform much better than AT spillovers in predicting the WID, even if the R^2 of the OLS regression of the two representations is almost identical when $\delta = 0.9$: the theoretical WID is almost indistinguishable from the observed one for the upper part of the support, while the two distributions slightly differ in the central and lower parts since the theoretical WID overpredicts the density mass of the world economy concentrated in the middle-income part.

To give a formal and quantitative meaning to the visual analysis of the "closeness" of the theoretical and observed WID, we perform a Kolmogorov-Smirnov test of the equality of the two distributions. The two-sample KS test is a nonparametric and distribution free test that assigns a probability distribution to the variable $\Delta = \sup_x |F_n(x) - G_m(x)|$ that measures the maximal distance between two empirical cumulative distributions $F_n(x)$ and $G_m(x)$ generated from unknown distributions F and G , where n and m are the number of observations in each sample: it is then possible to calculate explicitly a P-value

for a properly normalized Δ -statistics, and the hypothesis of the equality of the two distributions is rejected if P is "small". In general, it is possible to compute a threshold value for Δ under which the null hypothesis $F = G$ is accepted, but here we report simply the observed Δ and the P -value of the test: the higher the P -value, the closer the observed and theoretical WID.

Table 4 shows the results of the KS test of the equality of the observed WID $H_y(y_r)$ and the theoretical WID with appropriate technology, $H_y^{th}(y_r, \delta)$ given by (10) for each choice of δ , and backward spillovers, $H_y^{th}(y_r^{th})$ given by (68). We include also the test of the equality between the WID and the WFD, predicted by the common world technology hypothesis, as a benchmark for the performance of our specification of technology differences. The null hypothesis of the equality of the observed and theoretical WID is accepted at the 5% level only for three values of δ (0.1, 0.8 and 0.9) for the appropriate technology case, for the common world technology hypothesis WID=WFD and for backward spillovers: the highest P -value is obtained for the backward spillover specification (0.645), followed by the appropriate technology one with $\delta = 0.9$ (0.384).

5 Conclusion

We introduced a specification of the technology index in which knowledge spillovers can be either localized in a neighbourhood of a country current technological state, coherently with the appropriate technology hypothesis, or limited to the set of currently operated technologies (backward spillovers), 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, only countries operating similar technologies share useful technological information. Using aggregate capital intensity as an index of a country technological state, factor accumulation and productivity dynamics are linked. The resulting set of equilibria for a world economy characterized by appropriate technology 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 countries are technologically closer and information flows are higher, induce an acceleration in the growth rate of the world economy: in equilibrium growth and inequality are negatively related.

Calibrating the fundamental parameters using data on observed cross-country human and physical capital stocks, we finally show that both knowledge spillovers structures are able to explain over a half of observed static productivity differences. Evaluating the model's ability in replicating the observed WID, we show that backward spillovers are more successful in generating a theoretical WID close to the observed one.

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.
- 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.
- Nelson, R. and Phelps, E. (1966), "Investment in Humans, Technological Diffusion, and Economic Growth", *American Economic Review* 56, 69-75.
- 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.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London and New York.

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 $x^* = B^{-\frac{1}{\theta}} + 1 - \eta$.

Table 1 - Relative Factor Intensity, Relative Income and Relative Productivity

Country	Code	$k^\alpha h^{1-\alpha}$	y	A
Norway	NOR	1	1	1
United States	USA	0.9415	1.1389	1.2096
Switzerland	CHE	0.9186	0.8782	0.9559
Canada	CAN	0.8914	0.9011	1.0107
Sweden	SWE	0.8643	0.7981	0.9233
Australia	AUS	0.8476	0.9236	1.0896
Japan	JPN	0.8419	0.7550	0.8968
Finland	FIN	0.8390	0.7878	0.9390
Denmark	DNK	0.8374	0.8980	1.0723
New Zealand	NZL	0.8340	0.7472	0.8959
Belgium	BEL	0.8223	1.0064	1.2239
Austria	AUT	0.8071	0.9114	1.1292
Netherlands	NLD	0.8022	0.9137	1.1390
Hong Kong	HKG	0.8001	1.0278	1.2847
Republic of Korea	KOR	0.7923	0.6838	0.8631
France	FRA	0.7859	0.8981	1.1426
Singapore	SGP	0.7795	0.8584	1.1011
Israel	ISR	0.7784	0.8711	1.1190
Iceland	ISL	0.7523	0.7809	1.0380
United Kingdom	GBR	0.7287	0.8079	1.1087
Italy	ITA	0.7212	1.0156	1.4081
Ireland	IRL	0.7133	0.9542	1.3377
Greece	GRC	0.6961	0.6231	0.8950
Spain	ESP	0.6737	0.7763	1.1524
Cyprus	CYP	0.6698	0.6785	1.0130
Argentina	ARG	0.5863	0.5114	0.8723
Malaysia	MYS	0.5813	0.5194	0.8934
Barbados	BRB	0.5642	0.5957	1.0558

Table 1 (continued)

Name	Code	$k_i^\alpha h^{1-\alpha}$	y	A
Romania	ROM	0.5216	0.1996	0.3826
Chile	CHL	0.5163	0.4623	0.8954
Portugal	PRT	0.5097	0.5984	1.1738
Mexico	MEX	0.5031	0.4264	0.8476
Panama	PAN	0.4988	0.3045	0.6105
South Africa	ZAF	0.4928	0.4365	0.8857
Trinidad and Tobago	TTO	0.4749	0.4828	1.0166
Uruguay	URY	0.4617	0.4137	0.8948
Thailand	THA	0.4568	0.2661	0.5827
Venezuela	VEN	0.4496	0.3959	0.8805
Jordan	JOR	0.4294	0.3226	0.7512
Peru	PER	0.4294	0.2036	0.4743
Ecuador	ECU	0.4217	0.2518	0.5972
Mauritius	MUS	0.4155	0.5193	1.2497
Brazil	BRA	0.4132	0.3738	0.9047
Costa Rica	CRI	0.3983	0.2647	0.6646
Iran	IRN	0.3897	0.3565	0.9147
Botswana	BWA	0.3863	0.3588	0.9288
Turkey	TUR	0.3747	0.2948	0.7868
Algeria	DZA	0.3743	0.2994	0.7997
Philippines	PHL	0.3727	0.1551	0.4161
Guyana	GUY	0.3553	0.1549	0.4359
Syrian Arab Republic	SYR	0.3479	0.3216	0.9243
Tunisia	TUN	0.3479	0.3531	1.0148
Jamaica	JAM	0.3456	0.1529	0.4426
Paraguay	PRY	0.3431	0.2426	0.7069
Dominican Republic	DOM	0.3340	0.2487	0.7446
Colombia	COL	0.3245	0.2422	0.7464

Table 1 - (continued)

Name	Code	$k^\alpha h^{1-\alpha}$	y	A
Indonesia	IDN	0.3044	0.1729	0.6469
Sri Lanka	LKA	0.2934	0.1344	0.5218
Zimbabwe	ZWE	0.2870	0.1029	0.4086
El Salvador	SLV	0.2857	0.2370	0.9446
Nicaragua	NIC	0.2708	0.0995	0.4184
Honduras	HND	0.2696	0.1198	0.5062
Bolivia	BOL	0.2672	0.1170	0.4989
Lesotho	LSO	0.2521	0.0493	0.2228
Guatemala	GTM	0.2509	0.2343	1.063
Zambia	ZMB	0.2446	0.0437	0.2038
India	IND	0.2319	0.0946	0.4648
Papua New Guinea	PNG	0.2204	0.1306	0.6749
Pakistan	PAK	0.2144	0.1221	0.6487
Bangladesh	BGD	0.2000	0.1092	0.6217
Cameroon	CMR	0.1807	0.0671	0.4232
Kenya	KEN	0.1711	0.0453	0.3015
Ghana	GHA	0.1669	0.0465	0.3174
Togo	TGO	0.1506	0.0381	0.2885
Senegal	SEN	0.1491	0.0540	0.4128
Malawi	MWI	0.1425	0.0294	0.2349
Haiti	HTI	0.1339	0.0732	0.6225
Central African Republic	CAF	0.1338	0.0328	0.2796
Mali	MLI	0.1145	0.0295	0.2941
Niger	NER	0.1111	0.0288	0.2954
Uganda	UGA	0.0935	0.0307	0.3747
Mozambique	MOZ	0.0917	0.0305	0.3799

Table 2 - Relative Productivities and Appropriate Technology Knowledge Spillovers

Technological Neighbourhood Amplitude					
	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$	$\delta = 0.5$
Constant	0.9410196*** (0.1678781)	0.4680726* (0.1864839)	0.3651233** (0.1224665)	0.1304574 (0.0977648)	0.0877904 (0.086854)
$\ln D_\delta(z)$	0.5520611*** (0.0755041)	0.4732244*** (0.1102956)	0.5326067*** (0.089457)	0.4545924*** (0.0874277)	0.5315591*** (0.0935024)
R^2	0.3928	0.3025	0.2414	0.2119	0.2090
Obs.	82	82	82	82	82
	$\delta = 0.6$	$\delta = 0.7$	$\delta = 0.8$	$\delta = 0.9$	
Constant	0.1035439 (0.0927429)	0.0938162 (0.0742824)	0.3651233** (0.1224665)	-0.0849079* (0.0399052)	
$\ln D_\delta(z)$	0.7002127*** (0.1247104)	0.8833711*** (0.1255699)	0.9791215*** (0.1212171)	0.5988725*** (0.066256)	
R^2	0.2424	0.3283	0.4800	0.5446	
Obs.	82	82	82	82	

Note: OLS estimates of equation (62). Dependent variable is TFP relative to the technological leader (Norway). Robust standard errors in parentheses. *, ** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Table 3 - Relative Productivities and Backward/Technology Improving Knowledge Spillovers

Constant	0.0212248 (0.0475042)
$\ln H(z)$	0.3930646*** (0.0512411)
R^2	0.5534
Obs.	82

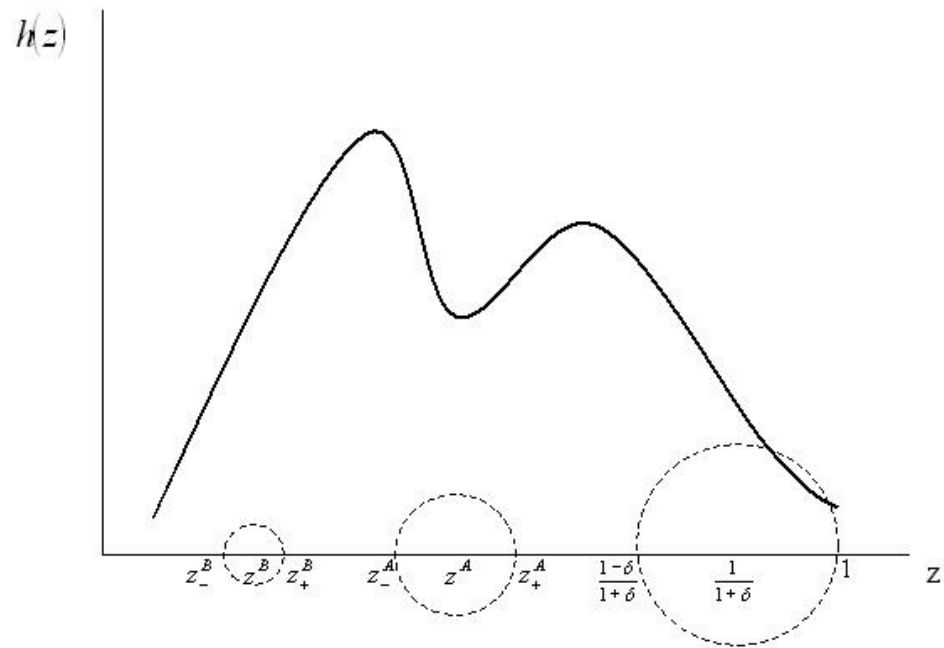
Note: OLS estimates of equation (66). Dependent variable is TFP relative to the technological leader (Norway). Robust standard errors in parentheses. *, ** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Table 4 - Kolmogorov-Smirnov Test for the Equality of Observed and Theoretical WID

	Technological Neighbourhood Amplitude				
	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$	$\delta = 0.5$
Δ	0.2195	0.2927	0.3171	0.3293	0.3415
<i>P - value</i>	0.026	0.001	0.001	0.000	0.000
	$\delta = 0.6$	$\delta = 0.7$	$\delta = 0.8$	$\delta = 0.9$	
Δ	0.3415	0.3049	0.2439	0.1341	
<i>P - value</i>	0.000	0.001	0.010	0.384	
	backward Spillovers	WFD			
Δ	0.1098	0.2561			
<i>P - value</i>	0.645	0.006			

Note: Kolmogorov-Smirnov test for the equality of the observed and theoretical WIDs with appropriate technology and backward spillovers. Δ is the maximum distance between the two distribution and *P - value* is the corrected P computed by Stata 8.2. WFD is the observed cross-country distribution of relative factor intensity $h(z)$, constructed as explained in the article. Sample of 82 countries for the year 1996.

Figure 1: Appropriate Technology Spillovers



Note: Kernel density estimation of per worker income (WID) taken from PWT 6.1 and physical-human capital aggregate $y_{kh} = k^\alpha h^{1-\alpha}$ (WFD), constructed as described in the main text, both relative to the observed maximum for the sample of 82 countries in the year 1996. Gaussian kernel and optimal bandwidth selected by Stata 8.2 in accord with Silverman (1986).

Figure 2: Equilibrium World Income Distribution with Appropriate Technology Spillovers

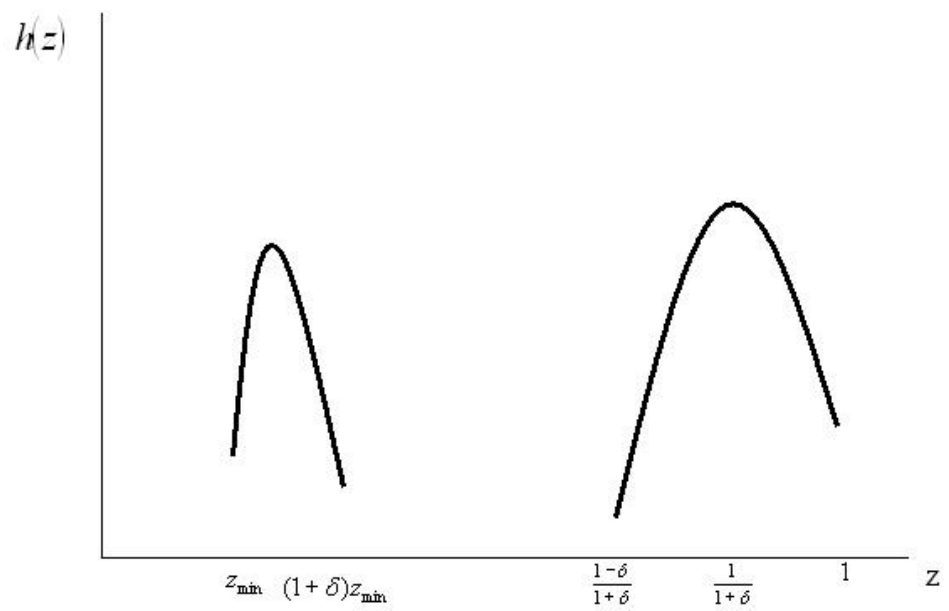


Figure 3: Equilibrium Growth of the World Economy

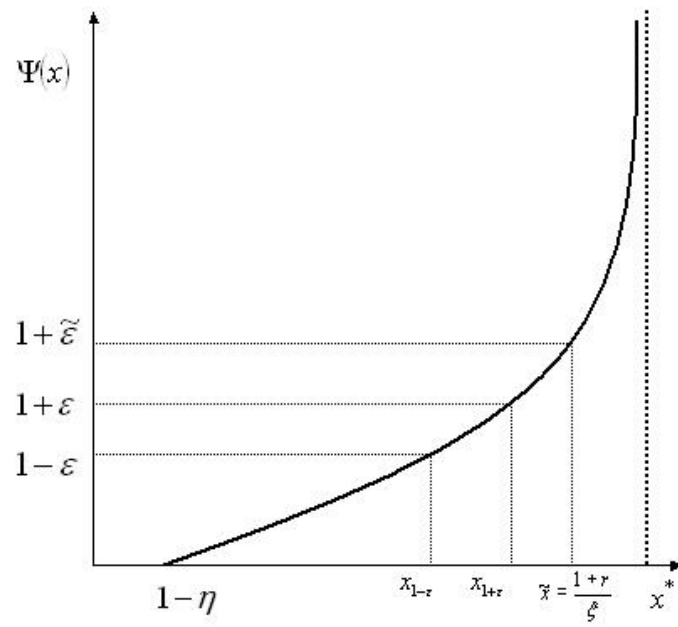


Figure 4: World Income Distribution and World Factor Distribution

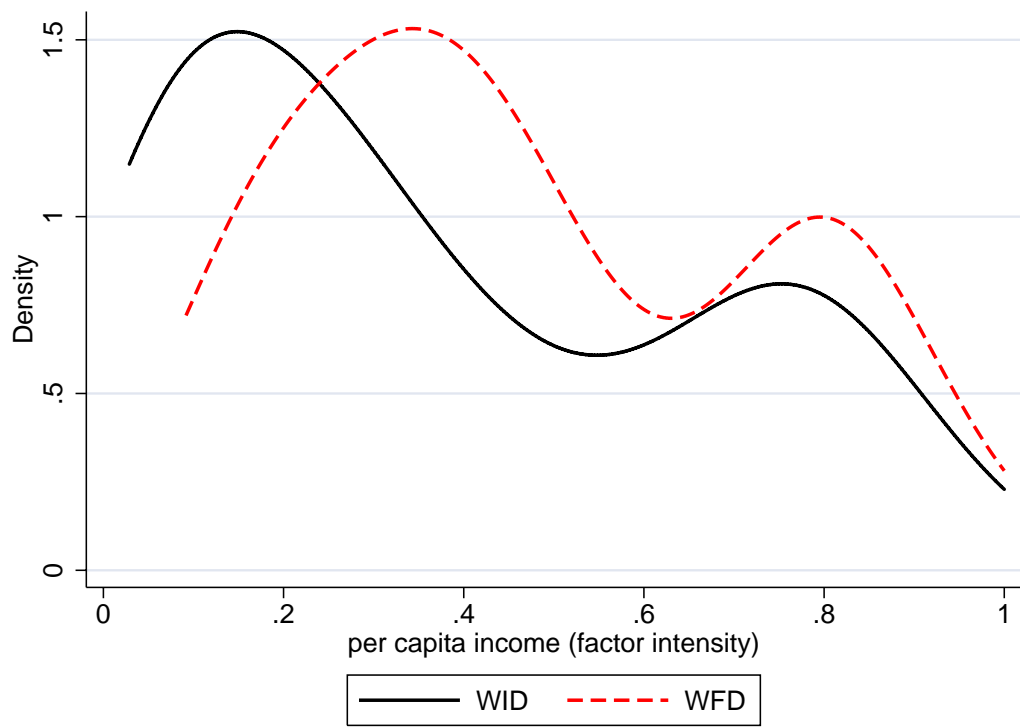


Figure 5 - Observed vs Theoretical WID with Appropriate Technology Knowledge Spillovers

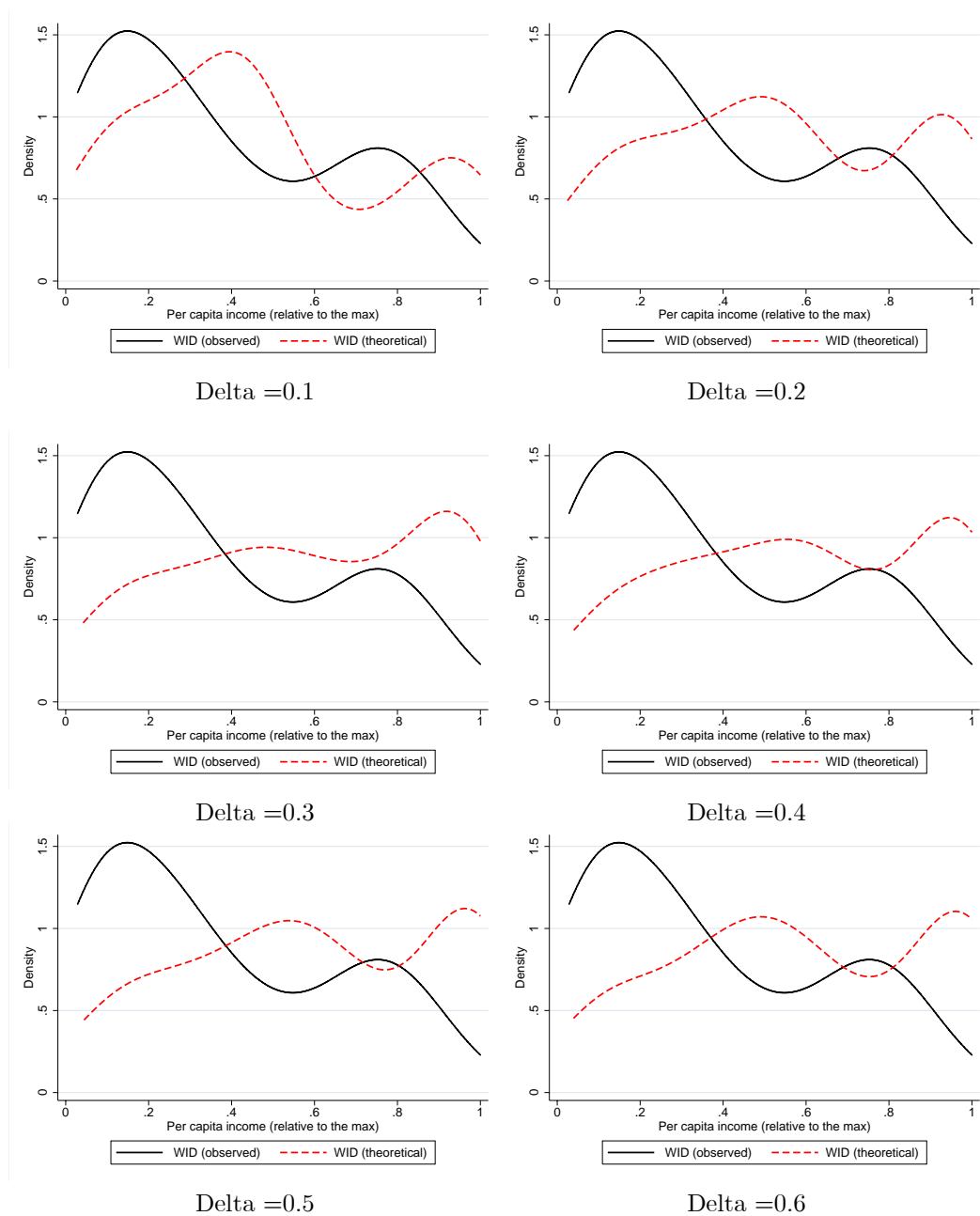
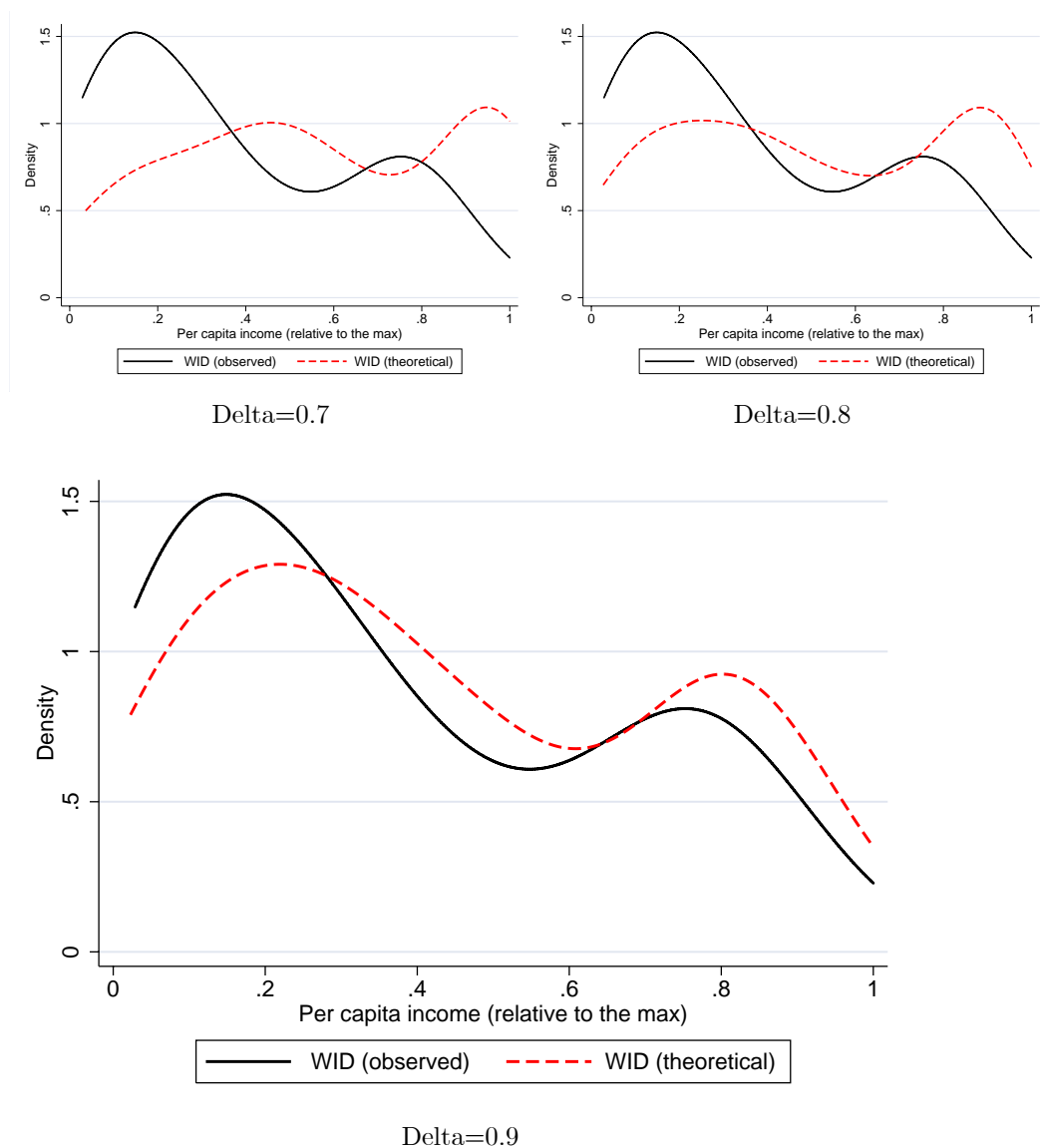
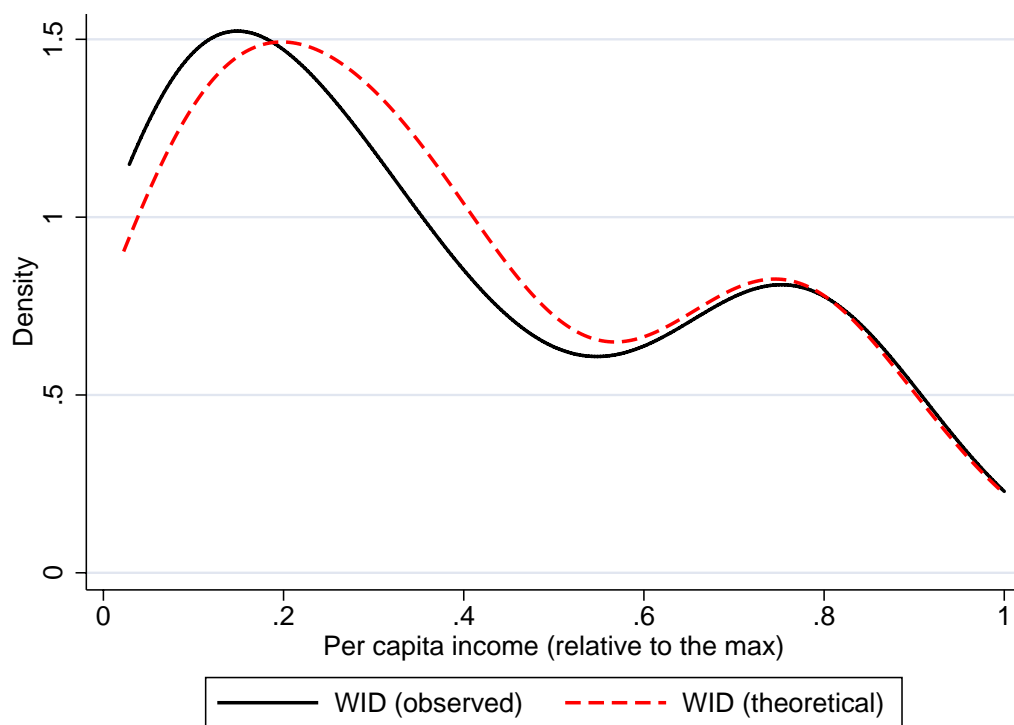


Figure 5 - Observed vs Theoretical WID with Appropriate Technology Knowledge Spillovers (continued)



Note: Kernel density estimation of observed per worker income (WID) taken from PWT 6.1 and theoretical per worker income, described by equation (63) in the main text, relative respectively to the observed and theoretical maximum for the sample of 82 countries in the year 1996 described in the main text. Gaussian kernel and optimal bandwidth selected by Stata 8.2 in accord with Silverman (1986).

Figure 6 - Observed and Theoretical WID with Backward Knowledge Spillovers



Note: Kernel density estimation of observed per worker income (WID) taken from PWT 6.1 and theoretical per worker income, described by equation (68) in the main text, relative respectively to the observed and theoretical maximum for the sample of 82 countries in the year 1996. Gaussian kernel and optimal bandwidth selected by Stata 8.2 in accord with Silverman (1986).