Technology and Inequality
Tesi di Dottorato in Economia Internazionale - XIX Ciclo

Stefano Gerosa

Coordinatore Dottorato: Prof. Giancarlo Marini

Supervisore tesi: Prof. Pasquale Scaramozzino

Università di Roma - "Tor Vergata"

Contents

Introduction

In economic modeling, the definition of a production technology is the formal specification of the relationship between a set of input and an associated production outcome. Empirical analysis imposes restrictions on the properties of the production technology: once observable inputs are defined and measured, the empirical consistency of the supposed technology can be tested with respect to its ability in reproducing observed time-series and cross-sectional behaviour of measurable outcomes. The assessment of the explanatory power of observable quantities in accounting for the observed dispersion of outcomes often shows the existence of a residual effect: the magnitude of such an effect then eventually calls for a theory transforming an unobservable force into an observable one.

This thesis is composed by three studies, whose common goal is advancing our knowledge of the properties of cross-country production technologies.

In the first part we focus on *across-country inequality*, tackling the issue of cross-country dispersion of incomes. The objective of growth theory is that of explaining the observed shape of the world income distribution (WID) and to eventually predict its future evolution. The existence of large cross-country productivity differences, measured by the residual dispersion of incomes left unexplained by the dispersion of observable quantities (physical and human capital), calls for the rejection of the hypothesis of a common world technology. We introduce a novel specification of the technology index, linking productivity to cross-country knowledge spillovers, that is empirically testable and has the potential to account for the observed pattern of productivity differences. We investigate two possible knowledge spillovers structures in a dynamic general equilibrium framework, and we characterize the equilibrium or long-run WID for each of them. We show that with *appropriate technology knowledge spillovers*, in which each country extracts useful knowledge only from countries operating similat technologies, the long-run WID is in general clustered and the world economy is splitted in distinct technological neighbourhoods, giving a possible explanation for the endogenous formation of convergence clubs. With *backward knowledge spillovers*, where the technology diffusion process is blocked by barriers to technology adoption measured by the aggregate capital intensity of an economy, the shape of the long-run WID is controlled by the strenght of the spillovers force and the degree of increasing returns of the world economy. We show that an increase of the spillovers force always amplifies the dispersion of the equilibrium WID and that growth and inequality are negatively related: the less dispersed the WID, the higher the equilibrium world rate of the world economy.

In the second part we analyse the pattern of cross-country productivity differences and we test the specification of the technology index introduced in the first part. In particular we test the knowledge spillovers structures introduced in the first part over 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 distribution. Using regression analysis to calibrate the fundamental parameters of our specification, we show that *both appropriate technology and backward spillovers can explain over half of the observed productivity differences*, but backward spillovers are more successful in replicating the actual shape of the WID.

In the third part we tackle the issue of *within-country inequality*, as measured by the skill premium, the wage of skilled workers relative to that of the unskilled. We study the ability of the capital-skill complementarity hypothesis (CSC), that assumes that capital substitutes unskilled labor more easily than skilled labor, to explain observed cross-country dispersion of the skill premium. We perform a steady-state analysis, novel to the literature about CSC, linking steady-state skill premia to the relative supply of unskilled labor and to observables that control the capital accumulation process (saving rates and barriers to capital accumulation, measured by the relative price of investments). We show that *CSC holds in non-OECD countries but not in the OECD subsample*, reinforcing a result obtained by other studies with different techniques: this result also show a fundamental cross-country parameter heterogeneity in the production function. As a by-product of our steady-state analysis we are also able to obtain new estimates for the elasticity of substitution between couples of inputs and to discriminate between alternative thresholds for the definition of skilled labor with respect to their consistentcy with plausible values of these elasticities. Finally, the fact that observable quantities are able to explain only a limited share of cross-country dispersion of skill.premia suggests that cross-country skill-biased technology differences are at work, and *capital accumulation alone cannot explain neither income differences nor cross-country differences in inequality.* If the analysis of cross-country income differences suggest that large international productivity differences are needed, the dispersion of cross-country skill premia reinforces this perspective and calls for a theory of productivity differences able to simultaneously explain growth and inequality facts.

Part 1

Knowledge Spillovers and the World Income Distribution

1. Introduction

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

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

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

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

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

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

(1995)(2003),using careful growth accounting analysis, to be cases of rapid factor accumulation rather than of exceptional productivity growth.

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

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

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

This argument can be applied by assuming that a country whose technological state is measured, as in Basu and Weil (1998), by the factor intensity k, which must be interpreted as an aggregate of human and physical capital, can only have access to the knowledge of countries that are in a neighborhood of k, simply because this is the only useful knowledge with respect to its technology level. As a result, the factor accumulation process and the dynamics of productivity are linked, and the growth experience of every country depends on the evolution of the entire cross-section distribution of factor intensity and not only 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 equilbria 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 equilbrium and the long run WID is in general clustered, its density mass concentrated over disconnected intervals of the support, delivering an explanation from the convergence club phenomenon.

We also examine another possible structure of technological spillovers, that reverts conventional wisdom about cross-country knowledge diffusion, analyzing the case in which spillovers are an increasing function of a country factor intensity: an economy can spill free knowledge only about technologies it has already developed, so that the more advanced a country the greater the information flow it can intercept. As Boldrin and Levine (2003) suggest, if the non rival character of ideas is a property of their immaterial nature, there exist also a cost associated with learning and implementing ideas into an actual production process: it is more likely that an advanced economy can integrate, freely or at near zero-cost, knowledge about inferior technologies into its framework rather than the contrary. Technological breakthroughs in nanotechnologies are unlikely to affect productivity in a developing country that lacks physical and human capital resources needed to adopt them, while often a technological advanced economy is able to extract useful knowledge even from traditional and low-tech sectors. A recent case is the one reported by the Economist (2006) on the conflict between Starbucks, the world's largest multinational chain of coffee shops, and Ethiopia, the poorest country in the world, over the recognition of intellectual property rights on three varieties of coffee-beans (Sidamo, Yirgacheffe and Harrar) originally developed in some ethiopian regions: Starbucks coffee shops sell Sidamo and Harrar coffees up to 27 $ a pound because of their exotic origin and because of the beans'speciality status, while "Ethiopian coffee farmers only earn between 0.58 and 1.16 $ for their crop, barely enough to cover the cost of production"[1].

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

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

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

This results shed a new light on those of Eeckhout and Jovanovic (2002): knowledge spillovers raise inequality not because of their direction, forward or backward along the distribution of factor intensity, but because of their mere existence. Moreover, we find that growth and inequality are negatively correlated: the more concentrated the distribution of relative factor intensity, the closer the technologies operated by countries interacting in the world economy, the higher the degree of increasing returns and the associated common growth rate of the world economy, consistently with empirical evidence documented by Sala-i-Martin (2006).

The rest of the paper is organized as follows. Section 2 presents the model, introduces the two possible structures of knowledge spillovers and characterizes optimal growth of the world economy. Section 3 studies the equilibrium WFD

[1]Tadesse 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.

and WID with appropriate and backward knowledge spillovers and analyses the relationship between spillovers, growth and inequality. Section 4 concludes.

2. The Model

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

The production function in per capita terms is given by

$$y_t = F_t(k) = A_t(k) \cdot k \tag{2.1}$$

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

$$A_t(k) = S_t(k)^{\beta} G_t^{1-\beta} \equiv \left[1 + \int_D \alpha(\frac{k}{K_t}) h_t(\frac{k}{K_t}) dz \right]^{\beta} G_t^{1-\beta} \qquad \text{for all } k \in [k_{\min}, K_t] \tag{2.2}$$

where G_t is an efficiency index which grows at the costant 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 (2.2) 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 tought as general knowledge, and a knowledge spillovers component S_t that comes from cross-country interactions. The knowledge spillovers part of (2.2) 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 $h_t(z)$: infact, 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

$$A_t^{'}(k) = 0 \qquad\qquad k \in [k_{\min}, K_t] \tag{2.3}$$

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

$$\begin{aligned} A_t^{'}(k) &= -\frac{\beta}{K_t} S_t(k)^{\beta-1} G_t^{1-\beta} \alpha\left(\frac{k}{K_t}\right) h_t\left(\frac{k}{K_t}\right) \\ &= -\frac{\beta}{K_t}\left(\frac{G_t}{S_t(k)}\right)^{1-\beta} \alpha(z)\, h_t(z) \leq 0 \end{aligned} \tag{2.4}$$

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

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

The appropriate technology assumption can be introduced in the model by modifying the specification of the technology index $A_t(k)$: a country total factor intensity k is also an index of its technological level, so that it can grab useful knowledge only from neighbouring countries along the world factor distribution , the ones which are "technologically close". One can let $D = [\underline{\delta}(k/K_t), \overline{\delta}(k/K_t)]$, where $\overline{\delta} = (1+\delta)$ and $\underline{\delta} = (1-\delta)$ so that

$$A_t(k) = \begin{cases} G_t^{1-\beta}\left[1 + \int_{z_{\min}}^{\overline{\delta}(k/K_t)} \alpha(z) h_t(z) dz\right]^{\beta} & k \in \left[k_{\min}, \frac{k_{\min}}{\underline{\delta}}\right] \\ G_t^{1-\beta}\left[1 + \int_{\underline{\delta}(k/K_t)}^{\overline{\delta}(k/K_t)} \alpha(z) h_t(z) dz\right]^{\beta} & k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\overline{\delta}}\right] \\ G_t^{1-\beta}\left[1 + \int_{\underline{\delta}(k/K_t)}^{1} \alpha(z) h_t(z) dz\right]^{\beta} & k \in \left[\frac{K_t}{\overline{\delta}}, K_t\right] \end{cases} \tag{2.5}$$

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 $\delta \to 0$ 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 $\delta \to \frac{1-z_{\min}}{z_{\min}}$ then $D \to [z_{\min}, 1]$ and the model converges to the Lucas-Romer setting.

Notice that in (2.5) 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 (2.5) has the potential to explain cross-country variation in both per capita income and TFP *through* cross-country variation in factor intensity: in the second part of this thesis we provide evidence supporting this specification.

The immediate consequence of the localisation 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

$$A_t^{'}(k) = \begin{cases} \frac{\beta}{K_t}\left(\frac{G_t}{S_t(k)}\right)^{1-\beta}\overline{\delta}\alpha\left(z_+\right)h_t\left(z_+\right) \geq 0 & k \in \left[k_{\min}, \frac{k_{\min}}{\underline{\delta}}\right] \\ \frac{\beta}{K_t}\left(\frac{G_t}{S_t(k)}\right)^{1-\beta}\left[\overline{\delta}\alpha\left(z_+\right)h_t\left(z_+\right) - \underline{\delta}\alpha\left(z_-\right)h_t\left(z_-\right)\right] & k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\overline{\delta}}\right] \\ -\frac{\beta}{K_t}\left(\frac{G_t}{S_t(k)}\right)^{1-\beta}\underline{\delta}\alpha\left(z_-\right)h_t\left(z_-\right) \leq 0 & k \in \left[\frac{K_t}{\overline{\delta}}, K_t\right] \end{cases} \tag{2.6}$$

where $z_+ = \overline{\delta}z$ 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 $k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\overline{\delta}}\right]$, 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 $h_t\left(z_+\right)$ and $h_t\left(z_-\right)$, 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 $D = [\frac{k_{\min}}{K_t}, \frac{k}{K_t}] \equiv [z_{\min}, z]$ so that the technology index becomes:

$$A_t(k) = S_t(k)^\beta G_t^{1-\beta} \equiv \left[1 + \int_{z_{\min}}^{k/K_t} \alpha(z) h_t(z) dz\right]^\beta G_t^{1-\beta} \qquad \text{for all } k \in [k_{\min}, K_t] \tag{2.7}$$

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

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

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

$$A_t^{'}(k) = \frac{\beta}{K_t} \left(\frac{G_t}{S_t(k)}\right)^{1-\beta} \alpha\left(z\right) h_t\left(z\right) \geq 0 \qquad \text{for all } k \in [k_{\min}, K_t] \tag{2.8}$$

By raising its own aggregate capital intensity a country also increases accessed knowledge since it can extract useful knowledge from a greater part of $h_t\left(z\right)$. In chapter 2 of this thesis we show that also (2.7) is consistent with empirical evidence.

Given (2.1), 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 form other countries; in particular

$$F_t^{'}\left(k\right) = A_t^{'} \cdot k + A_t\left(k\right) = A_t\left(k\right)\left[1 + \varepsilon_t\left(k\right)\right] \tag{2.9}$$

where $\varepsilon_t\left(k\right) = \frac{A_t^{'}(k) \cdot k}{A_t(k)}$ is the elasticity of the technology access function $A_t\left(k\right)$, 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 $\varepsilon_t\left(k\right)$ among countries (from (2.9) is also evident that $\varepsilon_t\left(k\right)$ must be greater than -1, ensuring that the marginal productivity of capital is not negative).

With forward spillovers *à* la Eeckhout-Jovanovic one has

$$\varepsilon_t(k) = -\beta z \frac{\alpha(z)\, h_t(z)}{S_t(k)} \tag{2.10}$$

which measure the percentage loss associated with a one per cent increase in k: in this case (2.9) 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

$$\varepsilon_t(k) = \begin{cases} \frac{\beta z}{S_t(k)} z\overline{\delta}\alpha(z_+)\, h_t(z_+) \geq 0 & k \in \left[k_{\min}, \frac{k_{\min}}{\underline{\delta}}\right] \\ \frac{\beta z}{S_t(k)} \left[\overline{\delta}\alpha(z_+)\, h_t(z_+) - \underline{\delta}\alpha(z-)\, h_t(z_-)\right] & k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\overline{\delta}}\right] \\ -\frac{\beta z}{S_t(k)} \underline{\delta}\alpha(z_-)\, h_t(z_-) \leq 0 & k \in \left[\frac{K_t}{\overline{\delta}}, K_t\right] \end{cases} \tag{2.11}$$

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

$$\varepsilon_t(k) = \beta z \frac{\alpha(z)\, h_t(z)}{S_t(k)} \tag{2.12}$$

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 sigle 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 $\widetilde{k}$ 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 (2.1): 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

$$v_t(k) = \max_{\tilde{k}} \left\{ \left[1 - C\left(\frac{\tilde{k}}{k}\right)\right] A_t(k) \cdot k + \frac{1}{1+r} v_{t+1}\left(\tilde{k}\right) \right\} \tag{2.13}$$

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

$$\begin{aligned}(1+r) C'\left(\frac{k_{t+1}}{k_t}\right) A_t(k_t) &= A_{t+1}(k_{t+1}) \left\{ \frac{k_{t+2}}{k_{t+1}} C'\left(\frac{k_{t+2}}{k_{t+1}}\right) + \right. \\ &\quad \left. + \left[1 - C\left(\frac{k_{t+2}}{k_{t+1}}\right)\right] \left[1 + \varepsilon_{t+1}(k_{t+1})\right] \right\}\end{aligned} \tag{2.14}$$

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 (2.2).

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

$$U = \sum_{t=0}^{\infty} \rho^t \frac{c_t^{1-\gamma} - 1}{1-\gamma} \qquad \rho < 1 \quad \gamma > 0 \tag{2.15}$$

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 (2.14) 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 (2.1).

The crucial assumption of this general equilibrium analysis is

$$\varepsilon_t(k) = \varepsilon \qquad \text{for all } k \text{ and for all } t \tag{2.16}$$

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 (2.16) holds, (2.14) reduces to an equation in the single variable $x_k = x$

$$\Psi(x) \equiv \frac{\left[\frac{(\xi x)^\gamma}{\xi\rho} - x\right] C'(x)}{1 - C(x)} = 1 + \varepsilon \tag{2.17}$$

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 $C(x)=B\left[x-(1-\eta)\right]^{\theta}$, 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 the two restrictions of the No-Ponzi game condition and of the non-negativity of net production.

Notice that:

- Since with forward spillovers model $\varepsilon_t(k)$ is always negative, the equlibrium 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 technolgy 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 (2.1) and (2.2) 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 $H_0(z)$. In the forward spilovers specification for $A_t(k)$ the *direction* of the force represented by this elasticity is the same for every country in the support, even if its magnitude can be different for each firm along its transition path: 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 (2.5), 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 (2.11) 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 equilbrium 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 $h(z) : [z_{\min}, 1] \to \mathbb{R}$ *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 (2.11) 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 equilbrium $\varepsilon \leq 0$.

Can an equilibrium with $\varepsilon < 0$ exist? It seems that at the lower tail of the WFD there is a symmetric problem, because for $k \in \left[k_{\min}, \frac{k_{\min}}{\underline{\delta}}\right]$ 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 $k_{\min} \to 0$ (2.11) is reduced to a system of two equations

$$\frac{\beta z}{S_t(k)}\left[\overline{\delta}\alpha\left(z_+\right)h_t\left(z_+\right) - \underline{\delta}\alpha\left(z-\right)h_t\left(z_-\right)\right] = -\varepsilon \qquad k \in \left(0, \frac{K_t}{\overline{\delta}}\right] \tag{3.1}$$

$$-\frac{\beta z}{S_t(k)}\underline{\delta}\alpha\left(z_-\right)h_t\left(z_-\right) = -\varepsilon \qquad k \in \left[\frac{K_t}{\overline{\delta}}, K_t\right] \tag{3.2}$$

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 (2.16) holds we have:

$$\frac{A_t^{'}\left(k\right)}{A_t\left(k\right)} = \frac{\varepsilon}{k} \to A_t\left(k\right) = \lambda_t k^{\varepsilon} \tag{3.3}$$

for some sequence of costants of integration $\{\lambda_t\}$.
Substituting from (2.5) gives:

$$G_t^{1-\beta}\left[(1 + \int_{\underline{\delta}(k/K_t)}^{\overline{\delta}(k/K_t)} \alpha(z)h(z)dz)\right]^{\beta} = \lambda_t k^{\varepsilon} \qquad k \in \left(0, \frac{K_t}{\overline{\delta}}\right] \tag{3.4}$$

and

$$G_t^{1-\beta}\left[(1 + \int_{\underline{\delta}(k/K_t)}^{1} \alpha(z)h(z)dz)\right]^{\beta} = \lambda_t k^{\varepsilon} \qquad k \in \left[\frac{K_t}{\overline{\delta}}, K_t\right] \tag{3.5}$$

and calculating the two sides of (3.4) and (3.5) for $k = \frac{K_t}{\overline{\delta}}$

$$\lambda_t = \left(\frac{K_t}{\overline{\delta}}\right)^{-\varepsilon} G_t^{1-\beta}\left[1 + \int_{\underline{\delta}/\overline{\delta}}^{1} \alpha(z)h(z)dz\right]^{\beta} \qquad k \in (0, K_t] \tag{3.6}$$

and eliminating λ_t from (3.4) and (3.5) leaves with an implicit characterization of $h(z)$ for the two parts of the support

$$1+\int_{z_-}^{z_+}\alpha(z)h(z)dz=\left(\overline{\delta}z\right)^{\frac{\varepsilon}{\beta}}\left[1+\int_{\underline{\delta}/\overline{\delta}}^{1}\alpha(z)h(z)dz\right]\qquad z\in\left(0,\frac{1}{1+\delta}\right] \tag{3.7}$$

$$1+\int_{z_-}^{1}\alpha(z)h(z)dz=\left(\overline{\delta}z\right)^{\frac{\varepsilon}{\beta}}\left[1+\int_{\underline{\delta}/\overline{\delta}}^{1}\alpha(z)h(z)dz\right]\qquad z\in\left[\frac{1}{1+\delta},1\right] \tag{3.8}$$

which can be rewritten as

$$[(1+\delta)\,z]^{\frac{\varepsilon}{\beta}}=\begin{cases}\dfrac{1+\int_{z_-}^{z_+}\alpha(z)h(z)dz}{1+\int_{\underline{\delta}/\overline{\delta}}^{1}\alpha(z)h(z)dz} & z\in\left(0,\frac{1}{1+\delta}\right]\\[2ex] \dfrac{1+\int_{z}^{1}\alpha(z)h(z)dz}{1+\int_{\underline{\delta}/\overline{\delta}}^{1}\alpha(z)h(z)dz} & z\in\left[\frac{1}{1+\delta},1\right]\end{cases} \tag{3.9}$$

But, while the left hand side of this equilibrium equation diverges as $z\rightarrow 0$ since $\varepsilon<0$, the right hand side of (3.9) 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 (3.9) follows from the fact that $\alpha(z)$ is bounded in $[z_{\min},1]$.

It follows that an equilibrium CFD 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 (2.11) the equilibrium condition $\varepsilon=0$ is verified at the boundaries of the support if and only if

$$h(z_+) = 0 \qquad z\in[z_{\min},\frac{z_{\min}}{1-\delta}] \tag{3.10}$$

$$h(z_-) = 0 \qquad z\in[\frac{1}{1+\delta},1] \tag{3.11}$$

so that

$$h(z) = 0 \qquad z \in [(1+\delta)z_{\min}, \frac{(1+\delta)}{(1-\delta)}z_{\min}] \tag{3.12}$$

$$h(z) = 0 \qquad z \in [\frac{(1-\delta)}{(1+\delta)}, 1-\delta] \tag{3.13}$$

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

$$(1+\delta)\,\alpha\,(z_+)\,h_t\,(z_+) = (1-\delta)\,\alpha\,(z-)\,h_t\,(z_-) \qquad k \in \left[\frac{k_{\min}}{\underline{\delta}}, \frac{K_t}{\overline{\delta}}\right] \tag{3.14}$$

so that in equilibrium

$$h_t\,(z_+) = 0 \Rightarrow h_t\,(z_-) = 0 \tag{3.15}$$

Since by definition $z_+ = \frac{(1+\delta)}{(1-\delta)}z_-$ (3.10) and (3.11) imply also[2]

$$h(z) = 0 \qquad z \in [\frac{(1+\delta)^2}{1-\delta}z_{\min}, \left(\frac{1+\delta}{1-\delta}\right)^2 z_{\min}] \tag{3.16}$$

$$h(z) = 0 \qquad z \in [\left(\frac{1-\delta}{1+\delta}\right)^2, \frac{(1-\delta)^2}{1+\delta}] \tag{3.17}$$

Iterating this passage one obtains

$$h\,(z) = 0 \text{ for } \begin{cases} z \in P_n = [\frac{(1+\delta)^n}{(1-\delta)^{n-1}}z_{\min}, \left(\frac{1+\delta}{1-\delta}\right)^n z_{\min}] \\ z \in Q_n = [\left(\frac{1-\delta}{1+\delta}\right)^n, \frac{(1-\delta)^n}{(1+\delta)^{n-1}}] \end{cases} \qquad n \in \mathbb{N} \tag{3.18}$$

so that

Proposition 2

The set of the equilibrium world factor distibutions for the world economy described by (1), (5) and (15) consists of all distributions $h(z)$ defined over intervals $[z_{\min}, 1]$, for some $z_{\min} \in [0,1]$, such that $h(z) = 0$ for $z \in \left(\bigcup_{n\in\mathbb{N}} P_n\right) \cup \left(\bigcup_{n\in\mathbb{N}} Q_n\right) \cap [z_{\min}, 1]$. The associated equilibrium world income distribution is then obtained through (1).

[2]One 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 (30) impose that also $h\left[\frac{(1+\delta)^2}{1-\delta}z_{\min}\right] = 0$.

Thus, the set of equilibria for a world economy characterized by the presence of local spillovers or appropriate technology is very different from that of a world economy in which knowledge spillovers are global, either because of a common world technology assumption à la Solow or because the technology index depends on the average factor intensity *à la* Romer-Lucas: even if preferences over consumption are identical for each country, absolute convergence in relative factor intensity and income levels is not the only equilibrium outcome, as in Solow, 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). 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 tendence 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 intensitiy 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 equilbrium 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 $h(z) : [z_{\min}, 1] \to \mathbb{R}$ *that verifies*

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

such that (16) and (17) hold.

Since (2.16) and (3.3) hold, for some sequence of costants of integration $\{\lambda_t\}$, substituting from (2.7) gives

$$G_t^{1-\beta}\left[1+\int_{z_{\min}}^{k/K_t}\alpha(z)h_t(z)dz\right]^{\beta}=\lambda_t k^{\varepsilon}\qquad\text{for all } k\in[k_{\min},K_t]\tag{3.19}$$

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

$$\lambda_t = G_t^{1-\beta}k_{\min}^{-\varepsilon}\tag{3.20}$$

and it is now possible to eliminate λ_t from (3.19)

$$1+\int_{z_{\min}}^{k/K_t}\alpha(z)h_t(z)dz=\left(\frac{k}{k_{\min}}\right)^{\frac{\varepsilon}{\beta}}\qquad\text{for all } k\in[k_{\min},K_t]\tag{3.21}$$

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

$$h(z)=\frac{\varepsilon}{\beta\alpha(z)z_{\min}^{\frac{\varepsilon}{\beta}}}z^{-\left(1-\frac{\varepsilon}{\beta}\right)}\tag{3.22}$$

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

$$\int_{z_{\min}}^{1}h(z)dz=\frac{\varepsilon}{\beta\alpha z_{\min}^{\frac{\varepsilon}{\beta}}}\int_{z_{\min}}^{1}z^{-\left(1-\frac{\varepsilon}{\beta}\right)}=1\tag{3.23}$$

from which

$$z_{\min}=\left(\frac{1}{1+\alpha}\right)^{\frac{\beta}{\varepsilon}}\tag{3.24}$$

and

$$h(z)=\frac{\varepsilon(1+\alpha)}{\beta\alpha}z^{-\left(1-\frac{\varepsilon}{\beta}\right)}\tag{3.25}$$

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. Infact, imposing the equilibrium condition $\varepsilon_t\left(k\right) = \varepsilon$

$$\varepsilon = \beta z \frac{\alpha h(z)}{[1+\alpha H(z)]} \tag{3.26}$$

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.

Intuitevely, for a given ε, a high β amplifies the volume of accessed knowledge S, 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 (3.24) shows, shifts back following an increase in β.

Proceding 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

$$1 - p = \int_{z_p}^{1} h(z)dz \equiv \Phi(\frac{\varepsilon}{\beta}, z_p) \tag{3.27}$$

by the implicit function theorem

$$\frac{\partial z_p}{\partial(\varepsilon/\beta)} = -\frac{\Phi_1}{\Phi_2} \quad \text{for all } p \in [0,1) \tag{3.28}$$

Since $\Phi_2 \equiv \frac{\partial \Phi}{\partial z_p} < 0$, the sign of (3.28) depends on $\Phi_1 \equiv \frac{\partial \Phi}{\partial(\varepsilon/\beta)}$. When copying is undirected this is given by

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

$$= \frac{(1+\alpha)}{\alpha}\int_{z_p}^{1} z^{w-1}dz + \frac{(1+\alpha)}{\alpha}\int_{z_p}^{1} w z^{w-1}\ln z dz =$$

$$= \frac{(1+\alpha)}{\alpha}\int_{z_p}^{1} z^{w-1}dz + \frac{(1+\alpha)}{\alpha}\left(\int_{z_p}^{1} w z^{w-1}\ln z dz\right) =$$

$$= \frac{(1+\alpha)}{\alpha}\int_{z_p}^{1} z^{w-1}dz + \frac{(1+\alpha)}{\alpha}\left([z^w \ln z]_{z_p}^{1} - \int_{z_p}^{1} z^{w-1}\ln z dz\right) =$$

$$= -\frac{(1+\alpha)}{\alpha} z_p^w \ln z_p > 0 \tag{3.29}$$

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 dissapear the WFD converges to a degenerate point mass centered in $z = 1$, since $\lim_{\beta \to 0} z_m = 1$. 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 productivities and per capita incomes. In an equilibrium with undirected copying

$$A_t(k) = S(k)^\beta G_t^{1-\beta} \equiv \left[1 + \alpha \int_{z_{\min}}^{k/K_t} h(z)dz \right]^\beta G_t^{1-\beta} = [1 + \alpha H(z)]^\beta G_t^{1-\beta} \tag{3.30}$$

and using (3.24) and (3.25)

$$H(z) = \frac{\varepsilon(1+\alpha)}{\beta\alpha} \int_{[1/(1+\alpha)]^{\frac{\beta}{\varepsilon}}}^{z} s^{-\left(1-\frac{\varepsilon}{\beta}\right)} ds = \frac{(1+\alpha)z^{\frac{\varepsilon}{\beta}} - 1}{\alpha} \tag{3.31}$$

so that

$$A_t(k) = G_t^{1-\beta} (1+\alpha)^\beta z^\varepsilon \tag{3.32}$$

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

$$a(z) = z^{\varepsilon} \tag{3.33}$$

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

$$H_A(a) = \Pr\{a(z) \leq a\} = \Pr\{z^{\varepsilon} \leq a\} = \Pr\left\{z \leq a^{\frac{1}{\varepsilon}}\right\} =$$

$$= H(a^{\frac{1}{\varepsilon}}) = \frac{(1+\alpha)a^{\frac{1}{\beta}} - 1}{\alpha} \tag{3.34}$$

that implies the density

$$h_A(a) = H_A^{'}(a) = \frac{(1+\alpha)a^{\frac{1}{\beta}-1}}{\beta\alpha} \qquad a \in \left[\left(\frac{1}{1+\alpha}\right)^{\beta}, 1\right] \tag{3.35}$$

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).

Proceding 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:

$$y_r = z^{1+\varepsilon} \tag{3.36}$$

where $y_r \in \left[y_r^{\min}, 1\right]$ 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 explictely derived proceding as in the case of the WPD, obtaining the density

$$h_Y(y_r) = \frac{\varepsilon\alpha}{\beta(1+\varepsilon)(1+\alpha)} y_r^{\frac{\varepsilon}{\beta(1+\varepsilon)}-1} \qquad y_r \in \left[\left(\frac{1}{1+\alpha}\right)^{\frac{\beta(1+\varepsilon)}{\varepsilon}}, 1\right] \tag{3.37}$$

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 proceding as in the WFD case

$$\frac{\partial y_r^p}{\partial\left(\frac{\varepsilon}{\beta(1+\varepsilon)}\right)} > 0 \quad \text{for all } p \in [0,1) \tag{3.38}$$

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, 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, intuitevely 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 continously increased.

4. Conclusion

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

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

5. 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 McMillan, 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.

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, Stratifcation, and Convergence Clubs", *Journal of Economic Growth* 2(1) 27-59.

Romer, P. (1986), "Increasing Returns and Long Time Growth", *Journal of Political Economy*, 94(5), 1002-1037.

Sala-i-Martin, X. (2006), "The World Distribution of Income: Falling Poverty and… Convergence, Period", *Quarterly Journal of Economics*, 121 (2), 351-397.

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

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

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

APPENDIX A

Existence and Uniqueness of the Equilibrium

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

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

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

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

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

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

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

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

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

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

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

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

Then (A4) becomes

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

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

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

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

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

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

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

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

$$\Psi'(x) = \left[\frac{\gamma (\xi x)^{\gamma-1}}{\rho} - 1\right] \frac{C'(x)}{1 - C(x)} + \tag{A11}$$

$$+ \left[\frac{(\xi x)^{\gamma-1}}{\rho} - 1\right] x \frac{C''(x)[1 - C(x)] + \left[C'(x)\right]^2}{[1 - C(x)]^2}$$

Equation (A1) shows that $1 - C(x) > 0$, otherwise the firm would sustain net losses. In an equlibrium 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

$$\gamma > \frac{x_c}{1+r} = \frac{\xi x}{1+r} \tag{A12}$$

then also the first term of the RHS of (A11) is positive, $\Psi'(x) > 0$ and the equation $\Psi(x) = 1 + \varepsilon$ has a unique solution for any given value of ε.

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

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

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

$$\lim_{x\to x^*}\Psi(x)=\infty \tag{A14}$$

so that x^* is reached only if $\varepsilon\to\infty$ (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

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

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

FIGURE 1. Appropriate Technology Spillovers

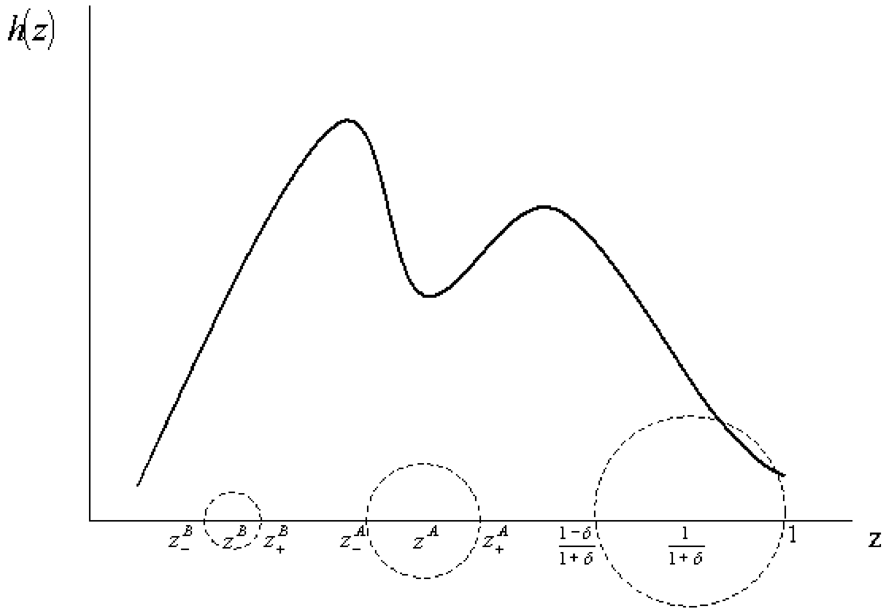


FIGURE 2. Equilibrium World Income Distribution with Appropriate Technology Spillovers

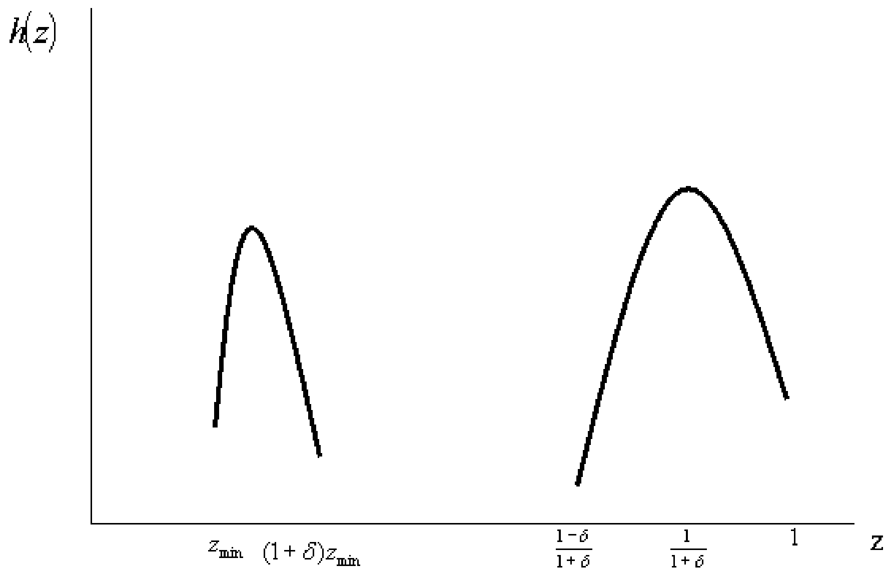
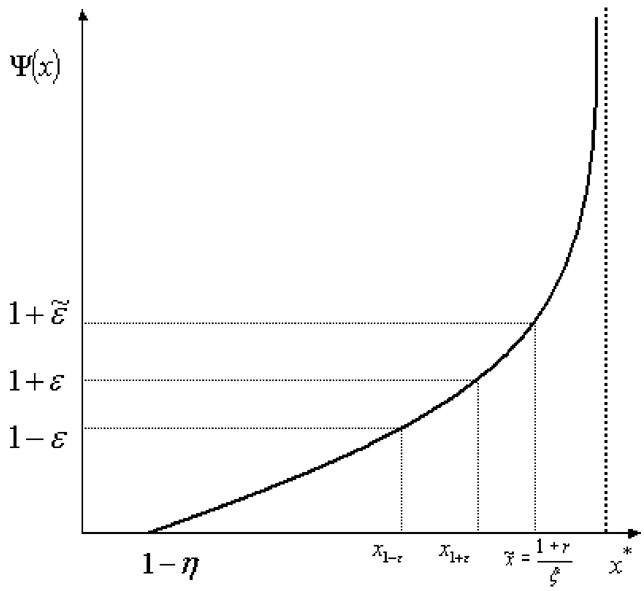


FIGURE 3. Equilibrium Growth of the World Economy



Part 2

Knowledge Spillovers and Productivity Differences

1. Introduction

Observed per capita income differences across countries are enormous: in 1996 average per capita income of the richest country in the world (USA) was about 50 times that of the poorest country (Ethiopia), while observed income dispersion across the 90-10 percentiles was given by a factor of 17. A succesful theory of growth and development should at least be able to account for cross-country income differences, i.e. to explain the actual shape of the world income distribution (WID). The conclusion of a decade of theoretical and empirical research is that, either through development accounting or using calibration techniques, the standard neoclassical framework is unable to explain observed income differences. The development accounting approach, summarized in Caselli (2005), introduces a particular functional form for the production function and uses cross-country data on per capita income and on factors of production to assess the relative contribution of observables input and unobservable efficiency. The general consensus here is that observed differences in physical and human capital stocks are not big enough to match observed differences in income, and differences in efficiency account for about half of income dispersion. Rich countries seem to be rich not only because of bigger stocks of physical and human capital, but largely because of a more efficient use of these stocks: the hypothesis of a common world technology should be rejected, while theoretical approaches explaining the existence and patterns of across-country technology differences are needed. Calibration exercises like those performed by Prescott (1998) reach the same conclusion: using plausible parameterizations of a neoclassical growth model with a common world technology and constant returns to scale in accumulable factors, differences in saving rates for physical and human capital accumulation needed to generate observed cross-country income dispersion are way too high to be consistent with observed differences in investment rates. As Prescott concludes, a theory of total factor productivity differences is needed.

First generation endogenous growth model à la Romer (1990) or Aghion and Howitt (1992) based on an explicit modeling of an R&D sector do provide a theory of TFP differences, but they predict that any parameter change that influences the steady-state value of the R&D volume directly translates into a variation of the long run growth rate. On the other hand a number of empirical studies, such as Evans (1996), suggest that for many countries growth rate differences show very little persistence over time despite the existence of large cross-country variation in rates of investment in physical and human capital. Klenow and Rodriguez (2004) intrepret this as an evidence of "very large international spillovers at the heart of the long run growth process", driving the world economy toward a common long-run growth rate and an equlibrium WID with long-run differences in income levels. Howitt (2000) builds a Schumpeterian model of the world economy in which all R&D performing countries converge to the same long-run growth rate, but differences in per capita income and aggregate productivity survive because of different long-run investment rates: the stability of the long-run WID is obtained through a technology diffusion process, showing that knowledge spillovers are essential in reconciling Schumpeterian growth theory with the empirical evidence.

In the first chapter of this thesis we developed a general equilibrium analysis of a world economy characterized by the existence of cross-country knowledge spillovers. The model predicts the existence of a long run WID, an equilibrium in which every country grows at the same rate but differences in efficiency, factor intensity and per capita income survive because of the existence of knowledge spillovers. Here we perform, using cross-country data, a first empirical evaluation of that theoretical framework aimed at testing the consistency of two knowledge spillovers structures. In particular we test a specification of the traditional technology index, that generalizes that of Eeckhout and Jovanovic (2002) aimed at explaing the distribution of firm size, allowing for two possible structures of cross-country knowledge spillovers: appropriate knowledge spillovers and backward spillovers.

The main insight of the appropriate technology literature, originated 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-print 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: 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. In this way 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, the world factor distribution (WFD), 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).

The traditional representations of knowledge spillovers suppose that followers can freely access technical knowledge created by the technological leader and often model technology diffusion, in line with the seminal contribution by Nelson and Phelps (1966), as an increasing function of the distance from the technological frontier: but this *advantage of backwardness hypothesis*, originally formulated by Gerschenkron (1962), seems to be challenged by the persistence of cross-country productivity differences and by the fact that even growth miracles such as those of some East Asian countries (Taiwan, Hing 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 also display low TFP levels.

We then examine another possible structure of technological spillovers, that reverts conventional wisdom about cross-country knowledge diffusion, analyzing the case in which spillovers are an increasing function of a country factor intensity: an economy can spill free knowledge only about technologies it has already developed, so that the more advanced is a country the greater is the information flow it can intercept. As Boldrin and Levine (2005) suggest, if the non rival character of ideas is a property of their immaterial nature, on the other hand there also exist 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. Another channel through which knowledge spillovers may be larger for 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 over a limited portion of the distribution of existing technologies, since they lack the physical and human capital infrastructures needed to adopt them. In this case, the barrier is given by aggregate capital intensity *relative* to the country with the maximum observed factor endowment, i.e. the technological leader.

We formalize these two knowledge spillovers structures and test them 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 distribution. We find that both knowledge spillovers structures are able to explain over half of the observed static productivity differences. Backward spillovers are in turn more succesful in generating a theoretical WID close to the observed one.

The rest of the paper is organized as follows. Section 2 presents the dataset, the observed pattern of productivity differences and the relationship between observed WID and WFD (world factor distribution). Section 3 introduces the two knowledge spillovers structures and derives the static relationship between relative productivity and relative factor intensity at a point in time. Section 4 present the estimation results and an analysis of the relationship between observed and theoretical WID for both spillovers structures. Section 5 concludes.

2. The World Income Distribution and Productivity Differences

The more direct way to assess the relative contribution of observables input and unobservable efficiency in shaping cross-country income differences is the developement 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

$$y = Ak^{\alpha}h^{1-\alpha} \tag{2.1}$$

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 extracted as a residual 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 (2.1) implies that the wage ω of a worker is such that $\ln\omega \propto \ln h$ and since the wage-schooling relationship is widely tought 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 educational-wage profile documented in Psacharopulos (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 Psacharopulos 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-country differences in income, factor intensity and productivity (TFP) resulting from a decomposition of (2.1) 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): as in the model presented in the first chapter, we interpret factor intensity as an index of the technological state of an economy and propose a specification for the efficiency index which depends on a country factor intensity *relative* to the technological leader, so that 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 $K_t = \max\left\{\text{observed } k^{\alpha}h^{1-\alpha} \text{ in } t = 1996\right\}$.

Denoting, as in Caselli(2005), by $y_{kh} \equiv k^{\alpha}h^{1-\alpha}$ 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 observed level of A for each country using y_{kh} it is possible to pursue a variance decomposition approach since

$$Var[\ln(y)] = Var[\ln(y_{kh})] + Var[\ln(A)] + 2Cov[\ln(A), \ln(y_{kh})] \qquad (2.2)$$

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

One can judge the explanatory power of the neoclassical-common world techology 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. Infact 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 dynamical 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

$$y_r = \frac{y}{y_{\max}} \qquad (2.3)$$

and

$$z = \frac{y_{kh}}{y_{kh}^{\max}} \tag{2.4}$$

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 WID[1]. 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.

3. The Structure of Knowledge Spillovers and Productivity Differences

Taken together, these observations suggest a representation of the technology index A as a function $A(z)$ of relative total factor intensity, with the property that $A'(z) > 0$. Interpreting factor intensity as a measure of a country technological state (the higher total factor intensity the more advanced the technology operated) a natural possibility is a representation in terms of cross-country knowledge spillovers, linking both statically and dynamically factor accumulation and productivity, in

which $A = A(z, h(z))$ and the entire cross-section distribution of factor intensity is an argument of the technology index.

Suppose, as in a continous-time version of the model presented in the first part of the thesis, that the world economy is composed by a unit mass of countries and that income per worker of a country with y_{kh}units of aggregate capital, where $y_{kh} \equiv k^{\alpha} h^{1-\alpha}$ as in (1), is given by

$$y(z,t) = A(z,t) y_{kh} \tag{3.1}$$

[1]The 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.

and total factor productivity is given by

$$A(z,t) = S(z,t)^{\beta} G(t)^{1-\beta} \equiv \left[1 + \int_{D} \alpha(z) h(z,t) dz \right]^{\beta} G(t)^{1-\beta} \qquad (3.2)$$

where $G(t)$ is an efficiency index which grows at the costant and exogenous rate g shared by each country, $z = \frac{y_{kh}}{y_{kh}^{\max}}$ is the country factor intensity relative to the supremum $y_{kh}^{\max}$ of the distribution of y_{kh} among countries, $H(z,t)$ is the distribution function of z defined over $[z_{\min}, 1]$, $h(z,t) = H^{'}(z,t)$ 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, if $\alpha(z) = \alpha$ copying is undirected) and D is the domain over which spillovers act.

Equation (3.2) specifies a 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 tought of as general knowledge, and a knowledge spillovers component $S(z,t)$ that comes from cross-country interactions. The knowledge spillovers part of (3.2) 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 $h(z,t)$: infact, interpreting $h(z,t)$ as a probability density, the integral in $S(z,t)$ represents 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 depends on the average level of the copying function $\alpha(z)$ over the entire relative cross-country factor distribution, as in the single country model of Lucas (1988), in which there is an externality based on the economy-wide average level of human capital, and in Romer (1986) . With those kind of *general or unlocalized knowledge spillovers* $\partial A / \partial z = 0$, there is no correlation between relative factor intensity and productivity and there is a common world technology shared by each country. Productivity can grow over time through cross-country interactions if $h(z,t)$ changes so that the mean value of $\alpha(z)$ increases (e.g. with undirected copying $\alpha(z) = \alpha$ and productivity grows if average relative factor intensity grows), but at a point in time this specification cannot explain systematic productivity differences.
- if $D = [z, 1]$, 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 z-countries have access to the knowledge of a large part of the cross-country distribution of techniques, while high z-countries can't copy much and have to rely more on investment. With this kind of *forward knowledge spillovers* $\partial A / \partial z < 0$ and relative factor intensity and TFP are negatively related: low-z country should display higher levels of efficiency. This is why the Eeckhout-Jovanovic model

cannot explain neither the positive correlation between income levels, levels of physical and human capital and TFP, nor the observed relationship between the WFD and the WID.

- if $D = [\underline{\delta}z, \overline{\delta}z]$, where $\overline{\delta} = (1+\delta)$ and $\underline{\delta} = (1-\delta)$, then A is characterized by *appropriate technology knowledge spillovers*: a country with relative factor intensity z extracts useful knowledge only from countries that are technologically near, where the amplitude of the technological neighbourhood over which spillovers flow is proportional to z and is controlled by the parameter δ. Note that since $h(s,t) = 0$ outside of $[z_m, 1]$, very low-z and very high-z countries don't have a complete technological neighbourhood and the intervals over which the integral is computed are $[z_m, \overline{\delta}z]$ for $z \in \left[z_{\min}, \frac{z_{\min}}{\underline{\delta}}\right]$ and $[\underline{\delta}z, 1]$ for $z \in \left[\frac{1}{\overline{\delta}}, 1\right]$. We argued in the first part that the fact that technological neighbourhoods are increasing in z, meaning that high-z countries spill knowledge from a larger portion of $h(z,t)$, seems logically plausible: advanced economies operate a larger set of technologies and can integrate in their framework also inferior technologies and related knowledge. In this case the sign of $\partial A/\partial z$ depends both on the shape of $h(z,t)$ and on the value δ: it follows that such a representation has the potential to account for static observed productivity differences.
- if $D = [z_m, z]$ then each country is supposed to extract free knowledge only from countries operating technologies it has already developed and we have *backward or technology improving knowledge spillovers.* This specification emphasizes the existence of *barriers to technology adoption* based on factor intensity: a country with low physical-human capital can't have access at zero cost to a large set of existing technologies, simply since it lacks the infrastructures and skills needed to implement them. In this case $\partial A/\partial z > 0$ so that also this representation of knowledge spillovers can account for static productivity differences.

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 first part, of undirected copying with $\alpha(z) = \alpha$: this assumption removes one degree of freedom and assigns exclusively to the WFD and to the free parameter β controlling the strenght 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

$$A_i(z,t,\delta) = S(z,t,\delta)^\beta G_i(t)^{1-\beta} \equiv \left[\alpha \int_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds\right]^\beta T(0)e^{\gamma t + \xi_i} \qquad (3.3)$$

where $T(0) = G(0)^{1-\beta}$ with $G(0)$ 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:

$$a_i(z,t) \equiv \frac{A_i(z,t,\delta)}{A_L(1,t,\delta)} = \left[\frac{\int_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds}{\int_{\underline{\delta}}^{1} h(s,t)ds} \right]^{\beta} e^{\xi_i - \xi_L} \tag{3.4}$$

and taking logs

$$\ln a_i(z,t) = C + \beta \ln D_\delta(z) + \nu \tag{3.5}$$

where $C = -\beta \ln \left[\int_{\underline{\delta}}^{1} h(s,t)ds \right]$, $D_\delta(z) = \left[\int_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds \right]$ 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 (2.1), on the observed quantity $D_\delta(z) = H(\overline{\delta}z) - H(\underline{\delta}z)$, 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 choiche of δ we estimate the associated $\widehat{\beta}$ and we calculate the theoretical or counterfactual income level relative to the technological leader for each country as

$$y^{th}_{R,\delta} = \frac{y^{th}_i(z,t,\delta)}{y^{th}_L(1,t,\delta)} = \left[\frac{\int_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds}{\int_{\underline{\delta}}^{1} h(s,t)ds} \right]^{\widehat{\beta}} z \tag{3.6}$$

and normalizing with respect to the maximum of $y^{th}_{R,\delta}$ (that may not coincide with $y^{th}_L(1,t,\delta)$), we obtain for each choice of the free parameter δ the theoretical world distribution of relative income levels $h^{th}_y(y^{th}_{r,\delta})$ that can be compared to the observed one $h_y(y_r)$.Theoretical or counterfactual relative income is the realtive 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

$$A_i(z,t) = S(z,t)^{\beta} G_i(t)^{1-\beta} \equiv \left[1 + \alpha \int_{z_m}^{z} h(s,t)ds\right]^{\beta} T(0)e^{\gamma t + \eta_i} \tag{3.7}$$

and TFP relative to the technological leader is

$$a_i(z,t) \equiv \frac{A_i(z,t)}{A_L(1,t)} = \left[\frac{1 + \alpha \int_{z_m}^{z} h(s,t)ds}{1+\alpha}\right]^{\beta} e^{\xi_i - \xi_L} \tag{3.8}$$

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

Supposing that $\alpha \gg 1$ and taking logs, it is possible to approximate (3.8) by

$$\ln a_i(z,t) \simeq \beta \ln H(z) + \epsilon \tag{3.9}$$

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

$$a(z_m,t) = \frac{A_l(z_m,t)}{A_L(1,t)} = \left[\frac{1}{1+\alpha}\right]^{\beta} \tag{3.10}$$

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 $\widehat{\beta}$ obtained from (3.9), it is possible to recover the normalizing constant $\widehat{\alpha}$ by matching theoretical and observed $a(z_m,t)$, so that $\widehat{\alpha}$ should be given by $\widehat{\alpha} = \left(\frac{1}{0.26}\right)^{1/\widehat{\beta}} - 1$, so that we can check te mutual consistency of $\widehat{\beta}$ and of the assumption $\alpha \gg 1$.

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

$$y_r^{th} = \frac{y_i^{th}(z,t)}{y_L^{th}(1,t)} = \left[\frac{1+\widehat{\alpha}\int\limits_{z_m}^{z} h(s,t)ds}{1+\widehat{\alpha}} \right]^{\widehat{\beta}} z \tag{3.11}$$

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 $h_y(y_r)$.

It should be noted that our specification of knowledge spillovers can be also tested with panel data about cross-country TFP growth rate differences over time. Time-differentiating (3.3) we obtain the TFP growth rate at time t of a country i with relative factor intensity z with appropriate technology spillovers

$$\frac{\overset{\bullet}{A}(z,t,\delta)}{A(z,t,\delta)} = (1-\beta)\,g + \beta\frac{\int\limits_{\underline{\delta}z}^{\overline{\delta}z} \overset{\bullet}{h}(s,t)ds}{\int\limits_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds} + \beta\left[\frac{\overline{\delta}h(\overline{\delta}z(t),t)}{\int\limits_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds} - \frac{\underline{\delta}h(\underline{\delta}z(t),t)}{\int\limits_{\underline{\delta}z}^{\overline{\delta}z} h(s,t)ds}\right]\overset{\bullet}{z} \tag{3.12}$$

where the first term captures the common growth rate of TFP due to increase in general knowledge, the second term the effect on knowledge spillovers due to the evolution of the WFD, the third term the effect on spillovers of the variation in a country relative factor intensity.

With backward spillovers one obtains

$$\frac{\overset{\bullet}{A}(z,t,\delta)}{A(z,t,\delta)} = (1-\beta)\,g + \beta\frac{\int\limits_{z_m}^{z} \overset{\bullet}{h}(s,t)ds}{\int\limits_{z_m}^{z} h(s,t)ds} + \beta\left[\frac{h(z,t)}{\int\limits_{z_m}^{z} h(s,t)ds}\overset{\bullet}{z} - \frac{h(z_m,t)}{\int\limits_{z_m}^{z} h(s,t)ds}\overset{\bullet}{z}_m\right] \tag{3.13}$$

that has the same interpretation.

We leave for future research the test of the ability of our specification of cross-country interactions to explain differences over time in TFP growth rates: here we focus only on the ability of our specification of technology in replicating observed cross-country differences in TFP levels at a point in time.

4. Estimations and Results

The two equations we estimate are equation (3.5) and (3.9). Data on productivity levels A (TFP) and relative factor intensity z are obtained as described in Section 2: 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 $D_\delta(z) = H_\delta(\overline{\delta}z) - H_\delta(\underline{\delta}z)$ is constructed as follows:

- We let δ vary in $[0, 1]$ taking the possible values $\{0.1, 0.2, ..., 0.9\}$.
- For each possible value of δ we add to the 82 observed values of z, the values $\overline{\delta}z$ and $\underline{\delta}z$ (obviously if $\underline{\delta}z < z_m$ or $\overline{\delta}z > 1$ 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 $D_\delta(z) = H_\delta(\overline{\delta}z) - H_\delta(\underline{\delta}z)$ (obviously if $\underline{\delta}z < z_m$ then $H_\delta(\underline{\delta}z) = H_\delta(z_m)$ and if $\overline{\delta}z > 1$ then $H_\delta(\overline{\delta}z) = 1$).

The last assumption, formulated in order to justify OLS estimates of (3.5) and (3.9), 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 (3.5) 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 significative only in 5 cases. The estimated $\widehat{\beta}$ 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 $\widehat{\beta}$ 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 $\widehat{\beta}$% 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 (3.9)

- The intercept is near zero as predicted by (3.9), even if not statistically significant. The estimated $\widehat{\beta}$ 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 $\widehat{\beta}$ entails a value of $\widehat{\alpha}$ consistent with the assumption $\alpha \gg 1$ used to obtain the regression equation (3.9): infact $\widehat{\alpha} = \left[\left(\frac{1}{0.26}\right)^{1/0.39} - 1\right] \simeq 61$, so that $\widehat{\beta}$ and the assumption $\alpha \gg 1$ are mutually consistent. We will use this value $\widehat{\alpha}$, together with the estimated $\widehat{\beta}$, in the computation of theoretical relative income levels given by (3.11).

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 half of the observed dispersion in relative TFP, hence a significative fraction.

The second dimension over which we evaluate these represenation 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 $\widehat{\beta}$ and observed values for $D_\delta(z)$ and $H(z)$ to calculate relative income levels with appropriate technology and backward spillovers, using respectively (3.6) and (3.11): Figures 2 and 3 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 unconsistently 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 undistinguishable 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 non-parametric 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 explicitely 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 whith appropriate tecnology, $H_y^{th}(y_{r,\delta}^{th})$ given by (10) for each choice of δ, and backward spillovers, $H_y^{th}(y_r^{th})$ given by (3.11). 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

The existence of large and systematic cross-country differences in TFP calls for a rejection of the common world technology assumption and for a theory of productivity differences. We introduce a simple specification for cross-country knowledge spillovers, that generalizes Eeckhout and Jovanovic (2002) and that has been extensively studied in first chapter of out thesis, and that has the potential to account for observed productivity differences. If a country capital intensity is an index of its technological state, then every economy extracts useful knowledge spillovers by sampling a portion of the world distribution of total factor intensity: either from a neighborhood of technologically close economies (appropriate technology spillovers), or from economies operating technologies already adopted (backward spillovers or spillovers with barriers based on factor endowments). In both cases factor accumulation and TFP levels are linked in a fundamental way, since the position of a country in the WFD determines the amount of technical kmowledge it can access. We show that both knowledge spillovers structures are able to explain more than 50% of observed static productivity differences. We then evaluate the consistency of our specification with the observed WID: we show that backward spillovers are more succesful in generating a theoretical WID close to the observed one. A further empirical test left for future research should be a panel data analysis of our technology specification, aimed at explaining cross-country differences in TFP growth rates over time.

6. References

Aghion, P. and P. Howitt (1992), "A Model of Growth Through Creative Destruction", *Econometrica 60*, 323-351.

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.

Barro, R.J. and Lee, J. (2001), "International Data on Educational Attainment: Updates and Implications", *Oxford Economic Papers*, 53 (3), 541-563 (July).

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 McMillan, pp. 151-200.

Eeckhout, J. and Jovanovic, B. (2002), "Knowledge Spillovers and Inequality", *American Economic Review*, 92(5), 1290-1307.

Gollin, D. (2002), "Getting Income Shares Right", *Journal of Political Economy 110(2)*, 458-474.

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.

Heston, A., Summers, R. and B. Aten (2002), *Penn World Tables Version 6.1.* Downloadable dataset. Center for International Comparisons at the University of Pennsylvania.

Howitt, P. (2000), "Endogenous Growth and Cross-Country Income Differences", *American Economic Review 90*, 829-846.if

Klenow, P.J., and Rodriguez-Clare, A. (1997), "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?", in B.S. Bernanke, and J.J. Rotemberg, eds., *NBER Macroeconomics Annual 1997*, MIT Press, Cambridge, 73-103.

Lucas, R. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.

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.

Prescott, E.C. (1998), "Needed: A Theory of Total Factor Productivity", *International Economic Review 39(3)*, 525-551.

Psacharopulos, G. (1994), “Returns to investment in education: A global update”, *World Development 22(9)*,1325-1343.

Quah, D. (1993), “Empirical Cross-Section Dynamics in Economic Growth”, *European Economic Review 37*, 426-34.

Quah, D. (2007), "Growth and Distribution", mimeo, London School of Economics.

Romer, P. (1986), "Increasing Returns and Long Time Growth", *Journal of Political Economy 94(5)*, 1002-1037.

Romer, P.M. (1990), “Endogenous Technological Change”, *Journal of Political Economy 98*, S71-S102.

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

Relative Factor Intensity, Relative Income and Relative Productivity

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	$k^{\alpha}h^{1-\alpha}$	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 (3.5). 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

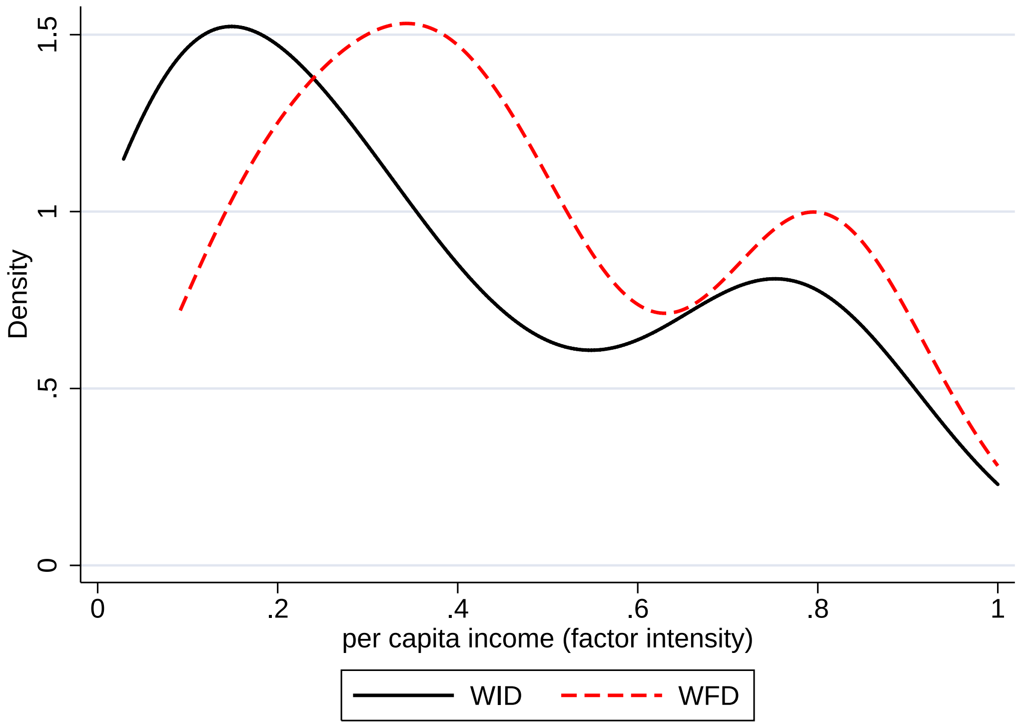
Note: OLS estimates of equation (3.9). Dependent variable is TFP relative to the technological leader (Norway). Robust standard errors in parentheses. *,** and *** mean significantly different from 0 at the $10\%, 5\%$ or 1% level.

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
$P - value$	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	
$P - value$	0.000	0.001	0.010	0.384	
	Backward Spillovers	WFD			
Δ	0.1098	0.2561			
$P - value$	0.645	0.006			

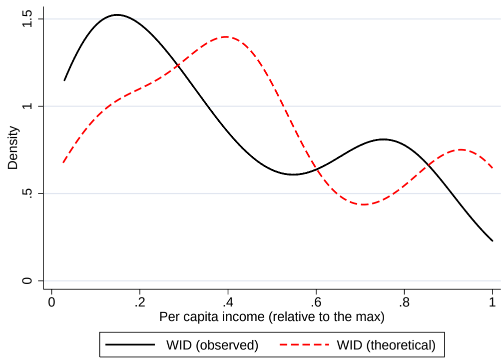
Note: Kolmogorov-Smirnov test for the equality of the observed and theoretical WIDs with appropriate technology and backward spillovers. Δ is the maximum distance between the two distribution and $P - value$ is the corrected P computed by Stata 8.2. WFD is the observed cross-country distribution of relative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

Figure 1 - World Factor Distribution and World Income Distribution

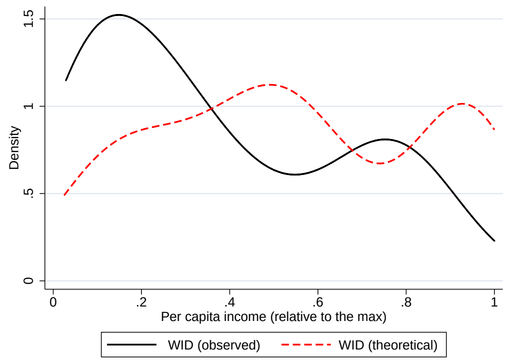


Note: Kernel density estimation of per worker income (WID) taken from PWT 6.1 and physical-human capital aggregate $y_{kh} = k^{\alpha}h^{1-\alpha}$(WFD), constructed as described in the main text, both relative to the observed maximum for the sample of 82 countries in the year 1996. Gaussian kernel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

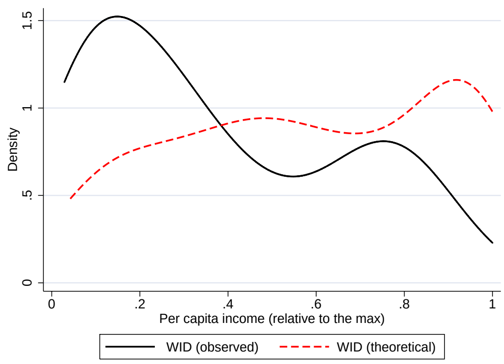
Figure 2 - Observed vs Theoretical WID with Appropriate Technology Knowledge Spillovers



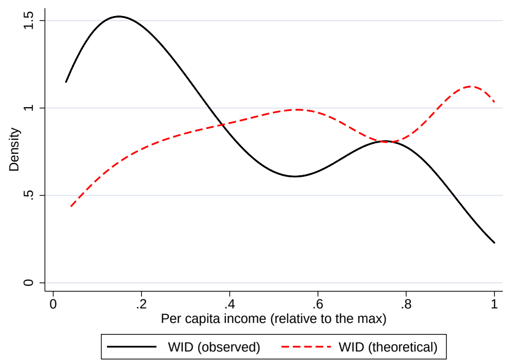
Delta =0.1



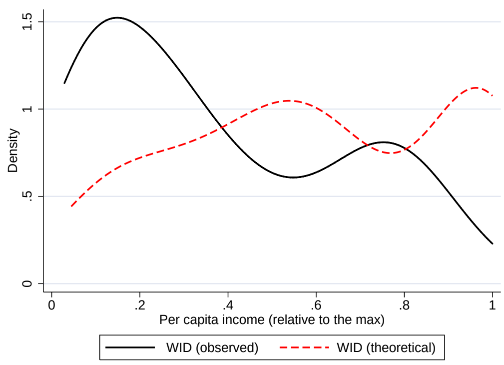
Delta =0.2



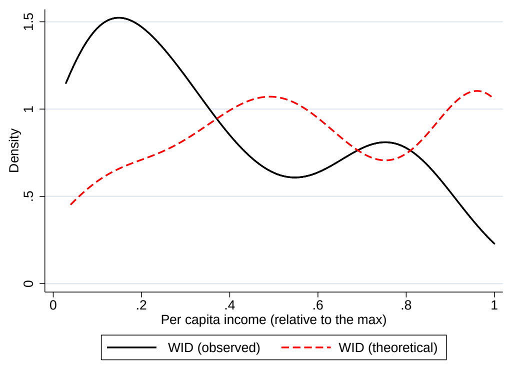
Delta =0.3



Delta =0.4

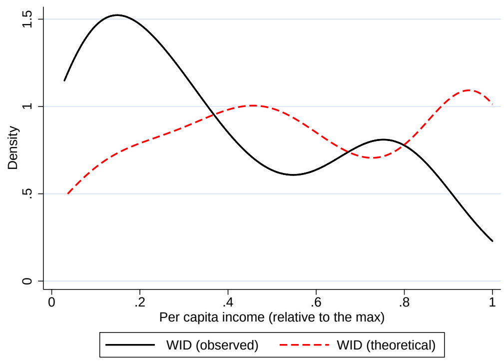


Delta =0.5

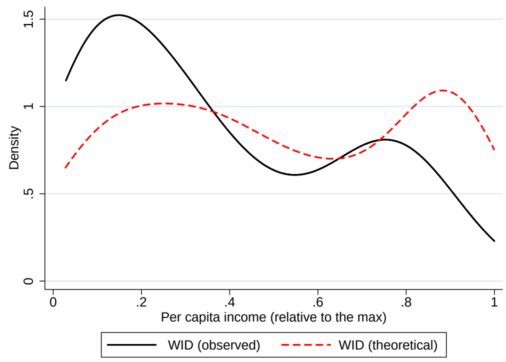


Delta =0.6

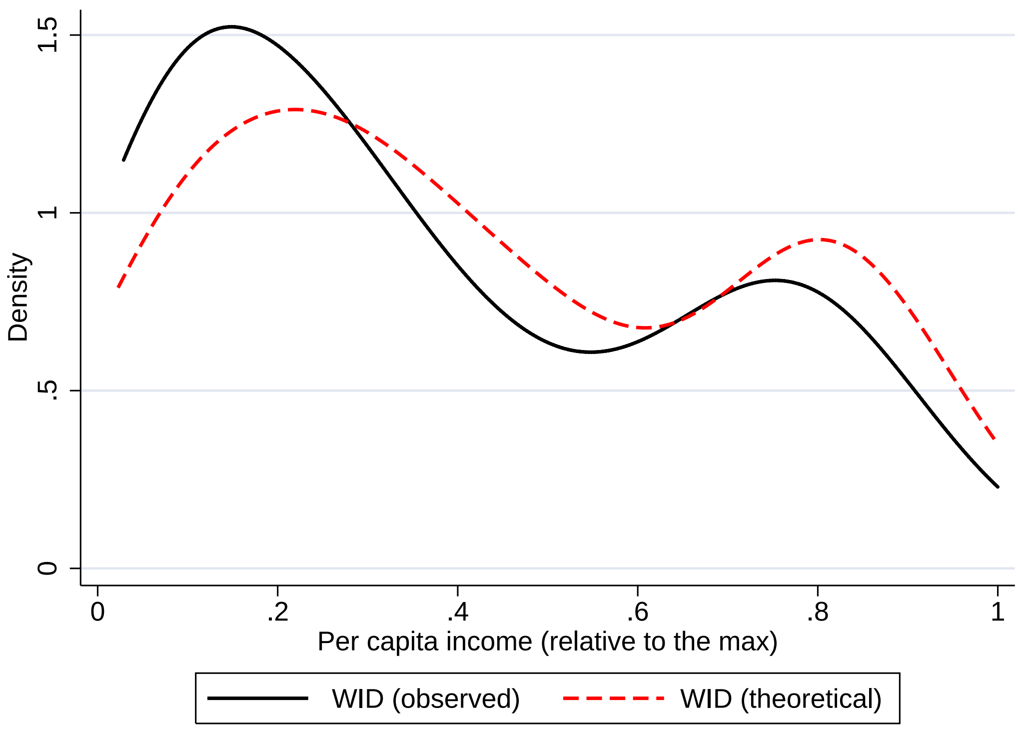
Figure 2 - Observed vs Theoretical WID with Appropriate Technology Knowledge Spillovers (continued)



Delta=0.7



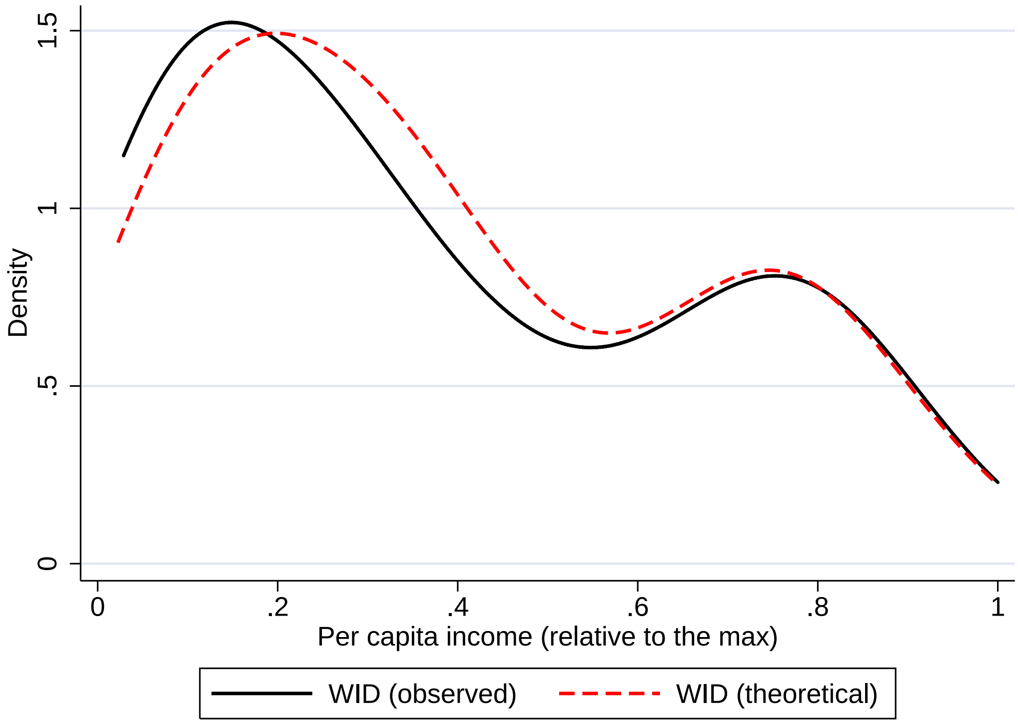
Delta=0.8



Delta=0.9

Note: Kernel density estimation of observed per worker income (WID) taken from PWT 6.1 and theoretical per worker income, described by equation (3.6) 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 3 - 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 (3.11) 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).

Part 3

Capital-Skill Complementarity and Cross-Country Skill Premia

1. Introduction

The Capital -Skill Complementarity hypothesis (CSC), originally formulated by Griliches (1969), can loosely be described as the statement that capital and skilled labor are relative complements or, the other way around, that capital and unskilled labor are relative subsitutes: new machines substitute unskilled workers more easily than skilled ones. Properly formalized, the CSC assumption links the dynamics of wage inequality, as measured by the skill premium, i.e. the wage of skilled workers relative to the wage of the unskilled, to the capital accumulation process: accumulation of capital tends to increase the relative marginal product of skilled labor and, in a competitive setting, the skill premium. The CSC assumption provides a framework for a quantitative analysis of time series and cross country patterns of skill premia in terms of two observable variables, the relative supply of skilled labor and the capital stock. Krusell et al (2000) evaluate the potential of CSC in explaining the dynamics of the U.S. wage inequality in the last 30 years, a period over which the skill premium has continuously grown in spite of a large increase in the relative supply of skilled labor: it turns out that with CSC, the rapid growth in the stock of capital equipment is able to generate a behaviour of the U.S. skill premium quite close to the observed one. Moreover Krusell et al (2000) show that CSC is also consistent with other stylized facts which characterize U.S. growth experience, such as the relative constancy of the aggregate labor share of income and a trendless rate of return on physical capital.

The linkage imposed by the CSC hypothesis between wage inequality and the capital accumulation process can be exploited to test the consistency of CSC using cross-country data on skill premia, relative supplies of skilled labor and observables that determine national capital stocks. Intuitively, the skill premium should be relatively higher in developing countries because of a low supply of skilled labor, a pure quantity effect. On the other hand, it should also be relatively lower there because of low saving rates and because of the existence of barriers to capital accumulation, policy distorsions and institutional settings that discourage investments.

We proceed by obtaining a steady-state approximation for aggregate capital intensity in a Solow-setting with an exogenous saving rate and we derive a linear relationship between the skill premium, the relative supply of skilled labor, the saving rates and barriers that can be used for a cross-country regression much in the spirit of Mankiw, Romer and Weil (1992). This approach allows us to test for the presence of CSC and at the same time to obtain an estimate for the elasticity of substitution between each couple of inputs: thus, our study also adds to the literature on the specification of the aggregate production function.

Is the CSC hypothesis also consistent with cross-country differences in the skill premium? Since CSC depends on the dynamics of capital accumulation the latter question can be reformulated as: is the cross-country pattern of skill premia consistent with existing differences in saving rates and with barriers to capital accumulation? Can these observables explain the existing variation of cross-country skill-premia or are technology differences between countries required?

These questions are very similar to those that characterize recent theoretical and empirical research in growth theory. This is in a sense obvious since in such a highly aggregate setting the skill premium is a pure macroeconomic phenomenon: the fundamental question is whether the cross-section of observables quantities (and in particular capital) are able to explain growth and inequality patterns, or if unobservable effects (technology) act in a systematic and quantitatively relevant way. This study can also be interpreted as an indirect test of the existence of a common world technology, using cross-country skill premia instead of cross-country income differences as a testing ground.

Our empirical approach to the evaluation of the CSC hypothesis is new, but a number of other studies have tried to test the consistency of CSC using cross-country data: Duffy, Papageorgiou and Perez-Sebastian (2004) directly estimate the parameters that control the elasticitiy of substitution between inputs of several version of a production function allowing for CSC, using a panel dataset of 73 countries over the period 1965-90, while Papageorgiou and Chmelarova (2005) use data on cross-country skill premia in a partial equilibrium framework, obtaining a testable relationship between the skilled labor share of the wage bill, the capital-output ratio and the skill premium.

We find weak evidence in support for the CSC hypothesis for the full sample, but strong evidence in favour of CSC in the non-OECD subsample and no evidence in the OECD subsample. This results confirm and reinforce those obtained by Papageorgiou and Chmelarova (2005): there is increasing evidence that developing countries are undergoing a transition similar to that documented by Goldin and Katz (1998) for the U.S. during the first part of the twentieth century, when the shift from the classical factory to continuous production processes seems to have been characterized by CSC. Furthermore, we obtain estimates of the elasticities of substitution close to those reported in the literature only when the threshold separating unskilled and skilled labor is relatively high: defining as skilled those workers who have completed only a primary cycle of education seems to be too weak a requirement, leading to unrealistic degree of substitutability between couple of inputs. Finally we are able to get an estimate of the steady-state or long run impact on the skill-premium of a change in the effective saving rate, i.e. the saving rate corrected for the magnitude of policy barriers discouraging investments: for OECD countries where CSC is uneffective the impact is null, while in non-OECD countries in which CSC acts a 10% increase in the effective saving rate it generates a 3% increase of the skill premium. As a final remark, obesrvables quantities are able to explain only up to 30% of observed cross-country dispersion of skill-premia, suggesting that unobservable forces have a role in shaping cross-country inequality. If the analysis of cross-country income differences suggest that large international productivity differences are needed, the dispersion of cross-country skill premia reinforces this perspective and calls for a theory of productivity differences able to simultaneously explain growth and inequality facts.

The rest of the paper is organized as follows. Section 2 reviews some key facts on CSC and the skill-premium. Section 3 presents our steady-state approximation of the skill premium and derives the estimating equation. Section 4 presents the data and the results of the estimations. Section 5 concludes.

2. Capital-Skill Complementarity and the Skill Premium

Suppose, as in Caselli and Coleman (2006), that per capita output is given by

$$y = F(k, L_u, L_s) = k^{\alpha} \left[(A_u l_u)^{\sigma} + (A_s l_s)^{\sigma} \right]^{\frac{1-\alpha}{\sigma}} \qquad \sigma < 1 \qquad (2.1)$$

where k is per worker capital, l_u and l_s are respectively the unskilled and skilled share of total labor $L = L_u + L_s$, while A_u and A_s are efficiency indexes of the two types of labor. This production function generalizes the Cobb-Douglas technology by splitting the labor force of an economy into two types of labor and allowing a variable degree of substitutability between them, given by the elasticity of substitution $\varepsilon = \frac{1}{1-\sigma}$. On the other hand the elasticity of substitution between physical capital and the CES labor aggregate is fixed to one by the Cobb-Douglas assumption and the partial elasticities of substitution between capital and skilled and unskilled labor are the same.

If markets are competitive, factors are paid their marginal product and the skill premium is given by

$$\frac{\omega_s}{\omega_u} = \frac{\partial F/\partial L_s}{\partial F/\partial L_u} = \left(\frac{A_s}{A_u}\right)^{\sigma} \left(\frac{L_s}{L_u}\right)^{\sigma-1} \tag{2.2}$$

Here the dynamics of the skill premium is controlled by a technology effect, given by the ratio $\frac{A_s}{A_u}$, which measures the degree of skill-bias of technology, and by a quantity effect, given by the relative supply of skilled labor $\frac{L_s}{L_u}$. An increase in the relative supply of skilled workers always lowers the skill premium, while an increase in the relative efficiency of skilled labor raises the skill premium only if $\sigma > 0$ and the elasticity of substitution ε is sufficiently high, which seems to be the empirically relevant case (Acemoglu (2003)). Note that the level of the skill premium is independent of the stock of per worker physical capital in the economy: the capital accumulation process plays no role in shaping the degree of wage inequality. This is a consequence of the assumption of a unique degree of substitutability between capital and both types of labor: intuitively, (1) represents machines as being skill-neutral, equally substituting (or complementing) unskilled and skilled labor.

Note also that in this framework the only measurable quantity that can be used to study the time series and cross-country behaviour of the skill premium is the relative supply of skilled workers, since the technology ratio $\frac{A_s}{A_u}$ is unobservable.

Suppose instead that per capita output is given by the two level CES[1]:

$$Y = \left[(A_x X)^{\sigma} + (A_u L_u)^{\sigma}\right]^{\frac{1}{\sigma}} \qquad \sigma < 1 \tag{2.3}$$

$$X = \left[bK^{\theta} + (1-b)\, L_s^{\theta}\right]^{\frac{1}{\theta}} \qquad \theta < 1 \tag{2.4}$$

which is obtained by nesting into a first CES aggregate with unskilled/raw labor, a second CES composite that aggregates physical capital and skilled labor: this second aggregate, denoted as X, should be interpreted as a measure of total capital, physical and human.

Here A_x and A_u are indexes of the technological state of the capital-skill composite and of unskilled labor respectively, b is a share parameter which is supposed

[1] Originally introduced by Sato (1967), particular versions have been used by Krusell et al. (2000), Goldin and Katz (1998) and Duffy, Papageorgiou and Perez-Sebastian (2004), while its general properties in a Solow setting are discussed in Papageorgiu and Saam (2005).

constant, σ and θ are curvature parameters that determine the elasticity of substitution between K, L_u and L_s.

In particular Duffy, Papageorgiou and Perez-Sebastian (2004) have shown that if $\sigma > \theta$ then Capital-Skill Complementarity holds, since both the partial and direct elasticities of substitution between L_u and K are bigger than the elasticity of substitution between L_s and K, and physical capital subsitutes unskilled labor more easily than skilled labor.

In a competitive setting the skill premium is given by:

$$\frac{\omega_s}{\omega_u} = \frac{\partial F/\partial L_s}{\partial F/\partial L_u} = (1-b)\left(\frac{A_x}{A_u}\right)^{\sigma}\left(\frac{L_s}{L_u}\right)^{\sigma-1} x^{\sigma-\theta} \tag{2.5}$$

There are now three components which influence the degree of wage inequality: a technology effect given by the ratio $\frac{A_x}{A_u}$, a labor supply effect given by the relative supply of skilled labor $\frac{L_s}{L_u}$, and the CSC effect given by the $x^{\sigma-\theta} \equiv \left[\frac{X}{L_s}\right]^{\sigma-\theta} = [bk + (1-b)]^{\frac{\sigma-\theta}{\theta}}$ term, which now links the magnitude of the skill premium to the capital accumulation process.

In particular an increase in the stock of aggregate capital per skilled worker $x = \frac{X}{L_s}$ increases the skill premium if and only if CSC holds ($\sigma > \theta$), while the percentage variation in the skilled premium induced by a one per cent increase in capital per unit of skilled labor $k = K/L_s$ id given by

$$\frac{\partial \ln\left(\frac{\omega_s}{\omega_u}\right)}{\partial \ln k} = (\sigma-\theta)\frac{bk^{\theta}}{bk^{\theta} + (1-b)} = (\sigma-\theta)\, b\left(\frac{k}{x}\right)^{\theta} = (\sigma-\theta)\, b\left(\frac{K}{X}\right)^{\theta} \tag{2.6}$$

which is positive if and only if CSC holds.

Its magnitude varies with the relative share of physical capital over aggregate (physical and human) capital X and with the degree of substitutability between physical and human capital controlled by θ: if θ is positive, and physical and human capital are relative substitutes, the higher the physical share of aggregate capital the greater the effect on the skill-premium, while the reverse is true when physical and human capital are relative complements ($\theta < 0$).

In the CSC framework capital accumulation does play a decisive role in shaping the degree of wage inequality of an economy and differences in the parameters which characterize this process (i.e. saving rates and barriers to capital accumulation) will translate in differences, over time and across ecomomies, in skill premia: it is possible to use cross-country data on skill premia at a point in time to test the consistency of the CSC hypothesis. In particular the skill premium should be relatively high in the skill scarce-low income countries because of the quantity effect acting through the $\frac{L_s}{L_u}$ ratio, but it should also be relatively high in the skill-abundant-high income countries because of higher saving rates and lower barriers to capital accumulation acting on the stock of aggregate capital x.

3. The Skill Premium in the Steady State

The general idea is to calculate the steady state value of x for an economy with a production function given by (3) and (4) and a Solow-like assumption on the saving rate, in order to obtain a steady-state relationship between the skill premium and the saving rate that could be used to evaluate the explanatory power of the CSC hypothesis. Since our focus is on the ability of the capital accumulation process and its determinants to explain observed international variation in the skill premium, we abstract from skill-biased technology differences and impose a Hicks neutral common world technology on the production function (2.3) and (2.4)

$$\begin{aligned} Y_t &= A_t \left\{ \alpha X_t^{\sigma} + (1-\alpha) L_{u,t}^{\sigma} \right\}^{\frac{1}{\sigma}} \\ &= A_t \left\{ \alpha \left[b K_t^{\theta} + (1-b) L_{s,t}^{\theta} \right]^{\frac{\sigma}{\theta}} + (1-\alpha) L_{u,t}^{\sigma} \right\}^{\frac{1}{\sigma}} \end{aligned} \tag{3.1}$$

Suppose the law of motion of *aggregate* capital is given by

$$\overset{\bullet}{X}_t = \frac{I_t}{\pi} - \delta X_t \tag{3.2}$$

where I_t is investment in term of the consumption good and π is a parameter that determines the rate at which foregone consumption is transformed into capital goods: π should be interpreted as a measure of barrier to capital accumulation, i.e. taxes, corruption, institutional settings that raise the relative price of investment. We assume in (3.2) that the economy has to accumulate through direct investment not only physical capital, but the whole X composite: here investment should be thought of as involving the combination of machines and human capital, and I_t is a measure that includes both capital equipment and intangibles as organizational and knowledge capital.

Finally we simplify the steady-state analysis by imposing that both unskilled and skilled labor supplies are given ($L_{s,t} = L_s$ and $L_{u,t} = L_u$), and that world technology evolves exogenously at a rate g, i.e. $\overset{\bullet}{A}_t / A_t = g$.

With a Solow assumption $I_t = sY_t$, where s is the exogenously given saving rate, we have:

$$\overset{\bullet}{X}_t = \frac{s}{\pi} Y_t - \delta X_t \tag{3.3}$$

and $\frac{s}{\pi}$ is the effective saving rate of a country with barriers π: for a given saving rate s, the higher the level of the barriers to capital accumulation π, the lower the fraction of total foregone consumption sY that is transformed into new capital goods.

Considering aggregate capital per unit of effective skilled labor $\widetilde{x} = X/AL_s$, (3.2) becomes

$$\overset{\bullet}{\widetilde{x}}_t = \frac{s}{\pi} \frac{Y_t}{A_t L_s} - (\delta + g) \widetilde{x}_t \tag{3.4}$$

so that the steady-state level $\widetilde{x}^*$ of aggregate capital is obtained by setting $\dot{\widetilde{x}}_t = 0$

$$\frac{s}{\pi}\left[\alpha(\widetilde{x}^*)^\sigma + (1-\alpha)\frac{L_u}{L_s}^\sigma\right]^{\frac{1}{\sigma}} = (\delta+g)\widetilde{x}^* \Rightarrow \widetilde{x}^* = \left[\frac{(1-\alpha)\left(\frac{L_u}{L_s}\right)^\sigma}{[\frac{\pi(\delta+g)}{s}]^\sigma - \alpha}\right]^{\frac{1}{\sigma}} \tag{3.5}$$

For all possible values of σ, an increase of the saving rate s or of the relative supply of unskilled labor raise the steady-state level of aggregate capital, while an increase of the depreciation rate, the technology growth rate or of the barriers magnitude decreases it. The effects of cross-country differences in saving rates and barriers on the steady-state value of the skill premium are straightforward: a higher saving rate (or lower barriers to capital accumulation) raises $\widetilde{x}^*$ and, if CSC holds, the steady-state skill premium increases.

It is now possible to express the steady-state level of the skill premium in terms of technology, the relative supply of skilled labor and the variables that control capital accumulation.

Since $x^* = A_t\widetilde{x}^*$, from (2.5) we obtain an expression for the steady-state value of the skill premium

$$\ln\left(\frac{\omega_s}{\omega_u}\right) = \ln(1-b) + \sigma\ln\left(\frac{\alpha}{1-\alpha}\right) + (1-\sigma)\ln\left(\frac{L_u}{L_s}\right) + (\sigma-\theta)\ln(A_t) + + (\sigma-\theta)\ln(\widetilde{x}^*) \tag{3.6}$$

Notice that if CSC holds ($\sigma > \theta$) then in a steady state with a constant relative supply of unskilled labor $\frac{L_u}{L_s}$, the skill premium grows at the exogenous rate of neutral technological progress g, while if capital and skilled labor are relative substitutes ($\sigma < \theta$) the skill premium decreases with technological change[2].

It is possible to approximate $\ln(\widetilde{x}^*)$ by a first-order Taylor approximation in the neighborhood of $\sigma = 0$ to obtain[3] (see Appendix A):

$$\ln(\widetilde{x}^*) \simeq \ln(\frac{L_u}{L_s}) + \frac{1}{1-\alpha}\ln(\delta+g) + \frac{1}{1-\alpha}\ln(\frac{s}{\pi}) \tag{3.7}$$

and inserting this approximation in (3.6)

---

[2]The observed oscillations of the skill premium along the growth path of some economies, documented by Krusell et al. (2000) for the U.S. and by Lindquist (2005) for Sweden are not consistent with this prediction, but they can be expression of out of steady-state behaviour of the economies. Moreover, here we neglect movements in relative supply of skilled labor, which directly influence the skill premium in and out of the steady-state.

[3]Notice that approximating for σ close to 0 means that the production function (7) gets closer and closer to a Cobb-Douglas aggregate of total capital and unskilled labor, $Y_t \simeq A_t X_t^a L_u^{1-a}$ as in Barro and Sala-i-Martin (1995, pag.56).

$$\ln\left(\frac{\omega_s}{\omega_u}\right) = \ln(1-b) + \sigma\ln\left(\frac{\alpha}{1-\alpha}\right) + (\sigma-\theta)\ln(A_t) + (1-\theta)\ln\left(\frac{L_u}{L_s}\right) + + \frac{(\sigma-\theta)}{(1-\alpha)}\ln(\delta+g) + \frac{(\sigma-\theta)}{(1-\alpha)}\ln(\frac{s}{\pi}) \quad (3.8)$$

Equation (3.8) is a steady-state *linear* approximation of the skill-premium that predicts the magnitude of the coefficients on the observables $\frac{L_u}{L_s}$ and $\frac{s}{\pi}$ and that can be estimated using cross-country data in order to obtain an estimate of the curvature parameters of the production function and at the same time to perform a new test on the consistency of the CSC hypothesis. Krusell et al (2000) use US time-series data to obtain an estimate of the curvature parameters of a two-level CES, using a simulation based estimation technique in a partial equilibrium framework, Duffy, Papageorgiu and Perez-Sebastian (2004) estimate directly a two-level CES production function using nonlinear least squares while Papageorgiu and Chmelarova (2005) use cross-country data on the skill-premium in a partial equilibrium framework known as the quasi-fixed cost function approach.

None of those studies focuses on steady-state or on parameters which influence capital accumulation: here household preferences over consumption are included in their simplest form, an exogenous and constant saving rate, and barriers to capital accumulation are taken into account, in order to assess the contribution of the capital accumulation process on wage inequality as measured by the skill premium.

4. Data, Estimation and Results

We use data taken fom a variety of sources:

- Data for the two main variables, the skill premium $\frac{\omega_s}{\omega_u}$ and the labor aggregates L_u and L_s, are obtained from Caselli and Coleman (2006). They build from the Barro and Lee (2001) dataset on the share of the labor force into seven different categories of educational attainment, three different measures of skilled and unskilled labor based on alternative thresholds: the primary completed threshold, by which is considered skilled everyone who has completed a primary cycle of education and unskilled who has not, the secondary completed threshold and the college completed one. These thresholds determine increasing requirements for a worker to be considered skilled and produce an increasing magnification of cross-country differences in the relative supply of skills. Once the threshold has been defined, each of the seven subgroup is then aggregated to form L_u and L_s using its wage relative (obtained as described below) to a base group as a weight: it follows that L_u and L_s are measured in efficiency units of the chosen base group. Data on educational attainment are from 1985.

  For each country Caselli and Coleman also estimate the difference in years of schooling between different subgroups, which, together with the cross-country estimates on Mincerian coefficients taken from Psacharopoulos (1994) and Bils and Klenow (2000), determines the skill-premium for each alternative skill threshold: infact the estimated Mincerian coefficient β_i, which is obtained by regressing in each country i the log wage on years of schooling, is simply the percentage gain associated with an extra year spent in school and if n is the difference in schooling years between skilled and unskilled labor the skill-premium is $\exp(\beta_i n)$ .Mincerian coefficients collected in Psacharopoulos (1994) and Bils and Klenow (2000) come from country-level studies using micro-data: this studies were executed in different years, that we report in Appendix B. Since they reflect institutional features of a country schooling system, the estimated Mincerian coefficients should also show a high degree of persistence over time and the fact that the estimates come from different years should not affect the analysis.

- Data for the saving rate s are taken from Mankiw, Romer and Weil (1992) that draw them from Penn World Tables 4.0 as the average share of real investment on real GDP, including government investment, over the period 1960-1985. Since we take s to be the rate of investment of the capital-skill composite X, this variable could well be underestimated since the average share of real investment does not include intangibles like knowledge and or studiesganizational capital.
- The parameter π which measures barriers to capital accumulation is empirically identified, as in Restuccia and Urrutia (2001), with the relative price of aggregated investment over consumption: cross-country differences in this relative price level should be interpreted as a measure of policy distorsions which discourage capital accumulation. We construct π as the average over the period 1960-1985 of the relative price of investment, measured as the ratio P_I/P_C where P_I and P_C are respectively the price level of consumption and investment given in the Penn World Tables 6.1 (Heston, Summers and Aten (2002)). We normalize this average, as in Restuccia and Urrutia (2001), by the average relative price of investment in the U.S to facilitate cross-country comparisons.

There are 52 countries for which data on all relevant variables are available. Appendix B lists the countries in the sample, the variables and presents some descriptive statistics for the full sample and the OECD and non-OECD subsamples.

In order to estimate the steady-state equation (3.8) we make four identifying assumptions :

- A common world technology assumption: the Hicks-neutral efficiency index grows exogenously at a constant rate shared by all countries of the world economy, so that $A_{i,t} = A_{i,0}\exp(gt)$.
- While every country shares the same technology growth rate g, the $A_{i,0}$ term reflects the initial technology conditions of each country and includes elements such as resource endowments, climate and institutions that may differ across countries: we assume, exactly as Mankiw, Romer and Weil (1992) in their cross-country growth regression specification, that

$$\ln A_{i,0} = a + \epsilon_i \tag{4.1}$$

where a is a constant and ϵ_i is a country-specific shock *independent* from the saving rate s and the level of barriers π: initial conditions on a country technology level are supposed not to affect, at least in the long run, the parameters that shape the future path of capital accumulation. This assumption makes it possible to estimate consistently equation (3.8) by ordinary least squares, but it should be noted that if saving rates, barriers to capital accumulation and initial technology conditions are correlated, then steady-state skill premia and the ratio $\frac{s}{\pi}$ are endogenous and the OLS coefficient on $\frac{s}{\pi}$ would be biased.

- As noted in Appendix A, approximating the skill premium for σ close to 0 means evaluating the steady-state in the limit in which the production function (3.1) approaches the Cobb-Douglas form $Y_t = A_t X_t^a L_u^{1-a}$, so that $(1-\alpha)$ is simply the share of output that goes to unskilled labor: we assume that this share is a *constant* across countries. Gollin (2002) shows that labor income share oscillates for most countries in the range 0.65-0.80, but he does not separately calculate unskilled and skilled labor share. Since we need an explicit value for $(1-\alpha)$ in order to identify σ and θ, we will make a rough guess and set $(1-\alpha) = \frac{1}{3}$ for all countries[4].
- Finally, we assume that also the depreciation rate δ is a constant that does not vary across countries.

Substituting for A_t into equation (3.8) and adding up the constant terms we get a cross-country steady-state approximation of the skill-premium at a point in time (here for simplicity $t=0$) in terms of the observables $\frac{L_u}{L_s}$ and $\frac{s}{\pi}$

$$\ln\left(\frac{\omega_s}{\omega_u}\right)_i \simeq C + (1-\theta)\ln\left(\frac{L_u}{L_s}\right)_i + \frac{(\sigma-\theta)}{(1-\alpha)}\ln\left(\frac{s}{\pi}\right)_i + \nu_i \tag{4.2}$$

where $\nu_i = (\sigma-\theta)\,\epsilon_i$.

The specification that will be tested using cross-country data on $\frac{\omega_s}{\omega_u}$, L_u and L_s, s and π will be the linear relationship

$$\ln\left(\frac{\omega_s}{\omega_u}\right)_i = \beta_0 + \beta_1\ln\left(\frac{L_u}{L_s}\right)_i + \beta_2\ln\left(\frac{s}{\pi}\right)_i + \nu_i \tag{4.3}$$

where the interpretation of the coefficients β_0, β_1 and β_2 is given by equation (4.2). Note that σ can be obtained using the OLS estimates of β_1 and β_2 through the linear transformation $\sigma = 1 + (1-\alpha)\,\beta_2 - \beta_1$. It is clear that if $\beta_2 > 0$, then $\sigma > \theta$ and CSC holds.

To assess the relevance of barriers to capital accumulation in explaining skill-premia dispersion, we ran the OLS regression of (4.3) with and without barriers: that is, we use alternatively $\ln(s)$, imposing $\pi = 1$ for each country in the sample as if there were no institutional distorsion over the capital accumulation process, and $\ln\left(\frac{s}{\pi}\right)$ as regressors. To address the issue of parameter heterogeneity, that questions the usage of a common production function for countries that are at different stages of their development path[5], we split the sample in an OECD and a non-OECD subsamples and run separate regressions for each of them.

---

[4] Incidentally, a value of α around $\frac{2}{3}$, which is plausible for a broad notion of capital including physical structures and intangibles like human capital, is also found by Restuccia and Urrutia (2001) to be the value needed in order to replicate observed income disparities using a simple one sector growth model with barriers to capital accumulation.

[5] See for example Durlauf, Kourtellos and Minkin (2001) and Masanjala and Papageorgiou (2004) for an evaluation of cross-country parameter heterogeneity for the Solow growth model, using respectively a Cobb-Douglas and CES specification for the production function.

There are several interesting features of the regression results, presented in Tables 1-6:

- There is very weak evidence in favour of the CSC hypothesis in the full sample, either if barriers to capital accumulation are included or excluded: the coefficient on $\ln(s)$ or $\ln\left(\frac{s}{\pi}\right)$ is always positive but never significantly different from zero, for all of the possible choices of the skill threshold (Tables 1 and 2). Considering separately the OECD and the non-OECD samples reveals a consistent pattern: CSC appears not to hold in the OECD subsample, since the relevant coefficient becomes negative, even if imprecisely estimated, for each skill threshold and with or without barriers (Tables 3 and 5). On the other hand, CSC appears to hold consistently in the non-OECD subsample: the relevant coefficient is always significantly different from zero (Tables 4 and 6)

  This finding replicates that of Papageorgiou and Chmelarova (2005), but is obtained using a completely different approach: while they use a partial equilibrium framework that generates a linear relationship between the skilled-labor share of the wage bill and the capital-output ratio, here we obtain a steady-state linear approximation relating skill premia to the determinants of capital accumulation. The evidence in favour of a *dynamic degree of substitutability/complementarity* between capital, skilled and unskilled labor along the growth path of an economy is reinforced. It seems that countries that are at lower stages of the development path experience a strong degree of complementarity between capital and skills, which is reflected in the behaviour of skill-premia among them as a subset of the world economy: skilled labor receives a premium which may be interpreted as a sign of its crucial role in complementing relatively advanced technologies in a relatively bacward technology environment. This is also consistent with Goldin and Katz (1998) finding of a high degree of capital-technology-skill complementarity in the transition from the classical factory to continuous process in the U.S. manufacturing sector in the period 1909-1940: it is conceivable that developing countries are experiencing a similar kind of transition, in which a relatively scarce skilled labor force commands an higher skill premium because of its essentiality in activating new technologies.

  On the other hand, the fact that there is no evidence for CSC in the OECD subsample does not necessarily contradict the empirical findings of Krusell et al. (2000) and Lindquist (2005) about the relevance of CSC in explaining the *dynamic behaviour* of the skill premium in the U.S. and in Sweden: it simply indicates that, in this subsample as a whole, the observed dispersion at a point in time of the skill premium is not related to capital intensities as predicted by CSC, and points out that alternative mechanisms for generating the observed dispersion in wage inequality within the subgroup of developed countries are needed.
- The steady-state relationship (4.2) not only does offer an explicit test over the existence of CSC, but it also allows for a separate identification of the substitution parameters σ and θ: the predicted elasticities of substitution between each couple of inputs can be compared with those obtained in the literature, in general through microeconomic studies as summarized in

Hamermesh (1993), to verify the ability of this pure cross-country macroeconomic setting to generate plausible values of these elasticities.

We use the Allen partial elasticity of substitution, that for each couple of inputs measures the percentage change in the ratio of the two inputs generated by one per cent increase in the ratio of their prices, holding the price of the other input and output quantity constant, while the quantity of the other input is free to vary. Intuitively, in a partial equilibrium framework the inverse of the elasticity of substitution ε_{L_s,L_u} between skilled and unskilled labor measures the percentage increase of the skill premium following a one percent increase in the ratio $\frac{L_u}{L_s}$.

Proceeding as in Sato (1967) one obtains the partial elasticities of substitution:

$$\begin{aligned} \varepsilon_{K,L_u} &= \varepsilon_{L_s,L_u} = \frac{1}{1-\sigma} \\ \varepsilon_{K,L_s} &= \frac{1}{1-\sigma} + \frac{1}{\zeta_X}\left[\frac{1}{1-\theta} - \frac{1}{1-\sigma}\right] \end{aligned} \tag{4.4}$$

where ζ_X is the relative share of K and L_s in total expenditure.

In a closed economy setting with competitive markets, ζ_X coincides with the share of GDP which goes to K and L_s: since we set the share of GDP which goes to unskilled labor to $\frac{1}{3}$, we impose $\zeta_X = \frac{2}{3}$ and we calculate the two elasticities using (18) and the estimates of σ and θ.

The values of the implied elasticities vary with the skill-threshold, with the inclusion or exclusion of barriers and with the samples: in general the full sample and the non-OECD subsample generate reasonable values, while the OECD subsample (which is composed of only 17 observations) does not. Using the primary school threshold the implied elasticities are in general way too high to be consistent with those reported by the literature, ranging from 15 to 50 for ε_{K,L_u} and ε_{L_s,L_u}, and from 10 to an astonishing and completely unrealistic 738 for ε_{K,L_s} (OECD subsample with barriers). Using the secondary school threshold the values become more realistic, especially for the full sample (around 4 for ε_{K,L_u} and ε_{L_s,L_u}, around 3 for ε_{K,L_s}) and for the non-OECD subsample (respectively around 4 and around 2). The high school threshold *always* generates realistic values for the elasticities of substitution between inputs, for all the samples with and without barriers: the estimated values of ε_{K,L_u} and ε_{L_s,L_u} are 2.5 for the full sample and 3.5 for the non-OECD subsample, while ε_{K,L_s} is around 2 for both samples.

One can guess fom these results that the primary school threshold is inappropriate as a definition of skilled labor, since it generates unrealistic values of the elasticities of substitution: defining skilled those workers "who can at least read a simple text (e.g. a simple set of instructions or a newspaper article) and perform some basic arithmetic operations", as Caselli and Coleman (2006) describe the primary school threshold, seems to be too weak a requirement[6].

[6]Caselli and Coleman (2006) results on the existence of cross-country skill-biased technology differences is stronger when the primary school threshold is used: our results could suggest that

The fact that, at least for the full and non-OECD samples, the steady-state analysis of such an higly aggregate setting is able to generate elasticities of substitution consistent with those estimated by the microeconomic literature is remarkable: Hamermesh (1993) reports values for ε_{L_s,L_u} between 0.5 and 3 and lower values for ε_{K,L_s}, which is quite close to our estimates.

Nevertheless, $\frac{1}{\varepsilon_{L_s,L_u}} = 1 - \sigma$ gives a measure of the impact of a variation of the relative supply of unskilled labor on the skill-premium in a partial equilibrium setting, while (4.2) shows that the steady-state effect of such a change is measured instead by the coefficient $1-\theta$: the estimated impact varies with the threshold and the sample and, not considering the primary school threshold, predicts for the full sample that a 10% increase in the $\frac{L_u}{L_s}$ ratio raises the steady-state skill-premium between 2.6% and 4.1%. On the other hand (4.2) also predicts the steady-state impact on the skill-premium of a variation in the effective saving rate: a 10% increase in $\frac{s}{\pi}$ has a near zero impact for the full sample in which there is weak evidence for CSC, while it generates a 3% increase of the skill premium in the non-OECD subsample, where CSC is effective.

- Regression analysis shows that observables have moderate explanatory power (R^2) with respect to cross-country dispersion of skill-premia: variation in the $\frac{L_u}{L_s}$ ratio and in effective saving rates $\frac{s}{\pi}$ can explain up to 27% of international variation in $\frac{\omega_s}{\omega_u}$ for the full sample an up to 32% for the non-OECD subsample, while they can explain very little for OECD countries. A great portion of observed dispersion is left unexplained, suggesting that unobservable forces have a role in shaping cross-country inequality. While the observation of cross-country income differences suggest that large productivity differences are needed, the dispersion of cross-country skill premia reinforces this perspective and calls for a theory of productivity differences able to simultaneously explain growth and inequality facts. Caselli and Coleman (2006) find, in a partial equilibrium framework, that the observed dispersion of skill-premia can be explained if skill-biased technological change is introduced: countries seem to use more efficiently the relative abundant factor, an observation that can be theoretically justified in an appropriate technology setting where a country technology choice depends on its factor endowment.
- The introduction of barriers to capital accumulation, contrary to the growth literature where it can change a lot in the empirical performance of some theoretical models, seems to have little effect, with the exception of the non-OECD subsample (Table 4 and 6): when barriers are introduced, the evidence for CSC increases (the difference $(\sigma - \theta)$ is estimated with more precision) and the R^2 grows a bit. This suggests that policy distorsions that discourage capital accumulation, as measured by the relative price of investment goods, have a direct influence on inequality in developing countries: the industrial transition they are experiencing, and in which CSC seems to have a role, raises both capital intensity and

the secondary or high scool thresholds should be used, which diminishes the amplitude of the observed skill-bias of technology differences.

the skill premium, so that lower barriers to capital accumulation increase wage inequality.

5. Conclusion

The CSC hypothesis links the skill premium to capital accumulation: observed cross-country differences in the skill premium are a natural testing ground for evaluating the existence of CSC. We derive a steady-state approximation of the skill premium as a linear function of observables that control the capital accumulation process, saving rates and barriers to investments. We find weak evidence in support for the CSC hypothesis for the full sample, but strong evidence in favour of CSC in the non-OECD subsample, strenghtening the results obtained by Papageorgiou and Chmelarova (2005). There is increasing evidence that developing countries are undergoing a transition similar to that documented by Goldin and Katz (1998) for the U.S. during the first part of the twentieth century, when the shift from the classical factory to continuous production processes seems to have been characterized by CSC. We also obtain estimates of the elasticities of substitution between couple of inputs, so that we can proceed to a robustness check of our analysis comparing them to estimates obtained in other studies. Estimated elasticities of substitutions are consistent with those reported in microeconomic studies only when the threshold separating unskilled and skilled labor is sufficiently high: defining as skilled those workers who have completed only a primary cycle of education seems to be too weak a requirement, leading to unrealistic degree of substitutability between couple of inputs. Finally, we are able to obtain an estimate of the steady-state or long run impact on the skill-premium of a change in the effective saving rate: for OECD countries where CSC is uneffective the impact is null, while in non-OECD countries, where CSC acts, a 10% increase in the effective saving rate it generates a 3% increase of the skill premium. As a final remark, obesrvables quantities are able to explain only up to 30% of observed cross-country dispersion of skill-premia, suggesting that unobservable forces have a role in shaping cross-country inequality. If the analysis of cross-country income differences suggest that large international productivity differences are needed, the dispersion of cross-country skill premia reinforces this perspective and calls for a theory of productivity differences able to simultaneously explain growth and inequality facts.

6. References

Acemoglu, D. (2003), "Patterns of Skill Premia," *Review of Economic Studies 70*, 199-230.

Barro, R.J. and J. Lee (2001), "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers, 53 (3)*, 541-563.

Barro, R. and Sala-i-Martin (1995), *Economic Growth,* McGraw-Hill, New York.

Caselli, F. and Coleman, W.J. (2006), "The World Technology Frontier", *American Economic Review Vol. 96 n° 3*, 499-522.

Duffy, J., C. Papageorgiou and F. Perez-Sebastian (2004), "Capital-Skill Complementarity? Evidence from a Panel of Countries", *Review of Economics and Statistics 86*, 327-344.

Durlauf, S.,Kourtellos, A. and Minkin, A. (2001), “The Local Solow Growth Model”, *European Economic Review 45*, 928-940.

Goldin, C. and L. Katz (1998), “The Origins of Capital-Skill Complementarity”, *Quarterly Journal of Economics 113*, 693-732.

Gollin, D. (2002), "Getting Income Shares Right”, *Journal of Political Economy 110(2),* 458-474.

Griliches, Z. (1969), “Capital-Skill Complementarity,” *Review of Economics and Statistics 5*, 465-68.

Hamermesh, D.S. (1993), *Labor Demand*, Princeton University Press, Princeton N.J.

Heston, A., Summers, R. and Aten, B. (2002), "Penn World Table Version 6.1", Center for International Comparisons at the University of Pennsylvania (CICUP).

Krusell, P., L. Ohanian, J.-V. Rios-Rull and G.L. Violante (2000), “Capital Skill Complementarity and Inequality: A Macroeconomic Analysis”, *Econometrica 68*, 1029-1053.

Lindquist, M.J. (2005), "Capital-Skill Complementarity and Inequality in Sweden", *Scandinavian Journal of Economics 107(4)*, 711-735.

Masanjala, W. and Papageorgiou, C. (2004), “The Solow Model with CES Technology: Nonlinearities and Parameter Heterogeneity”, *Journal of Applied Econometrics 19*, 171-201.

Papageorgiou, C. and Chmelarova, V. (2005), "Nonlinearities in Capital-Skill Complementarity", *Journal of Economic Growth 1,* 55-86.

Papageorgiou, C. and Saam, M. (2005) , "Two-Level CES Production Technology in the Solow and Diamond Growth Models", Working Paper, Louisiana State University.

Psacharopoulos, G. (1994). “Returns to Investment in Education: A Global Update,” *World Development 22*, 1325-1343.

Restuccia, D., Urrutia, C., (2001), "Relative Prices and Investment Rates", *Journal of Monetary Economics 47*, 93-121.

Sato, K. (1967), "A Two-Level Constant-Elasticity-of-Substitution Production Function", *Review of Economic Studies 43*, 201-218.

APPENDIX A

Steady-State Approximations

The logarithm of the steady-state level of total capital per unit of effective skilled labor is given by

$$\ln(\widetilde{x}^{*}) = \frac{1}{\sigma}\ln\left[\frac{(1-a)\left(\frac{L_u}{L_s}\right)^{\sigma}}{[\frac{\pi(\delta+g)}{s}]^{\sigma}-a}\right] \tag{A1}$$

which is a function $g(\sigma)$ of the elasticity parameter σ which measure the degree of substitutability between raw labor L_u and the physical-human capital aggregate X.

Approximating (A1) by a first-order Taylor expansion of $g(\sigma)$ in the neighborhood of $\sigma = 0$ gives

$$\ln(\widetilde{x}^{*}) \simeq \lim_{\sigma\to 0} g(\sigma) + \lim_{\sigma\to 0} g^{'}(\sigma) \tag{A2}$$

$$= 0 + \lim_{\sigma\to 0}\left\{ -g(\sigma) + \frac{\frac{\left([\frac{\pi(\delta+g)}{s}]^{\sigma}-a\right)(1-a)\left(\frac{L_u}{L_s}\right)^{\sigma}\ln\left(\frac{L_u}{L_s}\right)}{\left([\frac{\pi(\delta+g)}{s}]^{\sigma}-a\right)^{2}}}{\left[\frac{(1-a)\left(\frac{L_u}{L_s}\right)^{\sigma}}{[\frac{\pi(\delta+g)}{s}]^{\sigma}-a}\right]} - \frac{\frac{(1-a)\left(\frac{L_u}{L_s}\right)^{\sigma}[\frac{\pi(\delta+g)}{s}]^{\sigma}\ln[\frac{\pi(\delta+g)}{s}]}{\left([\frac{\pi(\delta+g)}{s}]^{\sigma}-a\right)^{2}}}{\left[\frac{(1-a)\left(\frac{L_u}{L_s}\right)^{\sigma}}{[\frac{\pi(\delta+g)}{s}]^{\sigma}-a}\right]} \right\}$$

$$= \ln\left(\frac{L_u}{L_s}\right) - \frac{1}{(1-a)}\ln[\frac{\pi(\delta+g)}{s}]$$

Notice also that in the limit for $\sigma \to 0$ the production function (7) becomes a Cobb-Douglas aggregate of the physical-human capital composite X and of unskilled labor L_u (Barro and Sala-i-Martin (1995, pag.56)):

$$\lim_{\sigma\to 0} Y_t = \lim_{\sigma\to 0} A_t\left\{\alpha X_t^{\sigma} + (1-\alpha)L_{u,t}^{\sigma}\right\}^{\frac{1}{\sigma}} = A_t X_t^{\alpha} L_u^{1-\alpha} \tag{A3}$$

and $(1-\alpha)$ becomes the share of GDP that goes to unskilled labor.

APPENDIX B

Data Description

Table B1 - Countries, Labor Stocks, Saving Rates and Barriers

Country	L_u^P	L_s^P	$\left(\frac{\omega_s}{\omega_u}\right)^P$	L_u^S	L_s^S	$\left(\frac{\omega_s}{\omega_u}\right)^S$	L_u^H	L_s^H	$\left(\frac{\omega_s}{\omega_u}\right)^H$	s	π	Year
Argentina	60	106	1.51	148	29	2.53	192	7	4.23	25.3	1.81	1989
Australia	17	129	1.24	85	56	1.63	146	15	2.13	31.5	1.08	1982
Bolivia	75	51	1.33	105	20	1.89	125	6	2.7	13.3	1.21	1989
Botswana	115	41	2.15	174	5	5.58	190	1	14.5	28.3	1.72	1979
Brazil	100	62	1.8	129	22	3.75	159	7	7.83	23.2	1.27	1989
Canada	7	134	1.23	83	56	1.6	159	6	2.07	23.3	1.23	1981
Chile	73	108	1.62	143	35	2.94	203	8	5.37	29.7	1.09	1989
China	65	50	1.22	105	13	1.57	124	1	2.01	20.3	1.69	1985
Colombia	83	76	1.75	138	22	3.53	180	5	7.1	18	1.35	1989
Costa Rica	83	77	1.55	126	28	2.67	156	10	4.6	14.7	1.61	1989
Cyprus	48	144	1.55	125	54	2.69	211	13	4.66	31.2	1.02	1984
Dominican Rep.	85	49	1.46	116	17	2.33	134	6	3.73	17.1	1.51	1989
Ecuador	69	107	1.6	126	39	2.89	171	13	5.22	24.4	0.85	1987
El Salvador	92	38	1.47	123	10	2.39	136	3	3.89	8	2.2	1990
France	46	112	1.49	130	33	2.46	183	7	4.06	26.2	1.04	1985
Ghana	84	35	1.4	123	5	2.15	130	1	3.29	9.1	2.02	1989
Greece	27	85	1.11	88	26	1.28	109	9	1.46	29.3	1.11	1985
Guatemala	98	43	1.81	133	11	3.82	151	3	8.05	8.8	1.84	1989
Honduras	102	75	2.02	152	21	4.87	217	3	11.75	13.8	1.56	1989
Hong Kong	38	99	1.28	97	39	1.73	151	6	2.35	19.9	1.54	1981
Hungary	37	89	1.29	111	22	1.47	128	8	1.83	16.7	1.75	1987
India	79	34	1.22	102	12	1.55	115	3	1.99	16.8	1.34	1981
Indonesia	85	72	1.97	177	11	4.62	222	1	10.8	13.9	1.33	1981
Israel	37	119	1.29	87	58	1.78	156	14	2.45	28.5	0.94	1979
Italy	43	66	1.1	91	20	1.23	110	4	1.38	24.9	1.03	1987
Jamaica	96	185	3.16	294	29	13.36	490	3	56.37	20.6	1.39	1989
Japan	28	119	1.3	106	43	1.79	152	12	2.48	36	1.21	1975
Kenya	109	34	1.93	159	4	4.38	166	1	9.93	17.4	1.68	1980

Table B1 - Continued

Country	L_u^P	L_s^P	$\left(\frac{\omega_s}{\omega_u}\right)^P$	L_u^S	L_s^S	$\left(\frac{\omega_s}{\omega_u}\right)^S$	L_u^H	L_s^H	$\left(\frac{\omega_s}{\omega_u}\right)^H$	s	π	Year
Malaysia	59	82	1.46	127	22	2.33	169	2	3.73	23.2	1.31	1979
Mexico	81	92	1.76	144	28	3.56	196	7	7.2	19.5	1.51	1984
Netherlands	24	128	1.34	119	40	1.95	169	10	2.82	25.8	1.14	1983
Nicaragua	91	40	1.47	114	15	2.39	126	6	3.89	14.5	1.96	1978
Pakistan	85	30	1.47	107	9	2.39	120	2	3.89	12.2	1.86	1979
Panama	63	139	1.73	142	47	3.43	227	11	6.81	26.1	1.23	1989
Paraguay	88	68	1.58	141	19	2.82	170	5	5	11.7	1.88	1989
Peru	65	75	1.38	106	30	2.07	139	10	3.11	12	1.45	1990
Philippines	46	97	1.38	103	37	2.05	141	13	3.06	14.9	1.87	1988
Poland	19	98	1.12	98	24	1.3	119	7	1.5	21.73	1.07	1986
Portugal	64	59	1.49	118	14	2.46	142	3	4.06	22.5	1.33	1985
Singapore	71	89	1.71	150	22	3.34	196	4	6.53	32.2	1.36	1974
South Korea	28	159	1.53	114	61	2.6	219	12	4.41	22.3	1.38	1986
Sri Lanka	51	76	1.32	117	18	1.88	149	1	2.66	14.8	2.62	1981
Sweden	28	133	1.31	78	67	1.83	170	13	2.55	24.5	1.15	1981
Switzerland	22	142	1.37	87	64	2.04	192	8	3.02	29.7	1.15	1987
Taiwan	36	97	1.27	95	37	1.72	143	7	2.32	20.7	1.50	1972
Thailand	87	65	1.52	143	16	2.55	152	8	4.29	18	1.72	1989
Tunisia	82	36	1.38	104	14	2.05	122	3	3.06	13.8	2.09	1980
UK	36	115	1.31	121	36	1.84	163	9	2.59	18.4	1.22	1972
USA	6	229	1.48	65	116	2.43	237	27	3.94	21.1	1	1989
Uruguay	62	97	1.47	139	28	2.39	176	7	3.89	11.8	2.60	1989
Venezuela	69	71	1.4	111	27	2.13	141	8	3.24	11.4	1.46	1989
West Germany	41	94	1.22	122	22	1.55	145	5	1.99	28.5	0.97	1988

Note: L_u^j and L_s^j for $j = P, S, H$ are the stocks of unskilled and skilled labor for the primary, secondary and high school threshold, taken from Caselli and Coleman (2006) and discussed in the main text. $\left(\frac{\omega_s}{\omega_u}\right)^j$ for $j = P, S, H$ is the skill premium for the primary, secondary and high school threshold, also taken from Caselli and Coleman (2006). s and π are respectively average investment share of GDP and average relative price of investment over consumption over the period 1960-1985, the former taken from Mankiw, Romer and Weil (1992), the latter computed from PWT 6.1."Year" refers to the time when the Mincerian coefficient used to calculate the skill premium has been estimated.

Table B2 - Descriptive stastistics - Full Sample

	Mean	Std. Dev.	Min	Max	Obs
$(L_u/L_s)^P$	0.99	0.83	0.026	3.21	52
$(L_u/L_s)^S$	6.95	7.43	0.56	39.75	52
$(L_u/L_s)^H$	45.28	50.86	8.78	222	52
s	20.41	6.9	8	36	52
π	1.47	0.4	0.85	2.62	52

Table B3 - Descriptive stastistics - OECD Subsample

	Mean	Std. Dev.	Min	Max	Obs
$(L_u/L_s)^P$	0.36	0.29	0.026	1.08	17
$(L_u/L_s)^S$	3.34	2.02	0.56	8.43	17
$(L_u/L_s)^H$	20.65	9.6	8.78	47.33	17
s	24.82	5.02	16.66	36	17
π	1.19	0.21	0.97	1.75	17

Table B3 - Descriptive stastistics - Non-OECD Subsample

	Mean	Std. Dev.	Min	Max	Obs
$(L_u/L_s)^P$	1.30	0.84	0.31	3.21	35
$(L_u/L_s)^S$	8.69	8.44	1.5	39.75	35
$(L_u/L_s)^H$	57.24	58.19	10.84	222	35
s	18.27	6.72	8	32.2	35
π	1.59	0.41	0.85	2.62	35

Note: $(L_u/L_s)^j$ for $j = P, S, H$ is relative supply of unskilled labor respectively for the primary, secondary and high school threshold (Source: Caselli and Coleman (2006)). s and π are respectively average investment share of GDP and average relative price of investment over consumption over the period 1960-1985, the former taken from Mankiw, Romer and Weil (1992), the latter computed from PWT 6.1.

Table 1 - Cross-country skill premia without barriers to capital accumulation ($\pi = 1$ for all countries in the sample)

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.2422378 (0.1898754)	0.044101 (0.4678897)	-0.2724844 (0.8573914)
$\ln\left(\frac{L_u}{L_s}\right)$	0.0740959^{**} (0.0291791)	0.2602965^{***} (0.0755516)	0.4196514^{***} (0.1451645)
$\ln(s)$	0.0603812 (0.0679365)	0.1417124 (0.1373441)	0.0698606 (0.1890973)
Implied θ	0.9259041^{***} (0.0291791)	0.7397035^{***} (0.0755516)	0.5803486^{***} (0.1451645)
Implied σ	0.9458299^{***} (0.0209356)	0.7864686^{***} (0.0610165)	0.6034026^{***} (0.1292275)
R^2	0.1254	0.2124	0.2700
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	18.18	4.67	2.51
ε_{K,L_s}	10.91	3.41	2.31
CSC	Yes (*weak*)	Yes (*weak*)	Yes (*weak*)
Obs.	52	52	52

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the $10\%, 5\%$ or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With ε_{K,L_u}, ε_{L_s,L_u} and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).

Table 2 - Cross-country skill premia with barriers to capital accumulation. Full Sample.

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.3349269*** (0.095667)	0.1867583 (0.3248323)	−0.285988 (0.6971067)
$\ln\left(\frac{L_u}{L_s}\right)$	0.07325 ** (0.0290676)	0.2691474*** (0.0805223)	0.4297168*** (0.1483474)
$\ln\left(\frac{s}{\pi}\right)$	0.0327603 (0.0411259)	0.1005613 (0.0964054)	0.0712556 (0.1345215)
Implied θ	0.926 75 *** (0.0290676)	0.730852 6*** (0.0805223)	0.5702832*** (0.1483474)
Implied σ	0.9375609*** (0.0218323)	0.7640379*** (0.0641453)	0.5937976*** (0.1325261)
R^2	0.1224	0.2141	0.2719
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	15.87	4.23	2.45
ε_{K,L_s}	12.33	3.43	2.25
CSC	Yes (weak)	Yes (weak)	Yes (weak)
Obs.	52	52	52

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With ε_{K,L_u}, ε_{L_s,L_u} and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).

Table 3 - Cross-country skill premia without barriers to capital accumulation ($\pi = 1$ for all countries in the sample). OECD countries.

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.734241 (0.5358456)	1.827379 (1.102006)	1.736924 (2.070185)
$\ln\left(\frac{L_u}{L_s}\right)$	0.0074225 (0.0384143)	-0.0577962 (0.0905088)	0.2024358 (0.2426112)
$\ln(s)$	-0.14015469 (0.1593802)	-0.3591214 (0.3412107)	-0.4269681 (0.5465613)
Implied θ	0.992577*** (0.0384143)	1. 057796*** (0.0905088)	0.7975642*** (0.2426112)
Implied σ	0.945906*** (0.0697545)	0.9382087*** (0.1617834)	0.6553838*** (0.267698)
R^2	0.0534	0.0688	0.0898
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	18.18	16.12	2.89
ε_{K,L_s}	178.40	-34.38	5.93
CSC	No (weak)	No (weak)	No (weak)
Obs.	17	17	17

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With ε_{K,L_u}, ε_{L_s,L_u} and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).

Table 4 - Cross-country skill premia without barriers to capital accumulation ($\pi = 1$ for all countries in the sample). Non-OECD countries.

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.0218487 (0.1896246)	-0.7159364 (0.4248941)	-0.8387523 (0.9218354)
$\ln\left(\frac{L_u}{L_s}\right)$	0.0689963 (0.0554479)	0.3193019*** (0.0682426)	0.3610596** (0.1541133)
$\ln(s)$	0.1474071** (0.0661237)	0.3939003*** (0.1345685)	0.3809114* (0.2126599)
Implied θ	09310037*** (0.0554479)	0.6806981*** (0.0682426)	0.6389404*** (0.1541133)
Implied σ	0.9800902*** (0.0477861)	0.8118668*** (0.0636234)	0.7657839*** (0.1397072)
R^2	0.0714	0.2882	0.2515
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	50	5.29	4.25
ε_{K,L_s}	-3.26	2.04	2.01
CSC	Yes	Yes	Yes (weak)
Obs.	35	35	35

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the $10\%, 5\%$ or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With ε_{K,L_u}, ε_{L_s,L_u} and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).

Table 5 - Cross-country skill premia with barriers to capital accumulation. OECD countries.

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.6470729 (0.4284682)	1.657689 (0.9471502)	1.573698 (1.79007)
$\ln\left(\frac{L_u}{L_s}\right)$	0.0015069 (0.03746)	−0.0754686 (0.0876136)	0.1832508 (0.2463101)
$\ln\left(\frac{s}{\pi}\right)$	−0.1218957 (0.1355466)	−0.3171386 (0.3036006)	−0.378183 (0.4614799)
Implied θ	0.998 493 1*** (0.03746)	1.075 469*** (0.0876136)	0.816 7492*** (0.2463101)
Implied σ	0.9582675*** (0.0582546)	0.9708128*** (0.1368016)	0.6907008*** (0.2480848)
R^2	0.0907	0.1200	0.1219
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	23.80	33.33	3.22
ε_{K,L_s}	738.09	−36.66	6.53
CSC	No (weak)	No (weak)	No (weak)
Obs.	17	17	17

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With $\varepsilon_{K,L_u}, \varepsilon_{L_s,L_u}$ and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).

Table 6 - Cross-country skill premia with barriers to capital accumulation. Non-OECD countries.

	Skill Threshold		
	Primary (P)	Secondary (S)	High School (H)
Constant	0.2043267^{**} (0.0971852)	-0.4040731 (0.2946899)	-0.6389869 (0.7441409)
$\ln\left(\frac{L_u}{L_s}\right)$	0.0700158 (0.056435)	0.3523^{***} (0.0749346)	0.3895418^{**} (0.1543199)
$\ln\left(\frac{s}{\pi}\right)$	0.0982335^{**} (0.0406909)	0.3100589^{***} (0.0923469)	0.3239627^{**} (0.1516815)
Implied θ	0.929984^{***} (0.056435)	0.6477^{***} (0.0749346)	$0.610\,458\,2^{***}$ (0.1543199)
Implied σ	0.9627254^{***} (0.050473)	0.7510427^{***} (0.0616687)	0.7184349^{***} (0.1366979)
R^2	0.0734	0.3245	0.2796
$\varepsilon_{K,L_u} = \varepsilon_{L_s,L_u}$	$26,31$	$4,01$	$3,54$
ε_{K,L_s}	$7,96$	$2,24$	$2,07$
CSC	Yes	Yes	Yes
Obs.	35	35	35

Note: OLS estimate of equation (4.3). Robust standard errors in parentheses.*,** and *** mean significantly different from 0 at the 10%, 5% or 1% level.

Weak means that the estimated difference $(\sigma - \theta)$ is not statistically significant different from 0 at the 1%, 5% or 10% level. With ε_{K,L_u}, ε_{L_s,L_u} and ε_{K,L_s} are denoted the implied Allen partial elasticities of substitution between each couple of inputs, computed using equation (4.4).