Universitat Autònoma de Barcelona

ESSAYS ON DEVELOPMENT ECONOMICS FROM A MACROECONOMIC PERSPECTIVE

Joaquin Naval Navarro

Universitat Autònoma de Barcelona

FACULTAT D'ECONOMIA I EMPRESA

DEPARTAMENT D'ECONOMIA I D'HISTÒRIA ECONÒMICA

PHD THESIS:

ESSAYS ON DEVELOPMENT ECONOMICS FROM A MACROECONOMIC PERSPECTIVE

Joaquin Naval Navarro

Supervised by:
Jordi Caballé Vilella
Chairman of the Department of Economics and Economic History at the Universitat Autònoma de Barcelona

Barcelona, Spring 2013

Agradecimientos

Esta tesis no hubiera sido posible sin el tiempo que me ha prestado Jordi Caballé Vilella.

Asimismo estoy endeudado con Jesús Fernández Huertas y Frédéric Docquier, tanto por sus comentarios a nivel académico como por transmitirme muchos ánimos a nivel personal.

Agradecer el apoyo recibido por Nuria Rodríguez en los inicios de esta aventura. A Tobias y Aysegul porque parte de este trabajo también es suyo. A Juan Carlos Conesa y al resto de profesores del departamento y del Instituto por lo que me enseñaron, muchas veces simplemente con el ejemplo. Y a Àngels y Mercè ya que han facilitado todo muchísimo.

Dar las gracias a todos mis compañeros de IDEA, en especial a Jan, Nuno, Pau y Rajesh, ya que hemos aprendido mucho los unos de los otros de manera totalmente desinteresada. Me escucharon, me ayudaron y me pidieron consejo.

También quiero agradecer a Dani, Ernesto, Font y Marc su apoyo incondicional. A los amigos de Estercuelgue y en general a todos los amig@s de Estercuel.

A Ferran, Francesc, Neus, Raulito, Vanesa y Víctor por los buenos ratos.

A Ángel, Eloy y el resto de amigos de Sant Andreu.

A mi familia, en especial a Alfonso, siempre tan orgulloso de su nieto.

Por último quiero agradecer a Laura por estar ahí en todo momento. A mi hermana Raquel, y a mis padres Joaquin y Pilar por apoyarme siempre, en todo y para todo. Gracias.

Contents

*“Compañera
usted sabe
que puede contar
conmigo
no hasta dos
ni hasta diez
sino contar
conmigo.”*

Fragmento de
Hagamos un trato
de Mario Benedetti

Chapter 1

Introduction

In this dissertation I explore forces that lie behind economic development over time and across countries. I use both theoretical and empirical approaches in order to study *i*) the effects of international migration on income inequality in the origin country of migrants, *ii*) the effects of the informal economy in the short and in the long run development of countries, and *iii*) why African countries have not experienced growth even though their levels of urban population and education have grown among the fastest in the world. In what follows, I present these three chapters of the dissertation in detail.

Chapter 2 discusses the emigration and economic inequality surrounding the emigrants of primarily developing countries. This chapter builds a theoretical framework inspired both by the work of Galor and Tsiddon (1997) explaining the evolution of inequality in origin countries, and Docquier, Rapoport and Shen´s (2010) model, which endogeneizes the decision to emigrate. Hence, the model departs from Kuznets' hypothesis on the inverted-U shape relationship between inequality and level of development, and highlights the effects of migration regarding the above relationship. The model also reflects how migration and education decisions are affected by liquidity constraints in migrants' home countries. Furthermore, it is able to replicate the stylized facts on relations between migration and education. These stylized facts are obtained using data from Docquier, Lowell and Marfouk (2009) which show that, overall, highly skilled workers on a global scale, emigrate more than medium and low skilled workers. In regards to developing countries however, a different pattern exists which is that, low skilled workers are in fact more likely to emigrate than medium skilled workers if migration costs are low, whereas the latter is reversed if migration costs are high.

In addition, it is established that the evolution of technology, together with migration costs determines the effects of migration on education, income and wealth inequality within different countries. The model associates migration and income in the origin country through remittances sent by migrants to their family. This relationship between migration and income allows the model to look at the effects of migration on inequality. Ergo, the predictions of the model claim that in the initial stages of development, migration rates increase and migration enlarges economic inequality over time for high migration costs. At more advanced stages of development or in the case that migration costs are low, migration rates and wealth inequality decline over time. To summarize, the main conclusions of the chapter suggest, the 'technological gap' between developed and developing countries and migration costs determine the effects of migration on inequality. It also demonstrates that wealth constraints play a crucial role both in the selection of migrants according to their education level, as well as in the effects of migration on wealth inequality.

Chapter 3 explores the informal sector, or shadow economy, as a source of multiplicity of equilibria in the long-run level of countries' development. The model establishes a relationship between wage inequality, human capital accumulation, child labor, and long-run growth, which is consistent with the following stylized facts: the informal sector is less productive and mainly uses low-skilled workers. As informality increases, the proportion of child labor is increased and the share of college educated individuals is decreased. The high level of wage inequality in developing countries has been found to be inconsistent with the predictions presented by theoretical models with complementarity between high- and low-skilled workers (research based on standard parameters).

The theoretical model also allows us to use quantitative theory to explore the implications of the above mentioned factors. Our model can generate transitory informality equilibria or informality-induced poverty traps. Once we bring the model to the data, its calibration reveals that the case for the poverty-trap hypothesis is strong: although informality serves to protect low-skilled workers from extreme poverty in the short-run, it prevents income convergence between developed and developing nations. Moreover, looking at the effects of a sudden elimination of informality, we obtain that it would induce severe welfare losses for poor people on the transition path but would generate welfare gains in the long run. In other words, the informal sector acts as a safety net for low-skilled workers in the short run that may harm the economy in the long run. The trade-off between child education and child income, combined with insufficient skill premium specifically relating to the informal sector, prevents households from investing in education. As a consequence, the aggregate level of human capital is far below what is required in order to increase an economy's development. Hence, we examine the effectiveness of different development policies in an attempt to exit the poverty trap. Our numerical experiments show that subsidizing education is the most cost-effective policy option, while subsidizing formal employment minimizes the length of the transition.

Chapter 4 is an attempt to explain why African countries have not experienced the same growth as more developed countries even though their rate and pace of human capital levels and urban population are among the highest in the world. The data reveals a positive association between human capital accumulation, urbanization and growth both over time and across countries. Taking the latter into account, an adjustment-cost hypothesis, which is comparatively more optimistic than other theories, is put forward to explain this paradox. According to this hypothesis, it is argued that low or negative social return to education in the short-run might be due to the transitory adjustment or urbanization costs. A simple model similar to Lucas (2009) is built based on this hypothesis. Furthermore, most of the structural parameters of the model are calibrated using panel data estimations from 1975 to 2000. In addition, the rest of structural and specific parameters are calibrated to match certain data moments. Once all parameters are calibrated, the model is used to compare simulations with the historical data on Africa and other regions to confirm the validity of the model. This was established by replicating the paths of human capital accumulation, urbanization and GDP per capita. The calibrated model predicts an economic take-off for the next few decades, however no full convergence with high-income countries is expected to be realized. Thus, both a positive and negative conclusion can be observed. Firstly, we can safely predict that African countries will, in the short-term, realize their latent growth potential but secondly that, African countries will never achieve the same status of wealth as more advanced economies.

References

Docquier, F., Rapoport, H., and I. Shen. 2010. "Remittances and inequality: A dynamic migration model." *Journal of Economic Inequality*, 8(2), pp. 197-220.

Docquier, F., Lowell, B.L., and A. Marfouk. 2009. "A gendered assessment of highly skilled emigration." *Population and Development Review*, 35(2), pp. 297-321.

Galor, O., and D. Tsiddon. 1997. "The Distribution of Human Capital and Economic Growth." *Journal of Economic Growth*, 2(1), pp. 93-124.

Kuznets, S. 1955. "Economic Growth and Income Inequality." *American Economic Review*, 45(1), pp. 1-28.

Lucas R. 2009. "Trade and the diffusion of the industrial revolution." *American Economic Journal: Macroeconomics",* 1(1), pp. 1-25.

Chapter 2

Migrant Selection and Inequality in Labor-Exporting Countries

2.1 Introduction

International labor flows arise as a consequence of the globalization process. As a matter of fact, labor movements affect both emigration and immigration countries. There can be changes in the labor supply, skill composition, wages and investment flows, among other outcomes. In this work, we discuss some effects of these movements on labor-exporting countries. These countries experience a reduction in their labor supply and an increase in capital flows due to remittances. These flows of labor and capital can alter economic inequality and growth in an open economy where capital moves freely while labor movements are subject to an exogenous fixed cost. We consider the migration decision as an alternative investment to produce economic growth, and study its consequences. Migration can increase investments in education and increase human capital acquisition. These changes in human capital have consequences on inequality in labor-sending developing countries. The questions that we will address in our analysis are the following: Do migration and remittances increase or decrease inequality in migrants' home countries? Do the inequality patterns change over time? Does migration affect the factors that generate inequality? We consider productivity differences to explain the different effects that migration and remittances can have on inequality.

It is increasingly accepted that "education and migration decisions are likely to be jointly determined", Hanson (2010, p. 239). For this reason, our baseline endogenous migration model expands upon Docquier et al.'s (2010) model where households choose production and migration as a single decision unit. In their model, remittances are the channel to equalize and maximize household utility, and migration costs are exogenous at a constant value. Unlike Docquier et al., we endogenize the decision concerning the generation of human capital within households through investment in education. Human capital units are obtained by means of a continuous and concave technology that presents indivisibilities in education investment. One of the main differences between our work and Docquier et al. (2010) is that we obtain migrants' earnings and remittances proportional to their education level. In their pioneering study, Lucas and Stark (1985) point out that income and remittances increase with education. They study the motivations to remit for the case of Botswana. Recently, Grogger and Hanson (2011) stated that migrants choose their destinations depending on their possible earnings, which suggests that migrants' wages highly depend on their human capital level.

Recent studies have pointed out an inverse U-shaped relation between migration and inequality as already suggested by Stark, Taylor and Yitzaki (1986). Mckenzie and Rapoport (2007) observe an inverted U-shaped relation between wealth and migration history. They consider the migrant networks to be a channel for reducing migration costs and creating the possibility for the poorest families to migrate and become richer than they would without migration. Docquier et al. (2010) also explain the inverse U-shaped relation between migration history and income inequality. They argue that remittances might increase income inequality and reduce wealth inequality in economies with little history of migration, whereas remittances decrease income and wealth inequality in economies with a long history of migration. We seek to complement these studies by considering technological progress and the technological gap as alternative explanations for the relation between migration and inequality. Therefore, we highlight the importance of development as an engine to influence inequality and the effects of migration on inequality.

The relation between economic inequality and growth has been widely studied since the seminal study by Galor and Zeira (1993). They use different interest rates for borrowers and lenders and indivisibilities in education to prove that the aggregate output and investment is affected by the initial distribution of wealth in the short and the long run. We develop a model similar to Galor and Tsiddon's (1997) to study the evolution of inequality and migration through technological process. Galor and Tsiddon (1997) consider the level of aggregate human capital to be the necessary input for the evolution of technology. The level of technology determines the evolution of inequality and the process of development. They study the evolution of human capital from parents to children and how the aggregate distribution of human capital evolves over time. The interplay between parents and children's education (local effect) and the state of technology and education (global effect) generate convergence towards an unequal distribution of human capital in low technological levels, and towards an equal distribution in high technological levels. Since we are interested in the effects of migration on economic inequality, we also consider how both distributions of income and wealth evolve over time, and how these distributions are affected by the migration investment and the technological gap between migrants' source and destination countries.

We use a combination of migration decisions and technological levels to characterize and compute remittances. This amount of money enables households to increase their level of consumption and investment in migrants' home countries. The option of migration as a family investment implies a trade-off between education and migration decisions because of financial constraints. If the return on education is higher than the return on migration, the model predicts that some families decide to invest in education and send a share of the family abroad. Due to the higher earnings that they receive abroad, educated migrants will remit more than migrants without education as Bollard et al. (2011) point out. We also take into account the size of the household that stays in the birth country, since the amount of remittances highly depends on the relatives who do not migrate (Bollard et al. 2011). Because of financial constraints, the interaction between migration and education returns in the home and destination countries determines the education level. Once remittances alleviate the liquidity constraints of the poorest and most uneducated households, they prefer to invest in education because of its higher returns. Moreover, the higher returns in the destination country increase the amount that households can invest in education, and a higher technological level in the home country is related to a lower minimum level of education that makes it profitable to become educated. The possibility of migration can increase the minimum amount of wealth to make it profitable to be educated.

As the development process evolves, we observe two different regimes of migration rates, income and wealth inequality. At early stages of migration, migration rates increase and the differences in wealth and income between poor and rich families expand. At more advanced stages of technology, migration rates decrease if migration costs are fixed. If migration costs are reduced, migration increases as long as it is profitable. If a technological barrier that generates convergence towards a single wealth distribution is overcome, the differences in income and wealth between rich and poor households are reduced because all the individuals become educated and converge into an equal society. We consider the possibility that migration accelerates this evolution. Remittances increase wealth inequality but also decrease the threshold technology that reduces inequality in the economy. As in Galor and Tsiddon (1997), these results are in line with Kuznets' hypothesis, which supports the fact that developing countries experience an increase in economic inequality, whereas more developed countries experience a reduction in economic inequality over time.

The remainder of the paper is organized as follows. In Section 2.2, we construct the model with and without migration. Section 2.3 presents some stylized facts. Section 2.4 presents the conclusions.

2.2 The model

2.2.1 Production

Consider a small open economy in discrete time where aggregate production of a single final good is given by a constant returns to scale neoclassical production function. In every period t, production takes place using aggregate physical capital K_t and aggregate human capital H_t. The final output depends on a factor productivity A_t related to the level of technology. We assume that productivity evolves according to an exogenous sequence $\{A_t\}_{t=0}^{\infty}$. Output Y_t produced at time t is

$$Y_t = A_t F(K_t, H_t) = A_t H_t f(k_t),$$

where $k_t = K_t/H_t$ is the physical capital to human capital ratio. The production function $f(k_t)$ is strictly increasing, strictly concave and satisfies the Inada conditions,

$$f(0) = 0, \lim_{x\to 0} f'(x) = +\infty, \lim_{x\to +\infty} f'(x) = 0.$$

We normalize the price of the final good to 1, i.e., we take the output good as the numéraire. In a competitive equilibrium each factor price equals its marginal product,

$$r_t = A_t F_K(K_t, H_t) = f'(k_t),$$

and

$$w_t = A_t F_H(K_t, H_t) = A_t \left(f(k_t) - f'(k_t)k_t\right) = A_t w(k_t),$$

where r_t is the return on physical capital and w_t is the return on human capital.

As a result of the assumption of a small open economy, physical capital flows into or out of the country so that the marginal product of physical capital is equal to the constant world interest rate $\overline{r}$. The constancy of the world interest rate implies that the physical to human capital ratio is fixed over time $K_t/H_t = k_t = \overline{k}$. Hence,

$$r_t = \overline{r}, \quad \text{and} \quad w_t = A_t\overline{w}.$$

The acquisition of human capital depends on investment in education. Individuals can acquire an education level e through wealth investment. We consider a continuous human capital technology as in Galor and Tsiddon (1997), Vidal (1998) and Ceroni (2001), among others. However, a minimum education investment is necessary to increase the human capital efficiency units. A share of the expenditure on goods is allocated to education. The human capital production function is given by

$$h(e) = \begin{cases} \mu & \text{if } e \le e_0 \\ \mu + (e - e_0)^{\eta} & \text{if } e > e_0 \end{cases}$$

where $\mu \ge 0$, $\eta \in (0, 1)$, and $e_0 \ge 0$. The human capital technology is such that $h(0) = \mu$, for all $e > e_0$, $h'(e) > 0$ and $h''(e) < 0$, and satisfies that

$\lim_{e \to e_0^+} h'(e) = \infty$ and $\lim_{e \to \infty} h'(e) = 0$.[1] Hence, the level of human capital is an increasing function of wealth investment with diminishing returns.

2.2.2 Economy without international labor mobility

The economy is modeled as an overlapping generation of dynasties. There is a continuum of households of size N composed of a continuum of individuals who live for two periods. Without loss of generality we assume that the size of the household is 1. Households differ in the initial distribution of wealth although we assume that all the members in the household equally share their wealth and education level.

In the first period, households receive a wealth endowment b_t^i from their parents. With this endowment, the household decides how much to invest in education e_t^i and savings s_t^i. In the second period, individuals work, consume, and make transfers to their offspring. Each individual has a single child, and parents retire at the end of the second period. There is no population growth.

Utility is derived from consumption and bequests to their offspring in the second period of life. The second-period utility function is

$$U_{t+1} = \log(c_{t+1}^i) + \gamma \log(b_{t+1}^i).$$

Therefore, the utility function exhibits "joy of giving" preferences since parents obtain utility from the transfers to their children rather than from the utility of their children. We impose that $\gamma \bar{r} < 1$, which implies that bequests do not grow unboundedly.

In period t, a household with inherited wealth b_t^i maximizes utility given prices $\bar{r}$, w_{t+1} subject to the following constraints:

$$s_t^i \geq 0; \tag{2.1}$$

$$e_t^i + s_t^i = b_t^i; \tag{2.2}$$

$$c_{t+1}^i + b_{t+1}^i = w_{t+1} h(e_t^i) + (1 + \bar{r}) s_t^i; \tag{2.3}$$

where constraint (2.1) states that households can invest in the international capital markets but they cannot borrow; constraint (2.2) says that households can invest their wealth endowment in education and/or savings; and constraint (2.3) means that households choose per capita consumption and bequests in the second period, which exhaust household income. At the same time, household income is the return on savings and salary, which depends on

[1] Among others, Galor and Tsiddon (1997) use this type of technology. They extend this technology to the case of there being diminishing complementarity between investment in wealth and parents' level of human capital.

the education level. Moreover, because of the functional form for the utility function, consumption and bequests are positive, i.e., $c_t > 0$ and $b_t > 0$.

Given the assumption regarding the utility function and the human capital production function, there is a single equilibrium solution for each amount of inherited wealth to the maximization problem. The homothetic preferences assumption over consumption and bequests implies that households maximize income, and spend a share $1/(1+\gamma)$ of income on consumption and a share $\gamma/(1+\gamma)$ on bequests.

Let $\underline{e}(A_t)$ be the minimum amount invested in education for which is worth becoming educated, so $\underline{e}(A_t) > e_0$. Therefore, $\underline{e}(A_t)$ is the minimum education investment for the wage income to be equal to the income without education, i.e.,

$$A_t\overline{w}h(\underline{e}(A_t)) = (1+\overline{r})\underline{e}(A_t) + A_t\overline{w}\mu. \tag{2.4}$$

Let $\overline{e}(A_t)$ be the maximum amount invested in education, $\overline{e}(A_t) \geq \underline{e}(A_t)$. It is the education investment in order for the return on education to be equal to the return on savings,

$$\overline{e}(A_t) = (h')^{-1}\left(\frac{1+\overline{r}}{A_t\overline{w}}\right) = e_0 + \left(\frac{\eta A_t\overline{w}}{1+\overline{r}}\right)^{\frac{1}{1-\eta}}. \tag{2.5}$$

We only consider sufficiently high technological levels in order to ensure the existence of $\underline{e}(A_t)$. Otherwise nobody chooses to become educated and the economy converges into a situation where nobody is educated. The human capital technology ensures the existence of $\overline{e}(A_t)$, since the marginal returns on education are decreasing from infinity to zero. We can observe from (2.4) that the minimum education level $\underline{e}(A_t)$ decreases as the level of technology A_t increases. In contrast, from (2.5) we observe that the maximum education level $\overline{e}(A_t)$ increases as the level of technology A_t increases.

When households maximize income, they choose the level of education and savings as a function of the inherited wealth.

a. If $b_t^i \leq \underline{e}(A_t)$, then $e_t^i = 0$ and $s_t^i = b_t^i$.
 If the wealth of the household is not high enough to make it profitable to become educated, they generate wealth without education.

b. If $\underline{e}(A_t) < b_t^i < \overline{e}(A_t)$, then $e_t^i = b_t^i$ and $s_t^i = 0$.
 If the wealth of the household is high enough to make it profitable to become educated, households choose the maximum possible level of education that they can pay for. As they cannot afford the optimal level of education, they do not save and invest all their wealth in education.

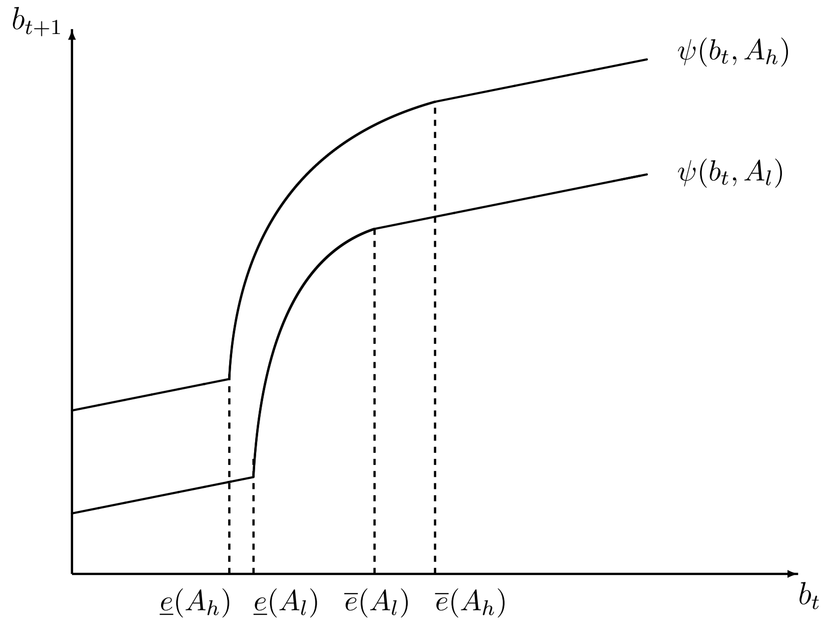
c. If $b_t^i \geq \overline{e}(A_t)$, then $e_t^i = \overline{e}(A_t)$ and $s_t^i = b_t^i - \overline{e}(A_t)$.
If the household wealth is higher than the optimal educational investment, households spend the optimal amount on education, and the rest of their wealth is saved and invested in international capital markets.

Therefore, given the technological level A_t and the level of inherited wealth b_t^i, we can obtain the optimal savings $s_t^i = s(b_t^i, A_t)$ and education $e_t^i = e(b_t^i, A_t)$ decisions, and the optimal human capital efficiency units $h_t^i = h(e_t^i, A_t) = h(e(b_t^i, A_t))$ that households supply.

The evolution of wealth within a dynasty

In the previous section we characterized the amount devoted to bequests for every dynasty as a function of the inherited wealth. At the same time, this amount is the inherited wealth of the next generation. Hence, the evolution of bequests characterizes the household wealth evolution. Figure 2.1 shows the evolution of bequests for two different technological levels: high and low, where subscript h denotes high and subscript l denotes low. As noticed before, the interval between the minimum and maximum values of education expands with the technological level A_t. Therefore, we can characterize the evolution of bequests within a dynasty as a function of their initial wealth.

Figure 2.1: Evolution of bequests across dynasties for A_h and A_l with $A_h > A_l$



The wealth dynamics of dynasties are given by

$$b_{t+1}^i = \psi(b_t^i, A_t)$$
$$= \begin{cases} \frac{\gamma}{1+\gamma}\left(A_t\overline{w}\mu + (1+\overline{r})b_t^i\right) & \text{if } b_t^i \leq \underline{e}(A_t) \\ \frac{\gamma}{1+\gamma}A_t\overline{w}(\mu + (b_t^i - e_0)^\eta) & \text{if } \underline{e}(A_t) < b_t^i \text{ and } \\ & b_t^i < \overline{e}(A_t) \\ \frac{\gamma}{1+\gamma}\left(A_t\overline{w}\left(\mu + \left(\frac{\eta A_t\overline{w}}{1+\overline{r}}\right)^{\frac{\eta}{1-\eta}}\right) + (1+\overline{r})(b_t^i - \overline{e}(A_t))\right) & \text{if } b_t^i \geq \overline{e}(A_t). \end{cases}$$

We can consider different technological levels. For the moment we only consider the case where the technological level is fixed over time at level A. Let $\widehat{A}$ be the technological level for which the long-run distribution converges towards a single steady state if the technological level is higher than $\widehat{A}$, and the long-run distribution converges to two steady states if the technological level is lower than $\widehat{A}$. The following proposition characterizes the evolution of dynasties when the technological level is fixed over time.[2]

Proposition 1. *Consider the dynamical system $b_{t+1}^i = \psi(b_t^i, A)$. If the dynamical system displays a single steady-state equilibrium, then*

$$\lim_{t\to\infty} b_t^i = b^s(A).$$

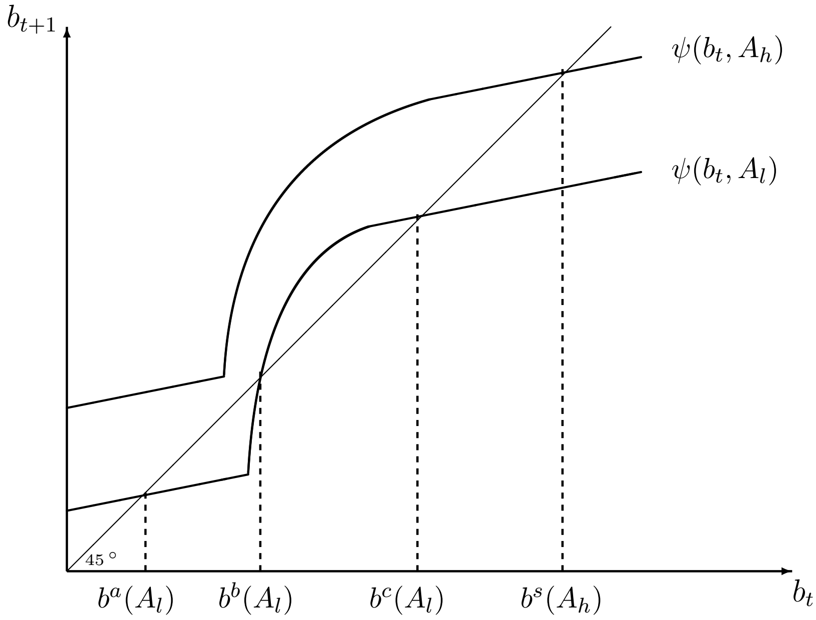
If the dynamical system displays three steady state equilibria, then

$$\lim_{t\to\infty} b_t^i = \begin{cases} b^a(A) & \textit{if } b_0^i \in [0, b^b(A)) \\ b^c(A) & \textit{if } b_0^i \in (b^b(A), \infty). \end{cases}$$

Proof. As can be seen in Figure 2.2, the dynamical system $b_{t+1}^i = \psi(b_t^i, A)$ converges to a single steady state or to three steady states. The technological level determines the number of steady states. If the level of technology $A \equiv A_h$ is higher than the threshold technology $\widehat{A}$ the system converges to a single steady state $b^s(A_h)$. Whereas, if the technological level $A \equiv A_l$ is lower or equal than $\widehat{A}$ the system has three steady states $b^a(A_l)$, $b^b(A_l)$, and $b^c(A_l)$, where $b^a(A_l)$ and $b^c(A_l)$ are stable and $b^b(A_l)$ is unstable. □

[2]Proposition equivalent to Proposition 2.2 in Galor and Tsiddon (1997).

Figure 2.2: Evolution of bequests across dynasties for $A_h > \widehat{A}$ and $A_l < \widehat{A}$



Given a technological level, the initial distribution of wealth characterizes the long-run distribution of wealth. If the dynamics of bequests are characterized by three steady states, the long-run distribution is polarized, whereas if the dynamics of bequests are characterized by a single steady state, the long-run distribution converges to a single point and the economy is equally distributed. Thus, there are two different wealth distribution patterns in the long run that depend on the technological level. Hence, if the technological process surpasses the technological level $\widehat{A}$, households converge towards an equally distributed economy, whereas if the technological process does not surpasses the technological level $\widehat{A}$, the economy converges towards a polarized and unequally distributed economy.

The evolution of the wealth distribution and technological progress

In the previous section we observed that the technological level determines wages, income and next generation wealth. Hence, the sequence of factor productivities $\{A_t\}_{t=0}^{\infty}$ characterizes the wealth evolution of dynasties and the aggregate wealth distribution over time. In this section we characterize the aggregate wealth distribution to obtain the evolution of wealth inequality in the economy.

Suppose that the density function $g_0(b_0^i)$ characterizes the wealth distribution of the parent generation in time 0. Because households cannot hold debt to their children in any period, i.e., $b_t > 0$ for all t, the positive real numbers are the support of the density function in all periods. We assume that the size of the continuum of households is constant and equal to N, therefore,

$$\int_0^\infty g_t(b_t^i)db_t^i = N, \quad t = 0, 1, 2, ...$$

For a stationary technological level A, the number of steady state equilibria of the wealth dynamics characterizes the aggregate human capital distribution in the long run. If there are multiple steady state equilibria, the number L^u of low-skilled households in the long run is

$$L^u = \int_0^{b^b(A)} g_t(b_0^i)db_t^i,$$

and the number L^s of high-skilled households is

$$L^s = \int_{b^b(A)}^\infty g_t(b_0^i)db_t^i.$$

The following proposition summarizes the long-run distribution for a level of technology A.[3]

Proposition 2. *Consider a stationary technological level A.*
If the dynamical system $b_{t+1}^i = \psi(b_t^i, A)$ displays a single steady state equilibrium $b^s(A)$, then

$$\lim_{t\to\infty} g_t(b_t^i) = b^s(A),$$

i.e., the long-run distribution of wealth is a point $b^s(A)$ of mass N.
If the dynamical system $b_{t+1}^i = \psi(b_t^i, A)$ displays three steady-state equilibria ($b^a(A)$ and $b^c(A)$ are locally stable, and $b^b(A)$ is locally unstable), then

$$\lim_{t\to\infty} g_t(b_t^i) = \begin{cases} b^a(A) & \text{if } b_0^i < b^a(A) \\ \\ b^c(A) & \text{if } b_0^i > b^b(A), \end{cases}$$

i.e., the long-run distribution of wealth is a two point distribution with a mass of L^u in $b^a(A)$ and a mass of L^s in $b^c(A)$.

Proof. It follows from Proposition 1. □

[3] Proposition equivalent to Proposition 3.1 in Galor and Tsiddon (1997).

For a given stationary level of technology, the initial distribution of wealth determines the long-run distribution of wealth. If the wealth dynamics display two locally stable steady states, the density function converges to a two mass points distribution. Hence, the economy converges towards a polarized economy with a mass L^u of poor households and a mass L^s of rich households, which may induce inequality in the distribution of income. However, if the wealth dynamics display a single steady state, the density function converges to a single mass point distribution. In this case, the economy converges towards a uniform and equally distributed economy composed of high-skilled households, regardless of the initial distribution.

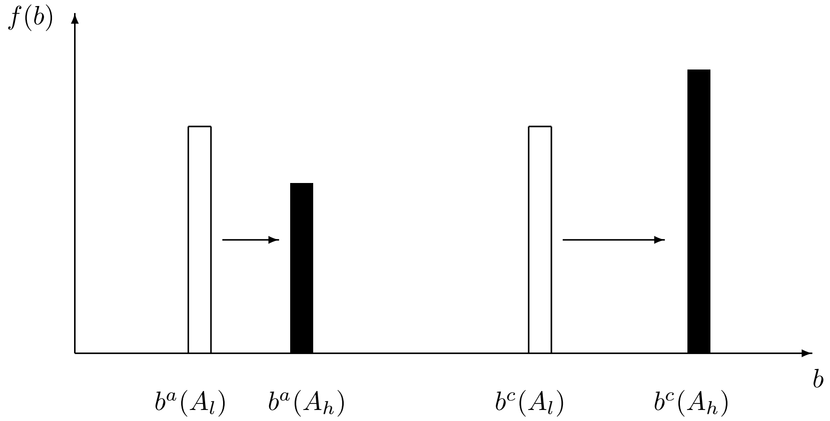
For every dynasty, given the initial level of wealth b_0^i and the technology level A_t, the function ψ characterizes the bequests that households leave to their children. Thus, given the sequence of technological levels $\{A_t\}_{t=0}^{\infty}$, the aggregate distribution in the next period follows from the distribution in the initial period, i.e., $g_1(b_1^i) = g_0(\psi^{-1}(b_1^i, A_0))$. In fact, we can generalize this relation for all the periods,

$$g_{t+1}(b_{t+1}^i) = g_t(\psi^{-1}(b_{t+1}^i, A_t)).$$

If technology is nondecreasing, i.e., $A_{t+1} \geq A_t$ for all t, for the same inherited wealth, households invest in education at least the same amount of wealth in the next period, because the returns on education are higher. Moreover, some of the households that do not invest in education, will invest in education in the next period because the minimum level of investment $\underline{e}(A_t)$ diminishes. Hence, the group of individuals L^s that tends to the higher-wealth steady state can increase because the wealth $b^b(A_t)$ that determines the size of the group in the long run diminishes as well, and, consequently, group L^u decreases. Therefore, the higher investment in education by the high-skilled group leads to increases in output in conjunction with an increase in the polarization of the distribution of wealth, human capital, and income. Figure 2.3 shows the long-run distribution of wealth for an economy that has experienced a TFP increase (from A_l to A_h). Moreover, the size of both unskilled and skilled workers is the same for the low TFP level. Thus, an increase in the TFP level increases the distance between the two mass points (increase in polarization) and the size of the high-skilled group (changes in inequality). In the long run, a change in TFP only changes the distance between the two mass points and not the size of each group. Therefore, inequality evolves as polarization does, inequality and polarization increase if TFP increases and is lower than the threshold level $\widehat{A}$, and inequality and polarization are reduced if TFP increases and surpasses $\widehat{A}$.

If we assume that there is a time period $\widehat{t}$ in which the technological level exceeds the threshold $\widehat{A}$, then, the dynamical system $b_{\widehat{t}+1}^i = \psi(b_t^i, A_{\widehat{t}})$ displays

Figure 2.3: Effects of an increase of the TFP in the long-run wealth distribution



a single steady state. From this period onwards, the economy converges towards a decrease in polarization in the distribution of wealth, human capital, and income.

Hence, in time periods in which the dynamical system $b^i_{t+1} = \psi(b^i_t, A_t)$ displays multiple steady-state equilibria, the distribution of wealth gravitates towards increased polarization, whereas in periods in which the dynamical system displays a single globally stable steady-state equilibrium the distribution of wealth gravitates towards diminished polarization. Therefore, in the long run, inequality increases or decreases if the technological threshold $\widehat{A}$ has or has not been overcome.

2.2.3 Economy with international labor mobility

In the previous section we assumed that labor was immobile. In the following sections we let households have access to migration. Households can achieve a higher wage in a destination country. Migration is an alternative investment that enables households to increase their income. This extra source of income can mitigate the borrowing constraints to invest in education for the next generation. We assume that the destination country has an equivalent production function except for the technological level. We also assume that the level of technology in the destination country is the technology frontier $\overline{A}$ and $\overline{A} \geq \widehat{A}$, then inequality decreases in the destination country. The home country's productivity evolves according to the sequence $\{A_t\}_{t=0}^{\infty}$. We treat the households' home country as a developing country, whereas the destination country is treated as a developed one. This assumption is in line with Kuznets' hypothesis: more advanced economies tend to be more equally distributed than developing ones

For simplicity's sake, we assume that the wage in the destination country does not change with migration. The arrival of new workers does not affect the labor share in the receiving economy. We make this assumption because we are interested in the effects in migrants' home countries. Furthermore, many studies reveal small effects or even no effects of migration on wages in the host country, although part of the literature still does not agree.[4]

In the previous section, households received a wealth endowment b from their parents in the first period. With this endowment, households decided education e and savings s. Now, they also decide the share m of the household that migrates at the end of the first period. In the second period, individuals consume and bequeath to their offspring. The share of the household can be related to the number of people who move, and the time spent in the migration country. If the household migrates, it has to pay a fixed cost F proportionate to the share of the household that migrates.[5]

Migrants also decide the quantity to remit that maximizes the household utility. Remittances R correspond to the direct transfers that migrants send to their family. All members of the household pay for the costs of migration, and migrants remit to compensate for these costs.

As in the economy without migration, utility is derived from consumption and bequests to offspring in the second period of life. Individuals decide the share m_t of the household that migrates, and utility depends on whether there is a share of the household that migrates or not. The second-period utility function is

$$U_{t+1} = m_t \log(c_{t+1}^{i,f}) + (1-m_t)\log(c_{t+1}^{i,h}) + \gamma\left(m_t \log(b_{t+1}^{i,f}) + (1-m_t)\log(b_{t+1}^{i,h})\right),$$

where superscript f denotes foreign consumption and bequests, those of migrants, and superscript h denotes home consumption and bequests, those of non-migrants. This utility function puts weights on the migrant and non-migrant part of the household, proportional to the size of the household that migrates. The weights equate the migrant and the non-migrant per capita consumption and bequests in equilibrium.[6]

In period t, a household with inherited wealth b_t^i maximizes utility given prices $\overline{r}$, $\overline{w}$ and technological levels A_t and $\overline{A}$ subject to the following con-

[4]Some examples of no effects of migration on wages are: Friedberg and Hunt (1995), Ottaviano and Peri (2012). An example of effects of migration on low-skilled wages is Borjas (2003).

[5]This type of cost is used by Docquier et al. (2010) in a dynamic framework and McKenzie and Rapoport (2007) in a static one.

[6]Docquier et al. (2010) use the same result but do not write the utility.

straints:

$$s_t^i \geq 0; \tag{2.6}$$

$$e_t^i + s_t^i + m_t^i F = b_t^i; \tag{2.7}$$

$$m_t^i \left(c_{t+1}^{i,f} + b_{t+1}^{i,f}\right) + R_{t+1}^i = m_t^i \left(\overline{A}\overline{w}h(e_t^i) + (1+\overline{r})s_t^i\right), \qquad \text{if } m_t^i > 0; \tag{2.8}$$

$$(1-m_t^i) \left(c_{t+1}^{i,h} + b_{t+1}^{i,h}\right) = (1-m_t^i) \left(A_{t+1}\overline{w}h(e_t^i) + (1+\overline{r})s_t^i\right) + R_{t+1}^i, \quad \text{if } m_t^i < 1; \tag{2.9}$$

$$m_t^i \in [0,1]. \tag{2.10}$$

Constraint (2.6) says that households can invest in international capital markets but they cannot borrow; (2.7) states that households can invest their wealth endowment in education, savings, and migration. Migration costs are proportional to the share of the household that migrates; (2.8) means that if there is a share of the household that migrates, the migrant share decides per capita consumption, bequests and remittances in the second period that exhaust the income in the destination country; constraint (2.9) means that, if the household has not fully migrated, the non-migrant share decides per capita consumption and bequests in the second period that exhaust the income in the birth country and the received remittances from the migrant share; and constraint (2.10) states that the migration share should be feasible.[7]

The assumption that households maximize utility as a unit implies that individuals within a household share costs and benefits. Let c_{t+1}^i be a share $1/(1+\gamma)$ of household income, i.e.,

$$c_{t+1}^i = \frac{1}{1+\gamma} \left(\left(\overline{A}m_t + A_{t+1}(1-m_t)\right)\overline{w}h(e) + (1+r_t)s_t\right), \tag{2.11}$$

and let b_{t+1}^i be the remaining income, i.e.,

$$b_{t+1}^i = \frac{\gamma}{1+\gamma} \left(\left(\overline{A}m_t + A_{t+1}(1-m_t)\right)\overline{w}h(e) + (1+r_t)s_t\right). \tag{2.12}$$

In the case of a share of the household migrating, i.e., for $m_t^i \in (0,1)$, we obtain

$$c_{t+1}^{i,h} = c_{t+1}^{i,f} \equiv c_{t+1}^i, \qquad \text{and} \qquad b_{t+1}^{i,h} = b_{t+1}^{i,f} \equiv b_{t+1}^i,$$

[7] One of the classical dichotomies in the migration literature is whether migrants maximize relative or absolute differences in income. The model uses absolute differences in wages and not relative differences to explain migration incentives. Grogger and Hanson (2011) highlight how absolute differences explain migration determinants better than relative differences as opposed to Borjas (1987).

the levels of consumption and bequests are equal for the individuals who remain in the birth country and the ones who migrate. In the case of the entire household migrating ($m = 1$), then there is neither consumption nor bequests at home and

$$c_{t+1}^{i,f} \equiv c_{t+1}^{i} \quad \text{and} \quad b_{t+1}^{i,f} \equiv b_{t+1}^{i}.$$

If nobody in the household migrates ($m = 0$), then there is neither consumption nor bequests abroad and

$$c_{t+1}^{i,h} \equiv c_{t+1}^{i} \quad \text{and} \quad b_{t+1}^{i,h} \equiv b_{t+1}^{i}.$$

In the case of a share of the household migrating, remittances are the channel to equalize the level of consumption at home and abroad. This amount of money that migrants send back home allows households to increase and equate the level of consumption and utility. In equilibrium,

$$R_{t+1}^{i} = m_t^i(1 - m_t^i)(\overline{A} - A_t)\overline{w}h(e_t^i), \tag{2.13}$$

remittances are a concave function of the migration share, and increase with the technological gap $\overline{A}-A_t$ between destination and source countries.[8] Moreover, note that if $m = 0$ or $m = 1$, then $R_{t+1}^i = 0$.

The education, savings and migration decisions are taken to maximize the second period household income. The previous problem can be compactly written as

$$\max_{m_t^i, e_t^i, s_t^i} m_t^i \overline{A}\overline{w}h(e_t^i) + (1 - m_t^i)A_t\overline{w}h(e_t^i) + (1 + \overline{r})s_t^i$$

subject to

$$s_t^i \geq 0,$$
$$e_t^i + s_t^i + Fm_t^i = b_t^i,$$
$$m_t^i \in [0, 1].$$

As in the economy without migration, there are education levels that determine which individuals invest in education and what the maximum amount they invest is. Let $\underline{e}(\overline{A} - A_t)$ be the minimum amount invested in education that makes it profitable to become educated. And let $\overline{e}(\overline{A} - A_t)$ be the maximum amount invested in education. The maximum amount invested in education is at least as high as without migration when households consider

[8] Expression (2.13) is equivalent to Docquier et al.'s (2010) expression with a quadratic production function. The expression is derived combining 2.8, 2.11 and 2.12.

migration to be an alternative investment. In fact, it is higher as long as migration is profitable for individuals. The possibility of higher returns in the destination country enlarges education investment, in other words, migration encourages higher education. In contrast, migration can be a disincentive for low levels of education. The minimum amount that makes it profitable to become educated is higher or equal to what it is without migration. This is because the possibility of migration makes households devote some resources to migration instead of starting to become educated.

In the case of poor households finding it optimal to invest in savings, this could be because it is only profitable to migrate to highly educated households, or because migration is not profitable to anyone. In the case of migration returns being higher than savings returns, the optimal migration m_t^i of a household with inherited wealth b_t^i can be obtained from the maximization problem above as the maximum of the function

$$\Phi(m_t^i|b_t^i) = (A_t + m_t^i(\overline{A} - A_t))h(b_t^i - Fm_t^i),$$

where $m_t^i \in [0, \min\{b_t^i/F, 1\}]$. The function Φ takes into account the income produced when there are no savings, and households can invest in education and/or migration. Its domain reflects the liquidity constraints that households face.

We can distinguish two parts in the function that characterize investment in migration. Figure 2.4 presents the function Φ for different domains. If $b_t^i \leq e_0$, this function is linear in m_t^i since the returns on education are constant and households can only invest in migration. Although in Figure 2.4 it is increasing, it could be increasing or decreasing depending on the technological gap, as we discuss later. If $b_t^i > e_0$, the function is linear for $m_t^i > (b_t^i - e_0)/F$, and it is strictly concave for $m_t^i \leq (b_t^i - e_0)/F$.

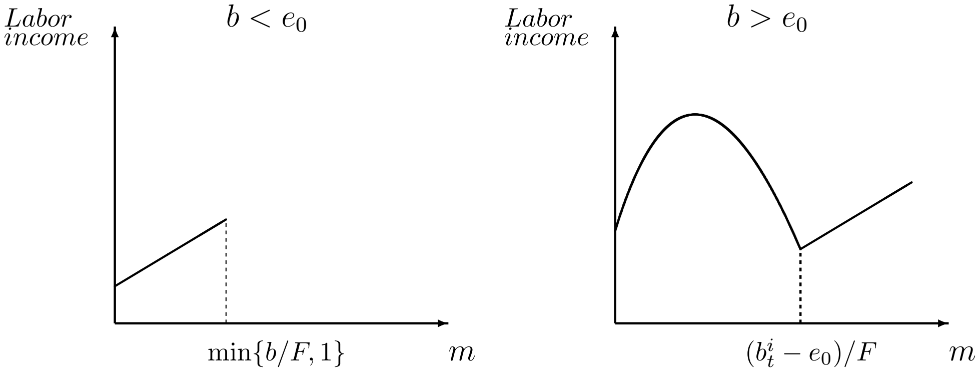
Therefore, the function Φ defines a function ϕ_m that gives the optimal migration investment for every amount of inherited wealth, i.e., $m_t^i = \phi_m(b_t^i)$.[9] In the case of there being two maximums we assume that the function ϕ_m gives the smallest migration share.

We can apply the same argument to the investment in education. Let ϕ_e be the function that gives the optimal education investment given the

[9] The maximum exists because Φ is continuous and its domain $\left[0, \min\{b_t^i/F, 1\}\right]$ is a compact set. Note that if we apply the Implicit Function Theorem to the strictly concave part, the migration share increases as bequests do, i.e.,

$$\frac{\partial \phi_m(b_t^i)}{\partial b_t^i} = \frac{(\overline{A} - A_t)h'(b_t^i - Fm_t^i) - F(m_t^i\overline{A} + (1 - m_t^i)A_t)h''(b_t^i - Fm_t^i)}{2(\overline{A} - A_t)h'(b_t^i - Fm_t^i) - F^2(m_t^i\overline{A} + (1 - m_t^i)A_t)h''(b_t^i - Fm_t^i)} \geq 0.$$

Figure 2.4: The function Φ



Note: the function Φ is not defined for values of m higher than $\min\{b_t^i/F, 1\}$.

inherited wealth, i.e., $e_t^i = \phi_e(b_t^i)$.[10]

We can distinguish different optimal decisions that depend on the technological gap to migration costs ratio. We can consider three different regimes: High-, Low- and Very low-migration regime. The returns on migration, education and savings characterize these regimes.

The High-migration regime occurs when the difference in technological levels is large enough and makes it worthwhile for all households to migrate. Migration costs are lower than the return on migration even for uneducated households,

$$(\overline{A} - A_t)\overline{w}\mu > (1 + \overline{r})F.$$

The Low-migration regime occurs when the difference in technological levels is not large enough, and it is only worthwhile for highly educated households to migrate. Migration costs are higher or equal than the return on migration for uneducated households, and the return on migration is higher than the migration costs and savings for highly educated households,

$$(\overline{A} - A_t)\overline{w}\mu \leq (1 + \overline{r})F,$$

[10]Note that the investment in education increases with wealth. If we use the Implicit Function Theorem we obtain

$$\frac{\partial \phi_e(b_t^i)}{\partial b_t^i} = \frac{\eta}{(1-\eta)} \frac{(e_t^i - e_0)^\eta}{h(e_t^i)} = \frac{\eta}{(1-\eta)} \frac{h(e_t^i) - \mu}{h(e_t^i)} > 0.$$

and

$$\overline{A}\overline{w}h(\overline{e}(\overline{A})) - A_t\overline{w}h(\overline{e}(A_t)) > (1+\overline{r})(F + \overline{e}(\overline{A}) - \overline{e}(A_t)).$$

Finally, the Very-low-migration regime is the difference in technology levels which makes it not worthwhile for any household to migrate. The return on migration is lower or equal than migration costs and savings even for highly educated households,

$$\overline{A}\overline{w}h(\overline{e}(\overline{A})) - A_t\overline{w}h(\overline{e}(A_t)) \leq (1+\overline{r})(F + \overline{e}(\overline{A}) - \overline{e}(A_t)).$$

The Very-low-migration regime is equivalent to the economy without labor mobility because migration is not profitable. The economy is at a high technological level and households do not benefit from migration. We will not consider this case in the remainder of the study since we are interested in the effects of migration.

1. High-migration regime.
   This case corresponds to countries with a low productivity level and/or with very low migration costs. The optimal choices of households depend on the inherited wealth of the household. The function Φ determines the optimal migration level. Because the poorest households profit from migration, the slope is positive in its linear part. Hence, the maximum of Φ is $\min\{b_t^i/F, 1\}$ if the maximum lies in the linear part. If the maximum lies in the strictly concave part, the maximum satisfies $\Phi'(m_t^i|b_t^i) = 0$ (or $m_t^i = 0$ if $\Phi'(m_t^i|b_t^i) \neq 0$, for all $m_t^i \in [0, (b_t^i - e_0)/F]$).

   In the high-migration regime we can distinguish two subcases. As can be seen in Figure 2.4 there can be a bequest level that has two maximums. We assume that in this case households prefer to send the smallest share possible to the destination country. This bequest level determines the minimum amount $\underline{e}(\overline{A} - A_t)$ invested in education. However, it could also be the case that the technological gap is so high that households want to migrate as much as possible, i.e., the maximum is $\min\{b_t^i/F, 1\}$. In this case, households always prefer to migrate than to become educated because the opportunities in their country are too low compared with the destination country. We assume that migration costs F are high enough as to ensure that not all the households prefer to migrate before they become educated.

   Households' choices can be summarized as follows:

   a. If $b_t^i \leq \underline{e}(\overline{A} - A_t)$, then $m_t^i = b_t^i/F$, $e_t^i = 0$, and $s_t^i = 0$.
   The poorest households do not save and invest all their wealth in migration.

b. If $\underline{e}(\overline{A}-A_t) < b_t^i < \overline{e}(\overline{A}-A_t)+F$ then $m_t^i = \phi_m(b_t^i) = (b_t^i - \phi_e(b_t^i))/F$, $e_t^i = \phi_e(b_t^i)$, and $s_t^i = 0$.
If the wealth of the household is high enough, households choose the optimal combination of education and migration in relation to their wealth. However, as they cannot afford the optimal level of education and entirely migrate, they do not save and invest all their wealth in education and migration.

c. If $b_t^i \geq \overline{e}(\overline{A}-A_t)+F$ then $m_t^i = 1$, $e_t^i = \overline{e}(\overline{A})$, and $s_t^i = b_t^i - \overline{e}(\overline{A}) - F$.
This case is for the unconstrained households who can choose the optimal level of education, migrate and still have some wealth to be saved.

Poor households invest all their wealth in migration. If the returns on education are higher than the returns on migration, the migration share falls and households start to invest in education. In the case of education being profitable, both education and the share of migration increase with wealth until households achieve their maximum optimal values. Then, households start saving.

2. Low-migration regime.
This case corresponds to countries with a high productivity level and high migration costs. In this case, only the households that are wealthy enough profit from migration. The poorest households do not migrate and save their endowments. The home technology determines the minimum education level. In this case, the slope of the linear part in the Φ function is negative, which prevents households from migrating unless they are educated.

Households' optimal decisions are the following:

a. If $b_t^i \leq \underline{e}(\overline{A} - A_t)$, then $s_t^i = b_t^i$, $e_t^i = 0$, and $m_t^i = 0$.
This is the case with the poorest households. If the wealth of the household is not high enough for it to be profitable to become educated or migrate, these households invest all their wealth in savings.

b. If $\underline{e}(\overline{A} - A_t) < b_t^i < \overline{e}(\overline{A} - A_t) + F$, then $s_t^i = 0$, $e_t^i = \phi_e(b_t^i)$, and $m_t^i = \phi_m(b_t^i)$.
If the wealth of the household is high enough, households choose the optimal combination of education and migration in relation to their wealth. There will be households that choose to become educated and do not migrate. As they cannot afford the optimal level of education and entirely migrate, they do not save and invest all their wealth in education and migration.

c. If $b_t^i \geq \overline{e}(\overline{A} - A_t) + F$ then $e_t^i = \overline{e}(\overline{A})$, $s_t^i = b_t^i - \overline{e}(\overline{A}) - F$, and $m_t^i = 1$. This case is for the unconstrained households who can choose the optimal level of education, migrate and still have some wealth to be saved.

We have observed that, if households invest in education and migration, the investment in education and the share of the household that migrates increase as the inherited wealth increases.

The following proposition summarizes the evolution of migration rates.

Proposition 3. *In an economy with international labor mobility,*

(i) migration is profitable if the technological gap to migration costs ratio $\overline{A} - A_t/F$ *is high enough;*

(ii) if the technological gap to migration costs ratio is low, only wealthy and highly educated households migrate and the migration share m_t *increases with wealth;*

(iii) if the technological gap to migration costs ratio is high, the migration share of uneducated households does not decrease with wealth unless the household becomes educated. If the household is educated, migration and education increase with wealth.

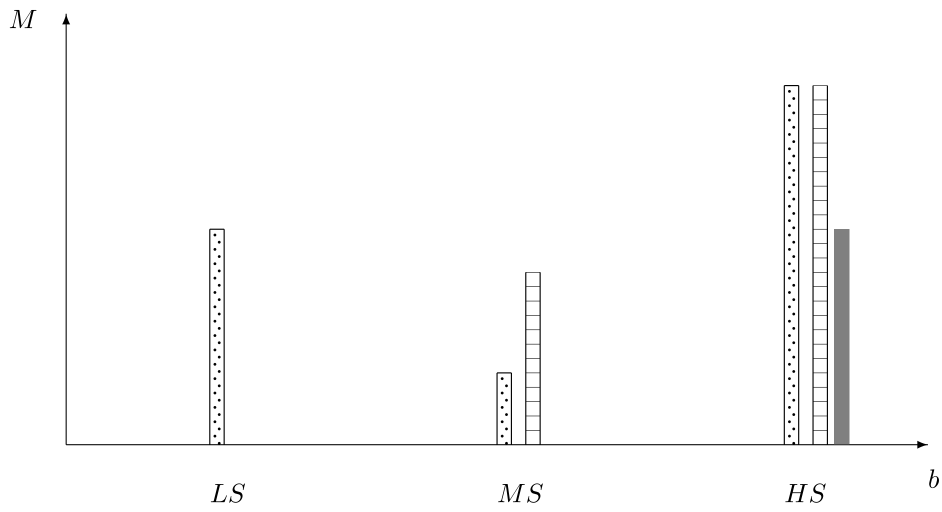
Proof. It follows from the optimal decisions of households above. □

This proposition characterizes the size and selection of migrants according to their level of education, and states that size and selection depend on the technological gap to migration costs ratio. Countries characterized by a low ratio select individuals with high skills to migrate, and countries characterized by a high ratio select individuals with high and low skills. Figure 2.5 illustrates the shape of selection and size of migration for different migration costs and a fixed technological gap in a given period. The gray color depicts a situation of high migration costs and, therefore, there is only migration of the highly skilled. The horizontal lines correspond to medium migration costs and migration increases as skills increase. Meanwhile, the dotes depict low migration costs and, therefore, migration is higher for the low and highly skilled than for the medium skilled.[11]

From the households' optimal decisions we can obtain aggregate migration rates. For fixed migration costs, if the technological gap is reduced,

[11]The gray color is named J-shape in the empirical part, the horizontal lines I-shape, and the dots U-shape.

Figure 2.5: Migrant size and selection for a fixed technological gap and different migration costs



less households benefit from migration, therefore, migration rates become lower. Similarly, for a fixed technological level, a reduction in migration costs increases the amount of households that can benefit from migration and increases the aggregate rate of migration. The following corollary summarizes the results.

Corollary 1. *Aggregate migration rates decline if the technological process reduces the technological gap for given migration costs F. Moreover, a reduction in migration costs increases aggregate migration rates for a given technological gap $\overline{A} - A_t$.*

Proof. Let R be the technological gap to migration costs ratio $(\overline{A} - A_t)/F$. The result follows from the fact that $\phi_m(b_t^i, R_1) \leq \phi_m(b_t^i, R_2)$ if and only if $R_1 \geq R_2$. □

The migration literature emphasizes the role of past migration as a way of reducing migration costs through networks. A reduction in migration costs changes the ratio that determines the size and the individuals who decide to migrate, higher migration reduces the ratio of high to low skilled migrants (Docquier et al. 2010). In our model, a reduction in migration costs can change countries from Low- to High-migration regimes. Hence, a reduction in migration costs in our model can explain why some countries change their pattern in the selection of migrants, countries with a high technological level close to the technology frontier can be in the high ratio due to a reduction in migration costs. A change from Low- to High-migration regime reduces the rate of high to low skilled migrants.

The evolution of wealth and income in an economy with international labor mobility

As in the economy without migration, the evolution of bequests determines the evolution of the dynasty's wealth. The possibility of migration generating an extra source of income can modify the evolution of bequests. In the economy without migration we can characterize the dynamics of bequests as a function of inherited wealth and the level of technology. In this case, we also take into account the level of technology in the destination country and the migration costs.

For the High- and Low-migration regimes, the wealth dynamics are given by:

$$b_{t+1}^i = \psi(b_t^i; A_t, \overline{A}, F) = \begin{cases} \frac{\gamma}{1+\gamma}\left(W(b_t^i) + R(b_t^i)\right) & \text{if } b_t^i \leq \underline{e}(\overline{A} - A_t) \\ \frac{\gamma}{1+\gamma} W(b_t^i) & \text{if } \underline{e}(\overline{A} - A_t) < b_t^i < \overline{e}(\overline{A} - A_t) + F \\ \frac{\gamma}{1+\gamma}\left(W(b_t^i) + R(b_t^i)\right) & \text{if } b_t^i \geq \overline{e}(\overline{A} - A_t) + F, \end{cases}$$

where

$$W(b_t^i) = \left(A_t + (\overline{A} - A_t)\phi_m(b_t^i)\right) \overline{w}(\mu + (\phi_e(b_t^i) - e_0)^\eta)$$

is labor income and

$$R(b_t^i) = \left(b_t^i - \phi_m(b_t^i)F - \phi_e(b_t^i)\right)(1 + \overline{r})$$

is capital income.

In the economy without migration, given a technological level, the initial distribution of wealth characterizes the long-run distribution of wealth. If the dynamics of bequests display three steady states, the long-run distribution is polarized, whereas if the dynamics of bequests display a single steady state, the long-run distribution converges to a single point. Thus, there are two different wealth distribution patterns in the long run that depend on the technological level.

In the economy without migration, $\widehat{A}$ is the threshold technology that determines whether the dynamical system $b_{t+1}^i = \psi(b_t^i, A)$ has one or two stable steady states. Let $\widehat{A}_m$ be the threshold technology that determines whether the dynamical system $b_{t+1}^i = \psi(b_t^i, A, \overline{A}, F)$ has one or two stable steady states.

Figure 2.6: Comparison of the evolution of bequests across dynasties

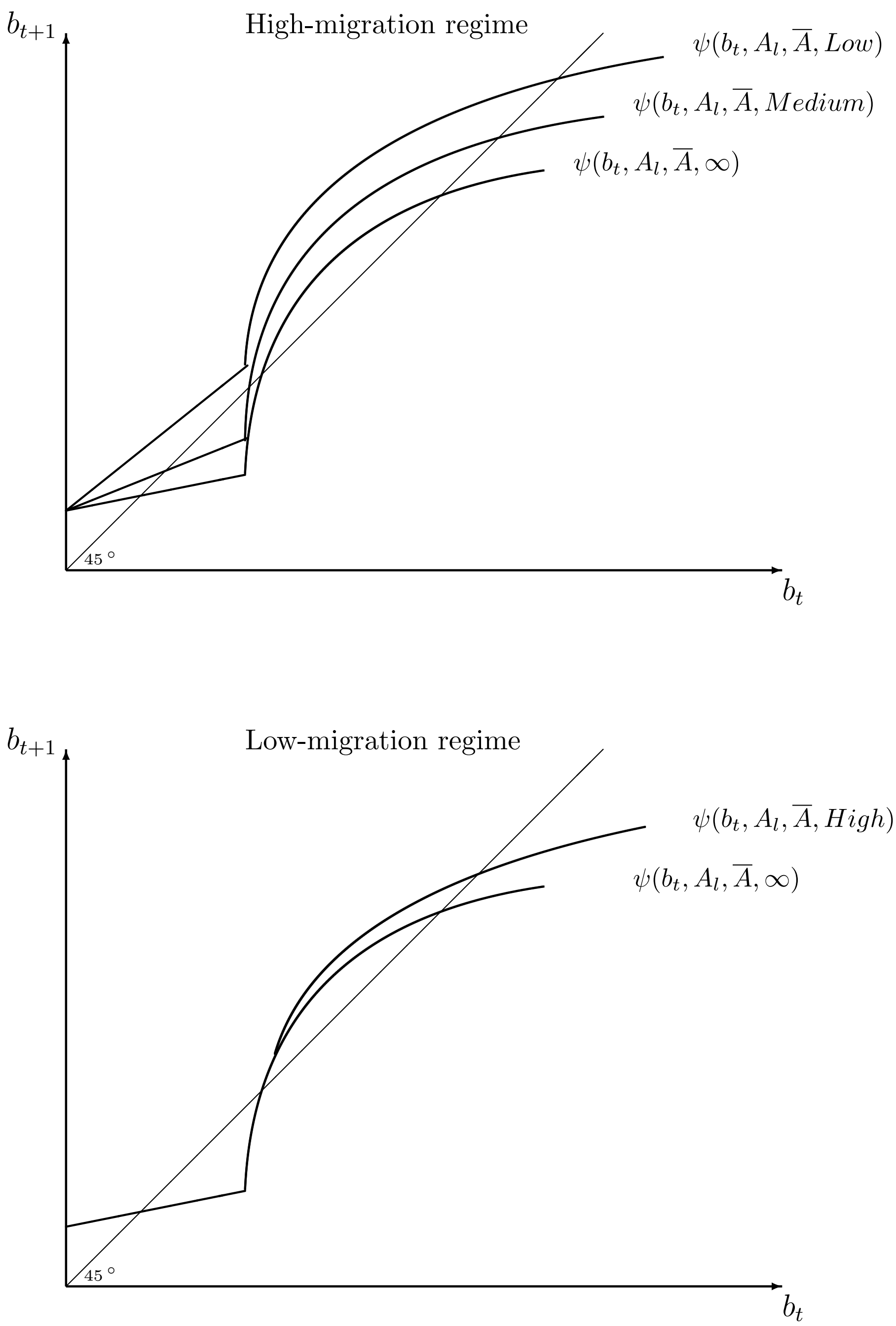
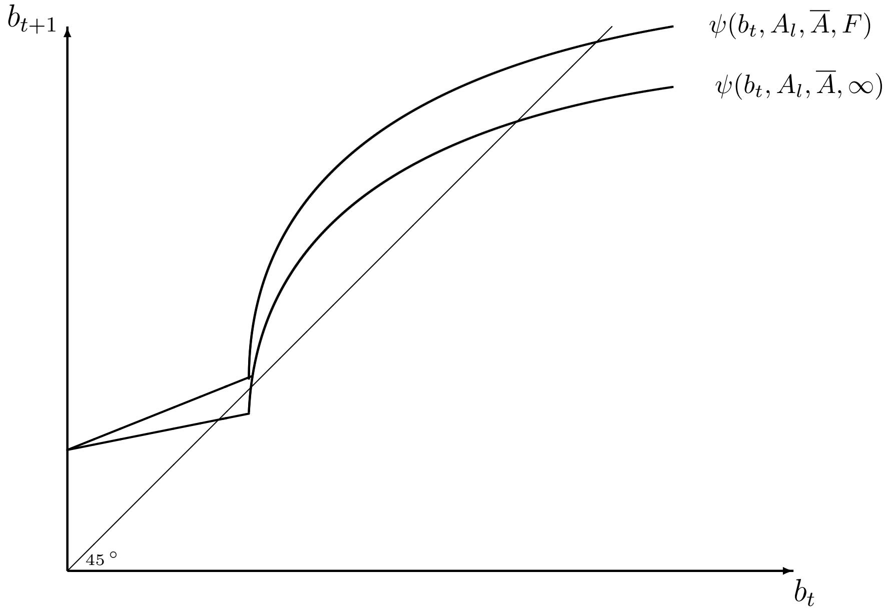


Figure 2.6 compares the evolution of bequests in an economy with a low productivity level, i.e., the dynamical system displays two stable steady-state points. The function $\psi(b_t, A_l, \overline{A}, \infty)$ shows the evolution without migration while the others show the evolution with migration. Two different migration regimes are presented, High-migration regime on the left side and Low-migration regime on the right. As we can observe, migration is profitable for all households in the High-migration regime whereas it is only profitable for educated households in the Low-migration regime. Thus, bequests left to children increase in the presence of migration as long as it is profitable. Since the productivity level is equal in both scenarios ($\psi(b_t, A_l, \overline{A}, \infty)$ is the same) what determines whether the economy is in the High- or Low-migration regime are the migration costs. Lower migration costs are related to the High-migration regime. Moreover, within the High-migration regime $\psi(b_t, A_l, \overline{A}, Low)$ shows lower migration costs than the costs of $\psi(b_t, A_l, \overline{A}, High)$.[12]

Figure 2.7: Comparison of the evolution of bequests across dynasties



We have observed that the economy without access to migration con-

[12]Galor and Tsiddon (1997) use a numerical example to illustrate the evolution in the dynamical system that relates parents and children's human capital. This example presents diminishing complementarity between resources invested and parents' human capital. In our scenario, diminishing migration costs due to past migration leads to a similar type of movement in the dynamical system of bequests.

verges to a polarized or equally distributed economy. We have also found that with access to migration the distribution of bequests within a dynasty changes. These changes affect the wealth and income distributions in the short and long run. The higher amount of bequests left to the next generation generates higher inequality in the distribution of wealth and income. Nevertheless, this higher amount reduces the technological threshold necessary to reduce polarization as shown in Figure 2.7. As in the economy without migration, we observe that in time periods in which the dynamical system $b^i_{t+1} = \psi(b^i_t, A_t, \overline{A}, F)$ displays multiple steady-state equilibria, the distribution of wealth and income gravitates towards increased polarization, whereas in periods in which the dynamical system displays a single globally stable steady-state equilibrium the distributions of wealth and income gravitates towards diminished polarization.

Proposition 4. *In an economy with international labor mobility, the following holds:*

(i) Inequality increases over time in countries with low technological levels and high migration costs.

(ii) Inequality decreases over time in countries with high technological levels or low migration costs.

(iii) Migration reduces the technological level that determines whether the economy increases or decreases economic inequality.

Proof. Part (i) follows from Figure 2.6. Part (ii) follows from Proposition 2 and Figure 2.6. Part (iii) follows from Figure 2.6.

□

2.3 Stylized facts

This section makes a link between the theoretical work presented above and some facts observed in the data. First, we present data sources and the relation of data with the model variables. Second, we contrast the size of migration in the data with the model predictions. The next section deals with the selection of migrants, what we mean by selection shape and what conclusions can be obtained. And finally, we comment on the relation between development and inequality, and the effects that migration can have on inequality.

2.3.1 Data

The main variables of interest are inequality, level of development, migration costs, and bilateral data on migration. Data on inequality are obtained from WIID (World Income Inequality Database) version 2.0c which extends the dataset developed by Deninger and Squire (1996) and computes the Gini index. The level of GDP per worker and the level of GDP per capita from World Bank indicators are a proxy for development. From Glick and Rose (2002) we obtain the log of distance between countries, whether countries have a common language, share colonial origins, or have common borders that we use as a proxy for migration costs. Bilateral data on migration come from Docquier, Lowell and Marfouk (2009). We use DLM to refer to the Docquier, Lowell and Marfouk database in the remainder of the paper.

DLM provides information on migration to 30 OECD countries and South Africa disaggregated by countries of origin and three educational levels for the years 1990 and 2000. This dataset considers migrants above 25 years old; for this reason, we can use it as a proxy for migrants who have finished their education. Nevertheless, DLM badly captures illegal immigration, except in the USA, therefore, it is expected to underestimate migration of the low-skilled.

There are also some other variables such as the human capital Gini index. We draw it from Castelló-Climent and Domenech (2002). Human capital Gini captures differences in the distribution of human capital across countries and time. Castelló-Climent and Domenech (2002) use educational acquisition from Barro and Lee's (1996) dataset to explain human capital distributions. We have constructed other indicator variables such as countries that signed the Schengen agreement, whether countries are in Africa or Latin America, and whether the host country is English speaking to improve the migration costs variable.

2.3.2 Migration size

There are several testable predictions of the model that are related to previous migration literature. Mayda (2010) and Grogger and Hanson (2011), among others, use different datasets on bilateral migration flows to OECD countries to explain the determinants of migration size. The main forces driving migration are differences in income and migration costs. At the same time, Beine et al. (2011) use diasporas as a channel to explain the network effects of reducing migration costs and the flow of migration. Nevertheless, we did not endogenize migration costs through previous migration in the previous sections to simplify the exercise, the introduction of decreasing costs would not change the main conclusions of the paper. We run both type of regressions with and without migrant networks. Similar to the previous works above we use differences in productivity levels that explain income differences and migration costs. In line with previous findings, our theoretical model predicts that higher differences in GDP per worker (or per capita) and lower migration costs increase migration. To highlight these results, we run a simple cross-country OLS regression of bilateral migration flows of the following form

$$\begin{aligned}\ln(m(i,j,t)) =& cte + a_1 * \ln(TechGap(i,j,t)) + a_2 * \ln(MigCost) \\ &+ a_3 * \ln(WealthLevel(i,t)),\end{aligned}$$

where $m(i,j,t)$ is the migration rate from country i to country j, i.e., the amount of people above 25 years from country i in country j divided by the population from country i. Population is estimated by the sum of the labor force participation and the total emigrants from country i.

There are many countries with no migration relation between them. To avoid the high number of zeros in the stock of migrants from one country to other, we use logarithms in order for countries with no migration relation not to be taken into account in the regression. Mayda (2010) and Grogger and Hanson (2011), among others, use this approach. Bertoli and Fernández-Huertas (2012) have developed another approach which requires more data to be used. According to Bertoli and Fernández-Huertas the coefficients from the regressions obtained following the former approach are biased. The results provide information on the direction of the effects of the variables considered but the magnitudes are not correct. Hence, from these regressions we obtain qualitative results to test the theoretical section but results are not valid quantitatively. In our case, we cannot use the latter approach because country fixed effects are needed to solve the problem. The inclusion of them would not allow us to identify the role of country-specific costs in determining the size of migration. There can also be other different types of problem, such as measurement errors or omitted variables. For example, the use of GDP per worker differences may not reflect the technological gap accurately enough.

Corollary 1 states that migration should increase with the technological gap to migration costs ratio. Hence, we expect the signs of coefficients to be $a_1 > 0$ and $a_2 < 0$. Proposition 3 also suggests that migration increases as the level of wealth does, i.e., we expect $a_3 > 0$. In Table 2.1 we can observe regression results that are in line with our predictions. The first two columns regress the log of migration rate from country i to country j to the log of technological gap (difference in GDP per worker), the log of average wealth level (GDP per worker in the source country), and migration costs (measured

Table 2.1: Bilateral migration rates

Dependent variable: ln(migration rate from country i to country j)					
ln(TecGap_pw)	.521	.523	.255	.603	.133
	.041**	.041**	.019**	.046**	.017**
ln(GDP_pw_source)	.930	.921	.146	.883	.046
	.041**	.042**	.019**	.061**	.022*
Human Capital Gini				-2.094	-.241
				.287**	.108*
border	1.187	1.184	.154	.516	.165
	.237**	.237**	.098	.303*	.099+
common language	1.452	1.450	.187	1.350	.148
	.113**	.113**	.051**	.123**	.042**
colony	2.376	2.382	.066	2.219	-.011
	.197**	.197**	.088	.211**	.072
ln(distance)	-.904	-.884	-.296	-1.069	-.107
	.045**	.050**	.020**	.057**	.020**
anglophone_dest	2.301	2.305	.347	2.459	.202
	.082**	.082**	.039**	.097**	.035**
Schengen		.117			
		.129			
ln(migration rate in 1990)			.857		.916
			.006**		.006**
constant	-16.752	-16.876	-2.549	-14.788	-1.341
	.787**	.798**	.364**	1.025**	.364**
number of observations	4413	4413	2174	2755	1351
R^2	0.3730	0.3731	0.9333	0.4644	0.9683

Notes: Standard errors are presented under each estimated coefficient; ** significant at 1% level, * significant at 5%, and + significant at 10%. The dependent variable is calculated as the amount of people above 25 years from country i in country j divided by the population from country i. Data sources: Docquier, Lowell and Marfouk (2009) for bilateral migration data; The World Bank indicators for GDP data; Castelló-Climent and Domenech (2002) for human capital gini: and Glick and Rose (2002) for the remainder of the variables.

by the log of distance from country i to country j and dummy variables with value 1 if the countries share a border, have a common language, have some colonial linkage, the destination country is English speaking, or the origin country has signed the Schengen agreement).

The third column uses 1990 migration rates to explain 2000 migration rates. If we use 1990 migration rates, the fit of the sample increases and still maintains the signs of the regression. However, half of the sample is lost when we include 1990 migration. Contrary to previous works, the dummy variable that indicates whether countries signed the Schengen agreement is not significant. Moreover, 1990 migration makes dummies for common border and colonial linkages to lose its significance.

In the last two columns in Table 2.1, we add the human capital Gini index in the origin country of migrants. As predicted by the model, a higher index leads to a reduction in migration due to wealth constraints that prevent individuals from migrating. In our model, a higher Gini denotes high educational costs and/or higher wealth differences, whereas a low Gini can reflect low educational costs and/or low wealth differences. The former implies a reduction in migration and the latter implies an increase in migration. Indeed, the reasoning behind educational costs and human capital inequality is as follows: given an aggregate human capital level and wealth distribution, higher educational costs are related to more people having less education and less people having more education than a country with lower educational costs; equivalently, given an aggregate human capital level and educational costs, a more polarized wealth distribution is related to more people having less education and less people having more education than a country with an egalitarian wealth distribution.[13]

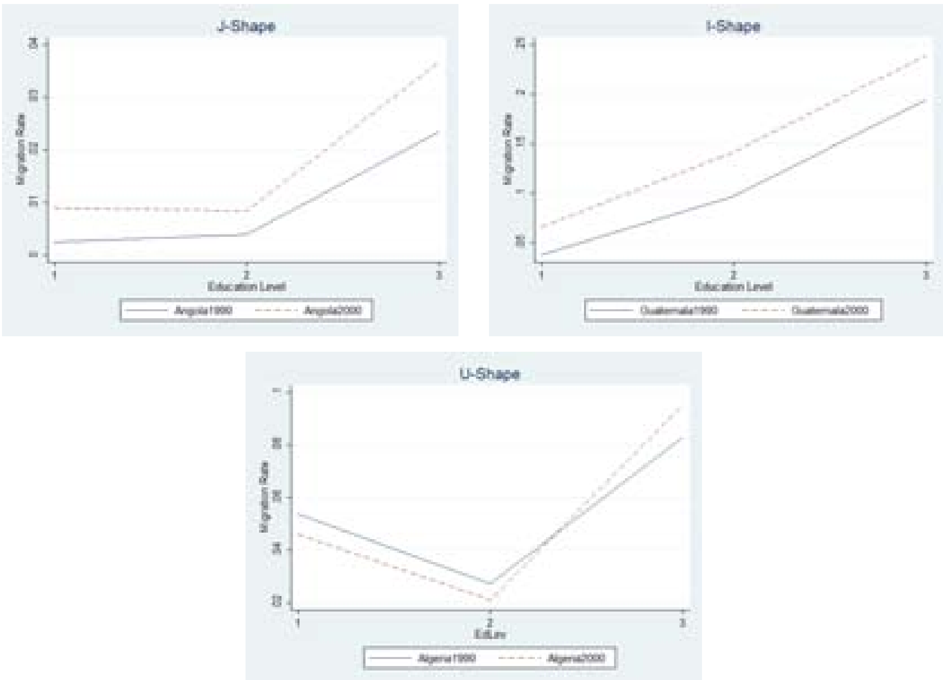
2.3.3 Selection

Migrant selection refers to the skills composition of migrants. Previous literature distinguishes between two types of workers –high- and low-skilled– in terms of their educational level. Previous literature states that there is positive (negative) selection of migrants if the ratio of high- to low-skilled migrants is higher (lower) than the same ratio for non-migrants. Here, we also consider the medium-skilled migrants to understand the composition and the wealth effects over selection better. In their characterization of migrant selection, Grogger and Hanson (2011) do not consider the medium-educated, instead they consider individuals with more or less than secondary education

[13] The same reasoning applies if we consider ability randomly and independently distributed among individuals.

to be a proxy for high- and low-skilled individuals, respectively. However, Caponi (2010) studies the shape of selection of Mexican migrants to the US on four educational levels: no education, primary, secondary, and tertiary education.[14] Similar to Caponi, we can look at different countries using the DLM dataset on three educational levels. Therefore, we refer to migrant selection as the shape of selection.

Figure 2.8: Migrant selection shapes



The DLM dataset distinguishes three general education groups: less than secondary education or low-skilled migrants; migrants with secondary education or medium-skilled migrants; and, more than secondary education or high-skilled migrants. We obtain three migration selection patterns with the exception of a few countries.[15] Let us denote each selection pattern: J-shape if migration is very low for individuals with secondary or less education com-

[14]Caponi (2010) explains the U-shaped relation between migration and education by the loss of human capital in migration and the altruism towards future generations.

[15]Some countries have a higher migration rate among individuals with secondary education than among other groups. These countries are Albania, Ecuador and Portugal for 1990 and 2000, Finland and Luxemburg for 1990, and El Salvador, Suriname and Venezuela for 2000 out of the 195 different countries considered.

pared to the highly skilled ones; I-shape if migration increases with education; and U-shape if individuals with secondary education migrate less than the remainder. Figure 2.8 shows an example of each possible shape.

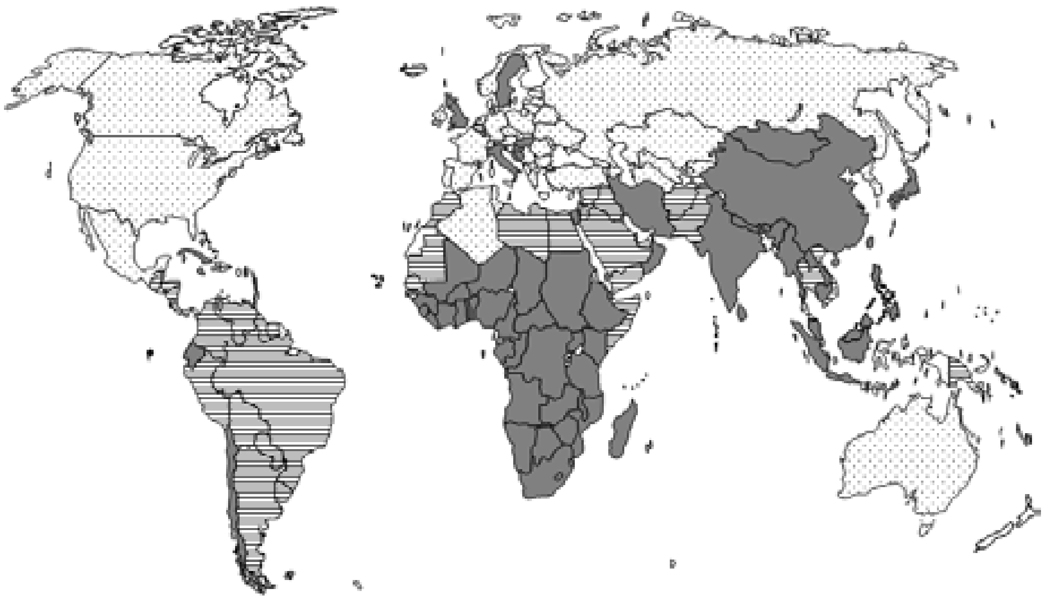
In order to classify countries in each group we should define the groups. In the J-shaped group there are countries with low rates of migration for the low and medium skilled in comparison to the rates of migration for the highly skilled. Thus, we include in the J-shaped group countries in which the difference between migration rates of low and medium skilled is lower than 10% of the difference between the highest (high skilled) and lowest (medium or low skilled) rates of migration. In this group there are countries such as Chile, the UK, and most of the sub-Saharan and Asian countries. In the I-shaped group there are countries where there is a strictly increasing relation between the migration rate and the education level. Hence, we include in the I-shaped group countries in which the migration rates increase with education and the difference between the low- and the medium-skilled migration rate is higher than 10% of the difference between the low- and high-skilled rates of migration. In this group there are countries such as Belgium, Thailand, and most Latin American and Arabian countries. Finally, in the U-shaped group there are countries with a low migration rate for the medium skilled compared to the low and high skilled, i.e., the medium skilled have the lowest migration rate, and the difference of migration rates between low and medium skilled is higher than 10% of the difference between medium and high skilled. Some examples are Mexico, Algeria, Turkey, and most of the OECD, former USSR, and former Yugoslavian countries.

In Figure 2.9 we can observe the distribution of migration shapes around the world. In the figure, J-shaped countries are colored in gray, I-shaped countries have horizontal lines, and U-shaped countries have dots. In addition, these shapes are highly persistent over time, there are only a few countries that change their shapes between 1990 and 2000.[16]

Figure 2.9 highlights the importance of wealth constraints in order to determine migrant selection. Although there may be other explanations, wealth constraints (due to low levels of development and high migration costs) seem

[16]In 1990 there were 86, 61 and 43 out of 190 countries with U-, I- and J-shape respectively. In 2000 there were 83, 69, and 37 out of 189. We exclude the 8 countries that present a different pattern. In fact, 146 countries out of 188 maintain the shape, 16 change from J to I (Bosnia and Herzegovina, Cambodia, Central African Republic, Chile, Indonesia, Kiribati, Laos, Liberia, Malawi, Mali, Marshall Islands, Federated States of Micronesia, Seychelles, Sierra Leone, Solomon Islands, and Zimbabwe), 10 countries from J to U (Australia, Azerbaijan, Cyprus, Estonia, Georgia, Latvia, Moldova, Monaco, Spain, and Uzbekistan), 7 from / to J (Egypt, Jordan, New Zealand, Pakistan, Qatar, Saudi Arabia, and United Arab Emirates), 4 from J to U (Croatia, Cuba, Italy, and Sweden), 2 from U to I (France, Switzerland), and 2 from I to U (Liechtenstein and San Marino).

Figure 2.9: Migrant selection in 1990



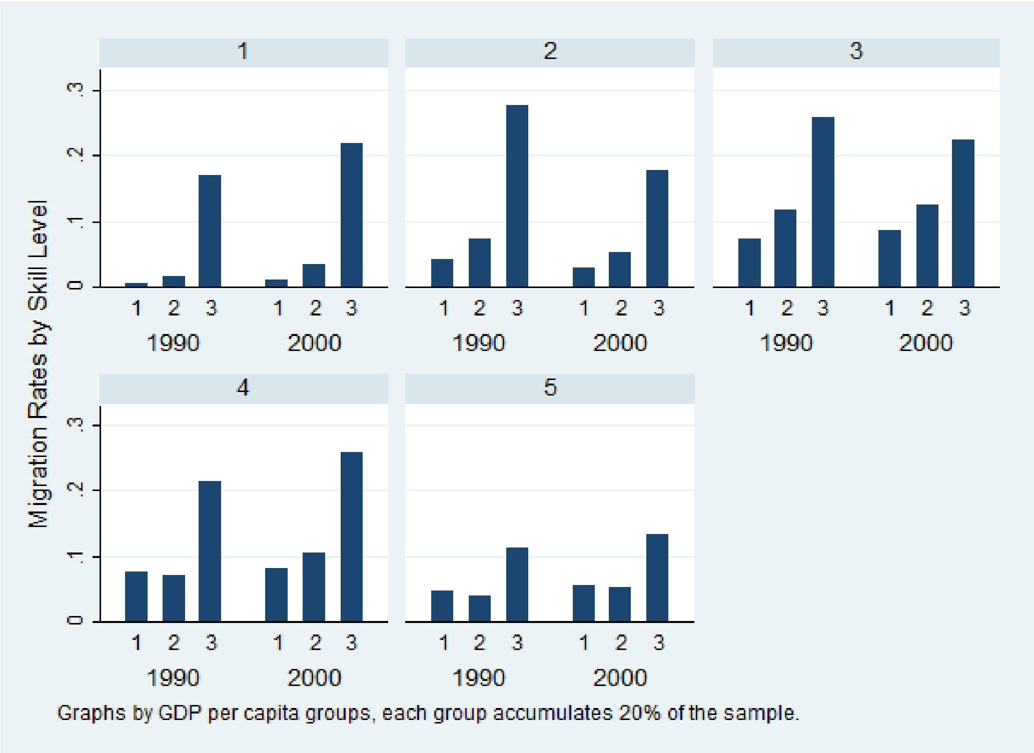
to play an important role in explaining the graph. There are other studies such as Fernández-Huertas (2012) that also emphasize the effects of the wealth constraints on determining migrant selection.[17] In our model, there are two main forces driving migration; wage differences and migration costs. However, we should also consider the wealth distribution and educational costs as two key elements of the model that determine the size of each education subgroup and the size of migration in each subgroup. Proposition 3 states that for a high technological gap: for high migration costs a J-shape appears since only wealthy and highly educated individuals will migrate; for a slightly lower migration costs an I-shape appears because the not so wealthy and highly educated also have incentives to migrate; and for very low migration costs we expect a U-shape because all individuals benefit from migration.[18]

In Figure 2.10 we can observe the average migration rates by educa-

[17]Fernández-Huertas (2012) concludes that higher skill prices in urban Mexico than in the US explain the negative migrant selection in urban Mexico. However, skill prices are not enough to generate negative selection in rural Mexico. He also concludes that the low prevalence of network effects lower the negative selection, and that wealth constraints must be added to explain the positive selection.

[18]See Figure 2.5 to compare the selection shape for different migration costs that are in line with Proposition 3.

Figure 2.10: Migrant selection by GDP groups



tional group, year, and level of GDP per capita. Each of the 5 subgraphs accumulates 20% of the countries ordered by GDP per capita, i.e., the first graph includes the 20% of countries with the lowest GDP per capita and the fifth graph includes the countries with the highest GDP per capita. Furthermore, bars 1, 2 and 3 refer to low-, medium-, and high-skilled individuals respectively for the years 1990 and 2000. Note that for low GDP per capita levels migration is much higher for the highly skilled than for the low skilled. Sub-graphs 1 to 3 show that as GDP per capita increases there is a higher increase in migration rates of low and medium skilled that we associate to smaller binding constraints. Sub-graphs 3 to 4 show that there is a small reduction in migration for the low and medium skilled, which is more pronounced for the medium skilled. This graph captures the fact that high skilled individuals profit more from migration than the other education groups due to low technological gaps and high migration costs. The final subgraph shows that very rich economies have lower migration rates as suggested in the previous subsection. The model seems to fail to capture selection in the most advanced economies, which tend to show a U-shaped relation although the model predicts that only highly educated individuals profit from migration due to low technological differences. Notwithstanding, this failure may reflect how in advanced economies, wealth constraints do not matter as much as in non-developed countries for migrant selection or how migration costs are much lower among developed countries.

2.3.4 Effects of migration on inequality

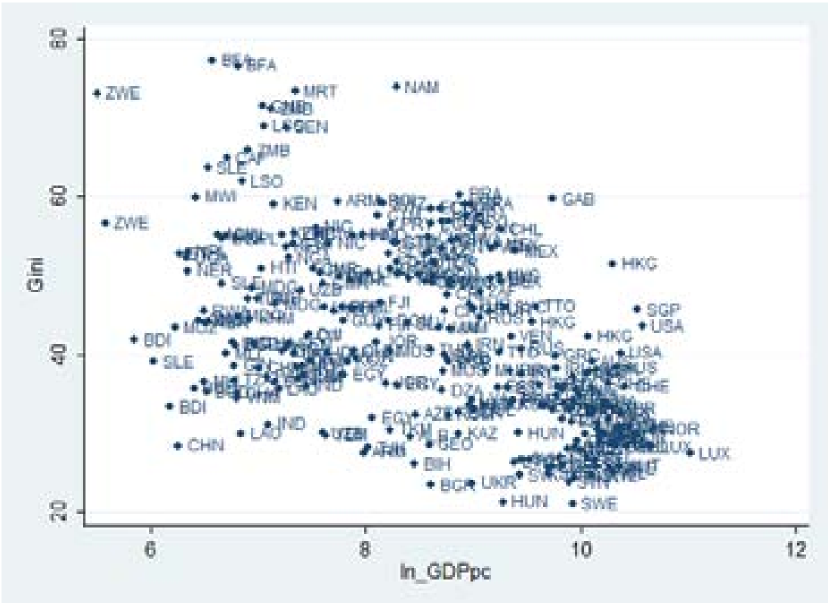
The study of the evolution of inequality started with the seminal work by Kuznets (1955). He pointed out that inequality increases in the first stages of development and decreases in the more advanced stages. Hence, inequality has an inverted-U shape in relation to economic development. Different authors have supported and criticized Kuznets' hypothesis.[19] However, Barro (2000) concludes that it seems to appear as an empirical regularity but the level of economic development does not explain the variations in inequality across countries or over time. Other authors have observed that inequality is increasing among developed countries because of the increasing demand for skills, but there is still a higher concentration of developed countries around low levels of inequality than the levels of inequality of not so developed countries. Figure 2.11 plots the relation between Gini index and log of GDP per capita (level of development) for different countries and years.[20] As pointed out by Barro (2000), the relation is not immediately clear but becomes clearer when controlling for other variables.

Although it is not directly considered in this work, the network effects of migration on inequality also seem to have an inverted-U shape. McKenzie and Rapoport (2007) describe this for two regions in rural Mexico. They find that migration increases inequality in the region with lower history of migration but decreases inequality in the region with a longer history of migration. However, they cannot observe the evolution of inequality and development in the same regions over time, they only observe each region twice in time. Previous authors have suggested this relation before but it could not be tested empirically due to a lack of data. As we have observed in the previous sections, selection is highly persistent. Migrant networks tend to increase the negative selection of migrants because previous migration may reduce the cost. Although the model is not directly reproducing migrant networks, we discussed a reduction in migration costs which generates the type of relation between migration and inequality. In line with our study, in order to observe

[19] Even the causal relationship between economic growth and inequality is criticized by Quah (2003).

[20] We use information about the Gini index from 1975 to 2006 for all the 143 different countries available in this period. In total there are 304 observations.

Figure 2.11: Inequality and development



such types of movement in the data we need countries with a high enough level of development and sufficiently low migration costs, which is what we observe in the case of Mexico. Recently, Ha et al. (2009) obtained an equivalent result for internal migration in China using village panel data. Both studies control for the effects of per capita income and find that it is not significant but increases inequality (positive coefficient in the regression).

As pointed out by Ha et al. (2009) it is not possible to carry out an empirical study that distinguishes the effects of migration not due to development in a dynamic setting. Recall that McKenzie and Rapoport (2007) obtained evidence on the effects of migration on inequality. However, they can not take into account the effects of development in a dynamic context because they only have data for two points in time for two different regions. This variable cannot be instrumented because of the direct relation between migration, income and inequality. From these regressions, we could not conclude that there is a causal relation between migration and inequality. There may be endogeneity problems. In particular, there may be reverse causality between migration costs and Gini, and collinearity between GDP levels and migration costs. Which is the hypothesis stressed along this paper.

In order to study the effects of migration on inequality, on the one hand we have constructed a theoretical model that satisfies the main stylized facts that determine migration, on the other hand we have introduced the migration decision to a framework that reproduces the evolution of inequality in relation to economic growth in line with some authors in the field.

2.4 Conclusion

This paper establishes a theoretical relation between economic inequality and migration in labor-exporting countries. It can thus provide an explanation for the evolution of inequality in countries with different productivity levels. It is shown that the effect of migration on inequality, through remittances and bequests left to children, depends on the productivity differences and the migration costs between source and destination countries of migrants.

On the one hand, the model includes the main drivers of migration such as education decisions. We are able to characterize the share of population that migrates (size) and the education characteristics of migrants (selection) consistently with previous studies and data. On the other hand, the dynamic part permits study of the long-term effects of migration on inequality in source countries. Following Kuznets' classic view of the relation between development and inequality, we have augmented this view by using migration as a possible source of variation in this relation. Countries with an equal level of development can show different levels of inequality due to different migration opportunities.

The model predicts that in countries with a low level of technology or countries with a high technological gap with respect to developed countries, the benefits of migration are higher but economic inequality increases unless the migration costs are very low. All households in the economy can benefit from migration but the wealthiest households can benefit more, and this results in an increasing difference in wealth across households. In contrast, more developed countries experience a reduction in inequality, but only highly-educated households can benefit from migration unless the migration costs are very low.

Some studies reveal the benefits of migration and remittances on growth. Since migrants increase wealth in our model, we take this wealth effect to study the impact on inequality. We have shown that migration and remittances reduce the productivity level necessary to reduce inequality compared to an economy without migration. In line with Kuznets' hypothesis, we depart from the baseline model that predicts that developing countries increase inequality and developed countries reduce inequality to show that migration can be a factor that accelerates this process. Migration helps and encourages households to accumulate wealth and invest in education in order to acquire human capital in migrants' homes.

Our analysis disregards several important features. For example, it does not include either brain drain or brain gain effects in a stochastic sense even though migration increases human capital in source countries in the model. Brain drain and brain gain can affect the technological process and the accumulation of human capital by decreasing or increasing in amount. Moreover, our analysis only considers the network effects of migrants in a vague sense. In line with the empirical literature, we could consider migration costs as a decreasing function of aggregate past migration. The introduction of decreasing costs would not change the main conclusions of the paper but convergence to an egalitarian society would be faster. As we have observed, lower migration costs increase the amount of people who can benefit from migration. Furthermore, a reduction in migration costs yields a reduction in the level of development necessary to reduce inequality.

References

Barro, R.J. 2000. "Inequality and Growth in a Panel of Countries." *Journal of Economic Growth*, 5, pp. 5-32.

Beine, M., Docquier, F., and C. Ozden. 2011. "Diasporas." *Journal of Development Economics*, 95(1), pp. 30-41.

Bertoli, S., and J. Fernández-Huertas. 2012. "Visa Policies, Networks and the Cliff at the Border." Mimeo, FEDEA.

Bollard, A., McKenzie, D., Morten, M., and H. Rapoport. 2011. "Remittances and the Brain Drain Revisited: The microdata show that more educated migrants remit more". *World Bank Economic Review*, 25(1), pp. 132-156.

Borjas, G.J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review*, 77(4), pp. 531-553.

Borjas, G.J. 2003. "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics*, 118(4), pp. 1335-74.

Caponi, V. 2010. "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US?" *Annales d'Economie et de Statistique*, 97-98, pp. 207-234.

Ceroni, B. 2001. "Poverty Traps and Human Capital Accumulation." *Economica*, 68, pp. 203-19.

Docquier, F., Rapoport, H., and I. Shen. 2010. "Remittances and inequality: A dynamic migration model." *Journal of Economic Inequality*, 8(2), pp. 197-220.

Docquier, F., Lowell, B.L., and A. Marfouk. 2009. "A gendered assessment of highly skilled emigration." *Population and Development Review*, 35(2), pp. 297-321.

Fernández-Huertas, J. 2012. "Understanding Different Migrant Selection Patterns in Rural and Urban Mexico." Mimeo, FEDEA.

Friedberg, R., and J. Hunt. 1995. "The Impact of Immigrants on Host Country Wages, Employment and Growth." *Journal of Economic Perspectives*, 9(2), pp. 23-44.

Galor, O., and D. Tsiddon. 1997. "The Distribution of Human Capital and Economic Growth." *Journal of Economic Growth*, 2(1), pp. 93-124.

Galor, O., and J. Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60, pp. 35-52.

Grogger, J., and G. Hanson. 2011. "Income Maximization and the Selection and Sorting of International Migrants." *Journal of Development Economics*, 95(1), pp. 42-57.

Ha W., Yi J., and J. Zhang. 2009. "Inequality and Internal Migration in China: Evidencefrom Village Panel Data." Human Development Research Paper, United Nations Development Programme.

Hanson, G. 2010. "International Migration and Development." *Equity and Growth in a Globalizing World* (eds. Ravi Kanbur and Michael Spence), ch. 10, pp. 229-262. Commission on Growth and Development, The World Bank.

Kuznets, S. 1955. "Economic Growth and Income Inequality." *American Economic Review*, 45(1), pp. 1-28.

Lucas, R., and O. Stark, O. 1985. "Motivations to remit: evidence from Botswana." *Journal of Political Economy*, 93, pp. 901-918.

Mayda, A. 2010. "International migration: a panel data analysis of the determinants of bilateral flows." *Journal of Population Economics*, 23(4), pp. 1249-1274

Mckenzie, D., and H. Rapoport. 2007. "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico." *Journal of Development Economics*, 84, pp. 1-24.

Ottaviano, G., and G. Peri. 2012. "Rethinking the Effects of Immigration on Wages". *Journal of the European Economic Association*, 10(1), pp. 152-197.

Quah, D. 2003. "One third of the world's growth and inequality." *Inequality and Growth: Theory and Policy Implications* (eds. T. Eicher and S. J. Turnovsky), ch. 2, pp. 27-58. MIT Press, Cambridge.

Stark, O., Taylor, J.E., and S. Yitzaki. 1986. "Remittances and Inequality." *The Economic Journal*, 96(383), pp. 722-740.

Vidal, J. 1998. "The effect of emigration on human capital formation." *Journal of Population Economics*, 11(4), pp. 1432-1475.

Chapter 3

Informality and Long-run Growth

Written jointly with Frédéric Docquier and Tobias Müller.

3.1 Introduction

The informal economy is defined as the part of an economy that is not taxed, monitored by any form of government, or included in any gross national product. It is obviously difficult to measure informality precisely, but informality is undoubtedly a much more severe phenomenon in developing countries.[1] For example, Buehn et al. (2010) estimated the average size of the shadow economy as a percentage of "official" GDP and obtained an average size of 38.4 percent in Sub-Saharan Africa, 36.5 percent in Europe and Central Asia, and 13.5 percent in high-income OECD countries. Schneider's data demonstrated similar results and showed that informality represents more than two-thirds of official GDP in the most affected countries (Schneider, 2005).

In addition, the nature of the informal economy also differs between rich and poor countries. In developed countries, the informal sector is characterized by unreported employment and sales. Informal activities are governed by the same production technology as in the formal sector and are simply hidden from the state for tax, social security or labor law purposes. Such *tax-based* informality ranges from 10 to 15 percent of official GDP in high-income countries (Schneider, 2005).

Informality is of a different nature in developing countries (although tax evasion also plays a role). It is seen as the only way to earn a living for people who are outside of the formal economy and not on anyone's payroll. Most of them live and work in this sector not because it is their wish or choice, but because they have no opportunities to be hired by an employer from the formal sector for an acceptable wage except for a few hours or days, with no right to be hired again. Such *poverty-based* informality is a way of life in poor countries. Many ubiquitous cottage microenterprises found on every street corner are not registered with authorities, and their production

[1] Measuring informality is a difficult task. People and firms who are engaged in illegal activities do not want to be known, or do not report their illegal activities. Measurement techniques can be grouped in direct and indirect methods; none of them are exempt from criticism. Indirect methods are more macroeconomic in essence: they look at the discrepancy between aggregate income and expenditure, electricity consumption versus economic activity, or monetary indicators (illegal activities conduct more transactions in cash). We also find authors who combine several indirect methods, as Schneider and coauthors. They use structural-equation estimation (MIMIC) that distinguishes between causes and indicators. The main causes are tax and social security contribution burdens, intensity of regulations, quality of public sector services, and state of the official economy. Among the indicators we can find monetary indicators, labor market indicators (comparison between total labor force and formal employment), or the state of the economy. An example of a direct method is the use of micro surveys. But again, all of them are subject to different criticisms.

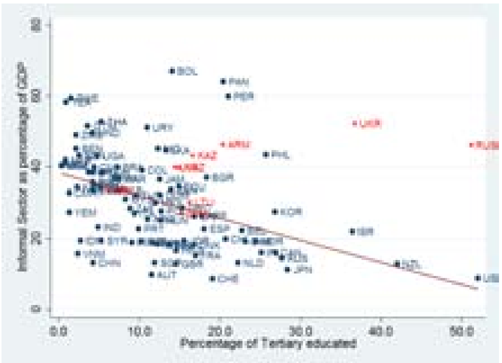
is governed by a specific technology, less efficient and intensive in low-skilled labor.

This paper investigates the dynamic implications of informality, a relatively neglected aspect in the existing literature, which mainly focuses on tax-evasion motive and possible coordination failures in entrepreneurs' decisions. As far as *tax-based* informality is concerned, a large amount of literature has formalized firms' and workers' decisions to join the informal sector to avoid taxation or regulation from the government. Among others, Zenou (2008) exploits a search-matching model a la Mortensen-Pissarides to explain its emergence. Increasing empirical literature aims at assessing the effect of taxes on informality in middle-income countries. Using a survey of firms in Brazil, De Paula and Scheinkman (2010, 2011) emphasized the role of value added taxes in transmitting informality through chain effects: informality of a firm is correlated to the informality of firms from which it buys or sells. Inspired by the seminal work of Rosenstein-Rodan (1943), another strand of literature (see Murphy et al. 1989, or Krugman 1991) has demonstrated that predominance of *poverty-based* informality can be seen as a result of a coordination failure, impeding the process of industrialization and productivity growth. They develop models of multiple equilibria, in which firms can choose to operate in the informal sector (characterized by low productivity and wages) or in the formal sector (characterized by high productivity and wages, and fixed equipment costs). Each firm has an incentive to move from informality to formality if the demand for the goods produced is large enough. This takes place when the economy-wide average income is high, i.e., when other firms industrialize and pay higher wages. Hence, a firm's decision whether to industrialize or not depends on its expectation of what other firms will do.

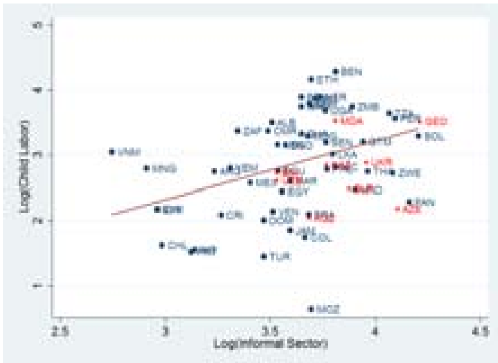
Our approach analyzes informality from a different angle. We want to explore the relationships between informality, wage inequality, human capital accumulation, child labor and long-run growth in a unified model. We build a two-sector model, in which people choose to join or not to join the informal sector. We disregard taxation and simply assume the existence of technological differences between sectors (as in Murphy et al. 1989, or Krugman 1991). Then we investigate the implications of *poverty-based* informality on welfare, inequality, growth, and effectiveness of development policies. Our philosophy is to use an abstract economic model, which highlights the major economic mechanisms underlying the formation and persistence of the informal sector and development. Incentives to invest in children's education and opportunities to obtain income from children will play a key role. We then relate the theory to the data, calibrate the parameters of our model and study its dynamic properties. Such a quantitative theory approach is now the dominant research paradigm used by economists incorporating rational expectations and dynamic choice into short-run macroeconomic and monetary economics models (King, 1995). However, very little has been done so far with this methodology in long-term macroeconomics and development economics.

We require our model to be compatible with five major stylized facts (SF1 to SF5) on *poverty-based* informality and development, as illustrated on Figures 3.1(a) to 3.1(d).[2]

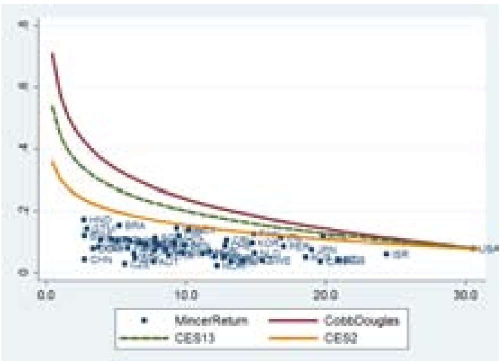
Figure 3.1: Stylized facts on informality, education and development



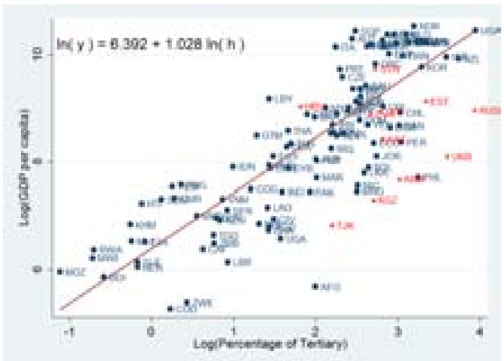
(a) Tertiary educated and informal sector size in 2000.



(b) Child labor and informal sector size in 2000.



(c) Tertiary educated and return to one year of schooling.



(d) GDP per capita and tertiary educated in 2000

Data sources. Education: Barro and Lee (2010); Informality: Schneider (2005); GDP: PWT 7.0; Child labor: World Development Indicators (2012); Returns to schooling: Hendricks (2004).

[2]In regression lines of Figures 3.1(a), 3.1(b), and 3.1(d), we exclude observations for socialist countries (marked with x) because informality in these countries seems to have different nature. Trends are steeper without these observations.

SF1. Informality decreases with development. Figure 3.1(a) shows the relation between the proportion of tertiary educated (individuals with some college education) and the ratio of output between the informal and formal sectors in year 2000. It shows a downward sloping relationship between informality and the economy-wide proportion of high-skilled workers. Our model will endogenize the size of the informal sector and be consistent with this fact. The rationale is the following: low-skilled workers are mobile across sectors. When the proportion of high-skilled workers is low in the formal economy, the demand for low-skilled labor is also low and formal firms pay low wages to the less educated. Many workers then move to the informal sector where wages are more attractive. Informality thus serves to protect low-skilled workers against very low levels of income offered in the formal sector and extreme poverty.

SF2. The informal sector exhibits lower productivity and uses low-skilled workers. This is a consensual hypothesis in informality models (Rosenstein-Rodan 1943, Murphy et al. 1989, Krugman 1991). It is supported by empirical studies. De Paula and Scheinkman (2011) showed that informal firms are managed by less able entrepreneurs, are smaller and exhibit low capital-labor ratios. They estimated that informal firms face at least 1.3 times the cost of capital of formal firms. Similarly, La Porta and Shleifer (2008) found evidence of a substantial difference between the registered and the unregistered firms regarding the skills of their managers, and suggested that this may drive many other differences, including the quality of inputs and access to finance. Rodrik (2011) points out that there is rapid unconditional convergence between rich and poor countries in manufacturing industries, but this phenomenon is hidden by a persistent specialization of poor countries in low-productivity (formal and informal) activities. Based on these facts, our model defines informality as a sector with lower productivity, low-skilled employment, and constant marginal productivity of labor. On the contrary, the formal sector combines high-skilled and less educated workers, exhibits decreasing marginal productivity, constant returns to scale, and higher total factor productivity.

SF3. Child labor increases with informality. One of the underlying aspects of informality is the existence of child labor. We can think of different forms of child labor, from shoeshine boys to children working in mining extraction. In general, children are not reported as part of the official labor force. Even if formal firms employ children, they are not recorded as part of their formal workers by the state agencies. Figure 3.1(b), plots the percentage of male children who work, and the size of informality as percentage of GDP in 2000. The World Bank data considers child labor to be children involved in economic activity for at least one hour in the reference week of the survey.

We can observe a positive correlation between informality and child labor, if informality increases, child labor increases as well. The reasons for child labor existing are related to wealth constrained households whose income is so low that they cannot send children to school. These families need their children to work in order to increase their income instead of investing in education, which can increase their future wealth and might help them to escape from poverty.

SF4. Skill premia are limited in poor countries, and no standard labor market model can account for such low skill premia. The relationship between the rate of return to one year of college (Hendricks 2004) and the proportion of college graduates in the labor force (Barro and Lee 2010) is represented on Figure 3.1(c).[3] Although return to education decrease with human capital, highest rates do not exceed 20 percent per year of schooling in low-income countries. Standard labor market models predict much larger return rates in developing countries. The CES representation is common in labor markets studies (such as Katz and Murphy, 1992, Card and Lemieux, 2001) and in cross-country analysis of relative productivity (Caselli and Coleman, 2006). The range between 1.3 and 2 for the elasticity of substitution can be seen across most labor market studies including Angrist (1995), Borjas and Katz (2007) and Katz and Murphy (1992). Assuming that college graduates have ten years more of education than the less educated and wages are equal to the marginal productivity of labor, the thin lines on Figure 3.1(c) represent the prediction of CES models with elasticities of substitution equal to one (Cobb-Douglas), 1.3 or 2.0. None of these models match the data. The average share of college graduates is around 3 percent in low-income countries. For such countries, the models predict a return to schooling comprised between 26 percent and 50 percent. The data provided in Hendricks (2004) show a maximal return to schooling of around 15 percent. We conclude that either the elasticities of substitution estimated for developed countries do not fit the production function of developing countries (an elasticity of 4.25 would be needed to match observations!), or the structure of the labor market differs across countries. We plead for the second hypothesis and see informality as a key factor limiting the skill premium and wage inequality in poor countries. Informality maintains a large skill ratio (i.e., ratio of college graduates to less educated workers) in the formal sector, thus keeping the return to schooling at a low level (Rodrik, 2011).

SF5. The elasticity of recorded GDP per capita to human capital is close to unity and school enrolment are lower in poor countries. Although many

[3] We use the most recent year of information of Mincerian returns in each country from Hendricks (2004).

studies point out that education has not generated as much growth as expected in developing countries, it is also reported that education is one of the necessary components to grow. As shown on Figure 3.1(d), the correlation between the proportion of college graduates in the labor force and GDP per capita is large, and the elasticity is close to unity. Despite scarcity in human capital and larger returns to schooling, contemporaneous school enrolment rates are lower in poor countries.

In this paper, we build a model compatible with these five stylized facts. The model may generate multiple equilibria or uniqueness, depending on the parameters' values. In the absence of informality, the model predicts long-run convergence in income across nations. Informality may slow down this convergence process or be the source of a poverty trap. The reason is that informality keeps skill premia at a relatively low level, reduces incentive to invest in education, and is conducive to child labor. Using the stylized facts above and other consensual parameters from the literature, we calibrate our model and study its quantitative properties. This allows us to discriminate between the poverty-trap and slow-convergence hypotheses. The calibration exercise reveals that the case for the poverty-trap hypothesis is strong: although informality serves to protect low-skilled workers in the short-run, it prevents income convergence across countries. On this basis, we assess the effectiveness of different policy options. Sudden elimination of informality would induce large welfare losses for several generations of poor people on the transition path. We thus compare different Pigouvian policies (subsidizing education to all families, or to low-income families, subsidizing high-skilled formal employment, or low-skilled formal employment) assuming that subsidies are financed by development assistance. Two criteria are used to evaluate these policies: cost-effectiveness and the length of the transition required to exit the poverty trap. Among the four subsidies considered, education subsidies to low-income families dominates the others in terms of cost efficiency. Moreover, only wage subsidies for low-skill jobs in the formal sector play a distinct and complementary role in the transition to the high-income equilibrium. Whereas the education and the high-skilled formal employment subsidies speed up the accumulation of human capital, the low-skill wage subsidy reduces the threshold at which the informal sector disappears. Therefore, targeted education subsidies are the cheapest single policy, but for medium time horizons, a combination of the two policies is found to be the most cost-efficient choice.

The remainder of this paper is organized as follows. Section 3.2 describes the model. The implications of informality are examined in Section 3.3. In Section 3.4, we calibrate the model and study its quantitative properties. Section 3.5 concludes.

3.2 Model

We develop a two-period overlapping generations model in infinite discrete time with children and working-age adults. In every period, a single homogeneous good can be produced in two different sectors, the formal and informal sectors (labeled f and i). Formal firms employ high- and low-skilled workers whereas the informal sector only employs low-skilled workers. In each period there is an endogenous number of adults of each type who choose how much to consume and how much to invest in the education of their children. All decisions are made in the adult period of life, i.e., children do not get to decide anything. Below, we describe the technology, preferences, the dynamics, and define the competitive equilibrium path of our economy.

3.2.1 Production

A single good can be produced in two sectors. The formal sector employs high- and low-skilled labor and the informal sector only uses low-skilled labor. Let h_t be the proportion of high-skilled adults at time t, and N_t the total labor force of adults. We denote by $\overline{H}_t = h_t N_t$ and $\overline{L}_t = (1 - h_t)N_t$ the size of high- and low-skilled labor forces, respectively. Low-skilled workers are assumed to be perfectly mobile across sectors, whereas high-skilled workers have no incentive to join the informal sector.[4] Output Y_t is the sum of output $Y_{f,t}$ produced in the formal sector and output $Y_{i,t}$ produced in the informal one. Output produced in each sector is

$$Y_{f,t} = A_t H_t^{\alpha} L_{f,t}^{1-\alpha}, \tag{3.1}$$

$$Y_{i,t} = B L_{i,t}, \tag{3.2}$$

where α is the share of output produced in the formal sector by high-skilled workers, A_t is a time-varying scale factor representing the state of technology, H_t is the quantity of high-skilled workers employed in the formal sector, $L_{f,t}$ and $L_{i,t}$ are the quantities of low-skilled workers employed in formal and informal sectors, respectively, and B is a scale factor associated to the technology in the informal sector, which is assumed to be constant. For simplicity purposes, we write $B = \widetilde{\gamma} A_0$, where $\widetilde{\gamma}$ is a parameter that allows us to write B in terms of the scale factor A_0.[5]

Firms choose inputs by maximizing profits

$$Y_{f,t} - w_{h,t} H_t - w_{l,t} L_{f,t} \tag{3.3}$$

[4] Our model does not account for brain waste, which may be responsible for employment of educated workers in informality.

[5] We require $\widetilde{\gamma} \in [0, 1 - \alpha]$ to be consistent with *SF2*.

and

$$Y_{i,t} - w_{l,t} L_{i,t}, \tag{3.4}$$

subject to $Y_{i,t} \geq 0$.[6] Under perfect competition, firms in formal and informal sectors choose employment levels by equalizing the marginal productivity of high- and low-skilled workers with their wage rates $w_{h,t}$ and $w_{l,t}$. In the formal sector, these conditions are

$$w_{h,t} = A_t \alpha \left(\frac{L_{f,t}}{H_t} \right)^{1-\alpha}, \tag{3.5}$$

$$w_{l,t} = A_t (1-\alpha) \left(\frac{L_{f,t}}{H_t} \right)^{-\alpha}. \tag{3.6}$$

The output and employment decisions in the informal sector can be described by the complementary slackness conditions

$$\frac{w_{l,t}}{\widetilde{\gamma} A_0} \geq 1, \quad Y_{i,t} \geq 0, \quad \text{and} \quad \left(\frac{w_{l,t}}{\widetilde{\gamma} A_0} - 1 \right) Y_{i,t} = 0, \tag{3.7}$$

which depict two alternatives that will give rise to two model regimes:

1. output in the informal sector is positive and marginal cost, $w_{l,t}/(\widetilde{\gamma} A_0)$, is equal to the (unitary) price of output (or, equivalently, marginal productivity of labor in the informal sector is equal to the low-skilled wage);

2. Firms in the informal sector produce no output and marginal cost exceeds the price of output (marginal productivity of labor in the informal sector is smaller than the low-skilled wage).

Moreover, we assume that total factor productivity (TFP) A_t in the formal sector is endogenous. It is a concave function of the skill ratio in the formal sector.[7] For simplicity and in reference to the AK model, the elasticity of TFP to the skill ratio equals $1 - \alpha$, i.e.,

$$A_t = A_0 \left(\frac{H_t}{L_{f,t}} \right)^{1-\alpha}. \tag{3.8}$$

[6] For simplicity, we omit the constraint $Y_{f,t} \geq 0$ because it is never binding in equilibrium.

[7] This assumption implies that the proportion of high-skilled individuals generates a positive externality on the aggregate productivity. It is a particular case of Lucas' model (Lucas 1988) and is also related to other AK models as the ones presented by Romer (1986) and Rebelo (1991).

3.2.2 Preferences

Each adult of type $k \in \{h, l\}$ at period t chooses consumption $c_{k,t}$ and the proportion $q_{k,t} \in [0, 1]$ of children sent to college to maximize utility. The utility function is logarithmic and depends on consumption $c_{k,t}$ and the average future wage $\overline{w}_{k,t+1}$ of children,

$$U_{k,t} = \ln(c_{k,t}) + \beta \ln(\overline{w}_{k,t+1}) \tag{3.9}$$

where β is the rate of preference for the income of children, and the average future wage of children is

$$\overline{w}_{k,t+1} = (1 - q_{k,t})w_{l,t+1} + q_{k,t}w_{h,t+1} = w_{l,t+1}(1 + q_{k,t}\sigma_{t+1}), \tag{3.10}$$

which depends on the value of the skill premium $\sigma_{t+1} = (w_{h,t+1} - w_{l,t+1})/w_{l,t+1}$ in the next period.

Educating a child incurs a monetary cost $\tilde{e}$.[8] Non-educated children can work in the informal sector as long as the informal sector exists, whereas educated children go to school and have no time left to work. In the informal sector, children receive a fraction $\eta \in [0, 1]$ of the low-skilled wage rate because they lack experience and physical strength compared to adults. The budget constraint is

$$c_{k,t} = w_{k,t} - n_k q_{k,t}\tilde{e} + n_k(1 - q_{k,t})\eta w_{l,t} d_t, \tag{3.11}$$

where n_k is the number of children of a k-type adult, and d_t is a dummy variable equal to 1 if some output is produced in the informal sector, and 0 otherwise.

Plugging (3.10) and (3.11) into (3.9) and maximizing utility with respect to $q_{k,t}$, we obtain

$$\hat{q}_{k,t} = \frac{\beta\sigma_{t+1}(w_{k,t} + n_k\eta w_{l,t}d_t) - n_k(\tilde{e} + \eta w_{l,t}d_t)}{(1 + \beta)n_k(\tilde{e} + \eta w_{l,t}d_t)\sigma_{t+1}}. \tag{3.12}$$

Therefore, the optimal level of education is

$$q^*_{k,t} = \begin{cases} 0 & \text{if } \hat{q}_{k,t} < 0 \\ \hat{q}_{k,t} & \text{if } 0 \leq \hat{q}_{k,t} \leq 1 \\ 1 & \text{if } \hat{q}_{k,t} > 1. \end{cases} \tag{3.13}$$

[8] As we will observe later, equilibrium high-skilled wages will be constant. Hence, a constant education cost is equivalent to education costs being proportional to high-skilled wages, which implies that education is more difficult to obtain for low-skilled than for high-skilled workers.

3.2.3 Dynamics and competitive equilibrium

In the previous section we obtained adults' optimal decision on the proportion of children to be educated. Hence, given the proportion h_t of high-skilled workers in period t, fertility rates n_h and n_l, and the equilibrium condition (3.13), we can compute the proportion h_{t+1} of high-skilled workers in the next period. For simplicity, we assume that high-skilled parents educate all their children, i.e., we assume that parameters are such that $\hat{q}_{h,t} \geq 1$, which implies that $q^*_{h,t} = 1$.[9] On the contrary, low-skilled parents only educate an endogenous fraction $q_{l,t} \in [0, 1)$ of their children. Therefore, the dynamics of the skill ratio across generations is governed by

$$\frac{h_{t+1}}{1-h_{t+1}} = \frac{n_h h_t + n_l q_{l,t}(1-h_t)}{n_l(1-q_{l,t})(1-h_t)} = \frac{n}{1-q_{l,t}}\frac{h_t}{1-h_t} + \frac{q_{l,t}}{1-q_{l,t}}, \tag{3.14}$$

where $n \equiv n_h/n_l$ is the fertility ratio of high- to low-skilled workers.

In addition, the labor-market-clearing conditions are

$$H_t = \overline{H}_t, \tag{3.15}$$

the supply and demand of high-skilled workers should be equal in equilibrium. In the next sections, we use H to denote the equilibrium amount of high-skilled workers. And

$$L_{f,t} + L_{i,t} = \overline{L}_t + \eta n_l(1-q_{l,t})\overline{L}_t, \tag{3.16}$$

low-skilled workers in formal and informal sectors should be equal to low-skilled adults and the efficiency units of children who work. Moreover, we impose the following extra condition:

$$L_{i,t} > \eta n_l(1-q_{l,t})\overline{L}_t \text{ whenever } L_{i,t} > 0. \tag{3.17}$$

Some adult workers are required for the functioning of the informal sector. This is a reasonable assumption since we are imposing that children need some infrastructure provided by adults to the informal sector in order to operate. We now define an equilibrium for our economy:

Definition 1. Given an initial population size N_0 and an initial number H_0 of high-skilled workers, an intertemporal equilibrium consists of sequences of prices $\{w_{h,t}, w_{l,t}\}$, aggregate quantities $\{N_t, \overline{H}_t, \overline{L}_t, H_t, L_{f,t}, L_{i,t}\}$, and household's decisions $\{c_{j,t}, q_{j,t}\}$ for $j = h, l$ and for all t such that:

[9] An alternative assumption to ensure that $\hat{q}_h \geq 1$ is to assume that h can not be higher than $\bar{h} < \alpha$ and parameters are such that $(A_0\alpha/(\tilde{e}n_h) - 1)\,\beta \geq 1 + (1-\alpha)\bar{h}/(\alpha - \bar{h})$. de la Croix and Docquier (2012) use the same simplifying assumption.

1. the household's decisions $c_{j,t}$ and $q_{j,t}$ maximize utility (3.9) subject to the constraints (3.10) and (3.11);

2. the firms' choices H_t, $L_{f,t}$, and $L_{i,t}$ maximize profits (3.3) and (3.4) subject to the constraint $Y_{i,t} \geq 0$;

3. the prices $w_{h,t}, w_{l,t}$, and aggregate quantities $\overline{H}_t$, $\overline{L}_t$ are such that markets clear, i.e., (3.15) and (3.16) hold;

4. aggregate variables N_t, H_t evolve according to (3.14);

5. $\overline{L}_t$, $L_{i,t}$, and q_t^l satisfy (3.17).

3.3 Implications of informality

In this section we characterize the existence of two possible transitory regimes and then study the implications of informality for human capital accumulation and long-run growth.

3.3.1 The *formality* and *informality* regimes

Two regimes arise as a consequence of informality. On the one hand, the *informality* regime arises if the formal and informal sector co-exist. On the other hand, the *formality* regime arises if all low-skilled adults opt for the formal sector and the informal sector disappears.

The *formality* regime is characterized by the absence of an informal sector. Then, plugging (3.8) into (3.5) - (3.7), wages and the skill premium in the *formality* regime are

$$w_{h,t} = A_0\alpha, \tag{3.18}$$

$$w_{l,t} = A_0(1-\alpha)\frac{h_t}{1-h_t}, \tag{3.19}$$

$$\sigma_t = \frac{\alpha(1-h_t)}{(1-\alpha)h_t} - 1. \tag{3.20}$$

Hence, in the *formality* regime, the skill premium σ_t decreases with the proportion of high-skilled workers in the economy, and the limit of the skill premium equals infinity when h_t tends to zero. A model with a single formal sector predicts huge wage disparities when human capital is low.

However, production in the informal sector becomes profitable if the marginal productivity of labor is not lower than the low-skilled wage. Combined with the assumption of perfect mobility of low-skilled workers across sectors, this implies that the number of low-skilled workers in the formal sector is proportional to the number of high-skilled workers in the economy, i.e., $L_{f,t} = \gamma H_t$ where $\gamma \equiv (1-\alpha)/\widetilde{\gamma}$ is a simple combination of parameters. Again, plugging (3.8) into (3.5) - (3.7) and taking into account that $Y_{i,t} > 0$, wages and the skill premium in the *informality* regime are

$$w_{h,t} = A_0 \alpha, \tag{3.21}$$

$$w_{l,t} = \frac{A_0(1-\alpha)}{\gamma}, \tag{3.22}$$

$$\sigma_t = \frac{\alpha\gamma}{1-\alpha} - 1 = \overline{\sigma}. \tag{3.23}$$

Therefore, the skill premium σ_t is constant when the informal sector is at work, so it does not depend on the proportion h_t of high-skilled workers. Informality explains why skill premia are limited in developing countries where the proportion of college graduates is low, as illustrated by stylized fact *SF4*.

The next lemma characterizes the emergence of the *informality* regime in terms of the proportion of high-skilled workers in the economy:

Lemma 1. *The informality regime (resp. formality regime) arises when the proportion of high-skilled workers is not too large (resp. large enough), i.e., when $h_t < 1/(1+\gamma)$ (resp. $h_t \geq 1/(1+\gamma)$).*

Proof. Low-skilled adults work in the informal sector if (3.19)<(3.22). Therefore, the *informality* regime arises if $h_t < 1/(1+\gamma)$. □

Informality was modeled in the production section as an alternative for low-skilled adults to supply their working-hours. Moreover, we observe that informality arises in economies with low levels of human capital. Let us denote GDP per capita and recorded GDP per capita by $y_t = Y_t/N_t$ and $y_{f,t} = Y_{f,t}/N_t$. Consistently with stylized fact *SF5*, our model predicts that the elasticity of formal output to human capital is equal to unity, as stated in the following proposition:

Proposition 5. *In the formality regime, GDP per capita is proportional to the share of high-skilled workers in the labor force, i.e., $y_t = A_0 h_t$, and recorded GDP is equal to GDP per capita, i.e., $y_{f,t} = y_t$. Meanwhile, in the informality regime, GDP per capita exceeds recorded GDP per capita, $y_t > y_{f,t}$, and recorded GDP per capita is proportional to the share of high-skilled workers, $y_{f,t} = A_0 h_t$.*

Proof. It follows from equations (3.1) and (3.8). □

In the *informality* regime, wages are constant. Hence, $q_{l,t}$ is equal to

$$q_{l,t}^{*} = \frac{\beta\,(1-\alpha)\,(1+\eta n_l)}{(1+\beta)\,[e\gamma+\eta(1-\alpha)]\,n_l} - \frac{1}{(1+\beta)\sigma_{t+1}}. \tag{3.24}$$

Note that in case that next period proportion h_{t+1} of high-skilled workers is not high enough so as to achieve the threshold proportion $1/(1+\gamma)$ that defines informality, then $q_{l,t}$ is constant and equal to

$$q_{l,t}^{*} = \frac{\beta\,[\alpha(1+\gamma)-1]\,(1-\alpha)\,(1+\eta n_l) - n_l\,(1-\alpha)\,[e\gamma+\eta(1-\alpha)]}{(1+\beta)\,[e\gamma+\eta(1-\alpha)]\,[\alpha(1+\gamma)-1]\,n_l} \equiv \overline{q}_l,$$

where $e = \tilde{e}/A_0$. Moreover, $q_{l,t}^{*} \leq \overline{q}_l$ because $\sigma_{t+1} \leq \overline{\sigma}$.

In line with some empirical papers as Buehn et al. (2010) or Schneider (2005), we define the informality level as the proportion of value added in the informal sector with respect to the value added in the official GDP, i.e., $I_t = Y_{i,t}/Y_{f,t}$. Note that $I_t \equiv 0$ in the *formality* regime. Consistently with stylized fact *SF1*, we have:

Proposition 6 (Short-run effects of informality)**.** *The informal sector increases low-skilled workers' wage, whereas high-skilled workers' wage is not modified. Moreover, the informality level I_t shows a decreasing relationship with respect to the proportion of high-skilled workers in the labor force in the informality regime.*

Proof. From (3.18) and (3.21) we can see that high-skilled wages are equivalent in both regimes. From (3.19) and (3.22), low-skilled wages within the *informality* regime are at least as high as in the *formality* regime if and only if $h_t/(1-h_t) < 1/\gamma$, and, by Lemma 1, the *informality* regime exists if and only if $h_t < 1/(1+\gamma)$, which is equivalent to $h_t/(1-h_t) < 1/\gamma$. Moreover, in the *informality* regime

$$I_t = \frac{Y_{i,t}}{Y_{f,t}} = \frac{1-\alpha}{\gamma}\left(\frac{(1-h_t)(1+\eta n_l(1-q_{l,t}))}{h_t} - \gamma\right).$$

Note that $q_{l,t}$ is characterized by equation (3.24). Since (3.23) and (3.20) characterize a continuous function $\sigma(h_t) = \sigma_t$ for $h_t \in [0,1]$, thus $q_{l,t}$ defined in equation (3.24) is continuous. Two cases arise, if $h_{t+1} < 1/(1+\gamma)$, then $q_{l,t} = \overline{q_l}$, and if $h_{t+1} \geq 1/(1+\gamma)$, then $q_{l,t}$ is defined by equation (3.24). In the former case $dq_{l,t}/dh_t = 0$, whereas in the latter case $dq_{l,t}/dh_t$ can be $\neq 0$. To compute this derivative let z_t be $h_t/(1-h_t)$. This monotonic variable transformation enables us to write equations (3.14) and (3.24) as

$$z_{t+1} = \frac{nz_t + q_{l,t}}{1-q_{l,t}}$$

and

$$q_{l,t}^{*} = \Omega - \frac{(1-\alpha)z_{t+1}}{(1+\beta)(\alpha-(1-\alpha)z_{t+1})},$$

where $\Omega = (\beta(1-\alpha)(1+\eta n_l)) / ((1+\beta)(e\gamma+\eta(1-\alpha))n_l)$. In order to compute the derivative $dq_{l,t}/dh_t$, we can plug the latter expression into the former expression and let H be a mapping from $\Re^2$ to $\Re$ such that

$$H(q_{l,t}, z_t) = \frac{nz_t + q_{l,t}}{1-q_{l,t}} - \frac{\alpha}{(1-\alpha)}\frac{(1+\beta)(\Omega-q_{l,t})}{1+(1+\beta)(\Omega-q_{l,t})}.$$

The vectors $(q_{l,t}, z_t)$ such that $H(q_{l,t}, z_t) = 0$ characterize the problem. Taking partial derivatives we obtain the Jacobian

$$\begin{aligned} DH(q_{l,t}, z_t) &= \left[\frac{\partial H(q_{l,t}, z_t)}{\partial q_{l,t}}, \frac{\partial H(q_{l,t}, z_t)}{\partial z_t}\right] = [DH_1, DH_2] \\ &= \left[\frac{1+nz_t}{(1-q_{l,t})^2} + \frac{\alpha}{(1-\alpha)}\frac{(1+\beta)}{(1+(1+\beta)(\Omega-q_{l,t}))^2}, \frac{n}{(1-q_{l,t})}\right]. \end{aligned}$$

Since $DH_1 > 0$, by the Implicit Function Theorem there exists a function $q_{l,t}(z_t)$ in a neighborhood of z_t and

$$\frac{dq_{l,t}}{dz_t} = -\frac{n(1-q_{l,t})}{1+nz_t+\frac{\alpha(1+\beta)(1-q_{l,t})^2}{(1-\alpha)\left(1+(1+\beta)(\Omega-q_{l,t})\right)^2}},$$

which implies that

$$\frac{dq_{l,t}}{dh_t} = -\frac{n(1-q_{l,t})}{(1-h_t)^2+nh_t(1-h_t)+\frac{\alpha(1+\beta)(1-q_{l,t})^2(1-h_t)^2}{(1-\alpha)\left(1+(1+\beta)(\Omega-q_{l,t})\right)^2}}.$$

Furthermore,

$$\frac{dI_t}{dh_t} = -\frac{1-\alpha}{\gamma}\left(\frac{1+\eta n_l(1-q_{l,t})}{h_t^2} + \eta n_l \frac{dq_{l,t}}{dh_t}\frac{1-h_t}{h_t}\right)$$

for all $h_{t+1} \neq 1/(1+\gamma)$. If $h_{t+1} < 1/(1+\gamma)$ then $dq_{l,t}/dh_t = 0$ and $dI_t/dh_t < 0$. If $h_{t+1} > 1/(1+\gamma)$ then

$$\frac{dI_t}{dh_t} = -\frac{1-\alpha}{\gamma h_t}\left[\frac{1}{h_t} + \frac{\eta n_l(1-q_{l,t})}{h_t} - \frac{\eta n_l(1-q_{l,t})}{h_t+\frac{(1-h_t)}{n}+\frac{\alpha(1+\beta)(1-q_{l,t})^2(1-h_t)n}{(1-\alpha)\left(1+(1+\beta)(\Omega-q_{l,t})\right)^2}}\right]$$

and $dI_t/dh_t < 0$, which implies that the informality level I_t shows a decreasing relationship with respect to h_t. □

The existence of the informal sector reduces inequality differences, which can be good for growth because of the negative association between high inequality and long-run growth pointed out by some authors.[10] However, informality allows firms to recruit illegal labor, which includes children of poor households. The following result makes the link between child labor and informality, consistent with stylized fact *SF3*:

Corollary 2 (Child labor)**.** *The proportion of children who work decreases as the proportion of high-skilled workers in the labor force increases in the informality regime. Hence, the proportion of children who work increases as the informality level increases.*

Proof. The proportion of children who work is

$$CL(h_t) = \frac{(1-q_{l,t})(1-h_t)n_l}{h_t n_h + (1-h_t)n_l} = \frac{(1-q_{l,t})(1-h_t)}{1-h_t(1-n)}.$$

Hence, taking the derivative with respect to h_t we obtain

$$CL'(h_t) = -\frac{dq_{l,t}}{dh_t}\frac{1-h_t}{1-h_t(1-n)} - \frac{(1-q_{l,t})n}{(1-h_t(1-n))^2}.$$

As in the previous Proposition, if $h_{t+1} < 1/(1+\gamma)$ then $dq_{l,t}/dh_t = 0$ and $CL'(h_t) < 0$. Whereas if $h_{t+1} > 1/(1+\gamma)$ then

$$CL'(h_t) = (1-q_{l,t})n\left(\frac{1}{(1-h_t(1-n))^2+\Upsilon} - \frac{1}{(1-h_t(1-n))^2}\right) < 0,$$

where

$$\Upsilon = (1-h_t(1-n))\left(\frac{\alpha(1+\beta)(1-q_{l,t})^2(1-h_t)}{(1-\alpha)\left(1+(1+\beta)(\Omega-q_{l,t})\right)^2}\right) > 0$$

and $\Omega = \left(\beta\left(1-\alpha\right)\left(1+\eta n_l\right)\right)/\left(\left(1+\beta\right)\left(e\gamma+\eta(1-\alpha)\right)n_l\right)$. Moreover, from the previous Proposition we know that I increases as h decreases, which implies that the proportion of children who work increases with informality. □

3.3.2 Effect on long-run growth

Now we study the long-run effects of informality, in particular, we study its effects on human capital accumulation. We distinguish two important

[10] See Galor and Zeira (1993) or Alesina and Rodrik (1994) among others.

channels. First, as informality limits the returns to schooling, it is likely to reduce the incentive to acquire human capital. Second, informality allows firms to recruit illegal labor, which includes children of poor households.

In the *formality* regime, i.e., $h_t \geq 1/(1+\gamma)$, substituting wage rates (3.18)-(3.20) into (3.13) yields:

$$q^*_{l,t} = \frac{\beta(1-\alpha)h_t}{(1+\beta)en_l(1-h_t)} - \frac{\alpha - h_{t+1}}{(1+\beta)(1-\alpha)h_{t+1}} \equiv q_l(h_t, h_{t+1}). \tag{3.25}$$

Moreover, human capital dynamics for an economy without informality are governed by

$$\frac{h_{t+1}}{1-h_{t+1}} = \frac{n}{1-q_l(h_t,h_{t+1})}\frac{h_t}{1-h_t} + \frac{q_l(h_t,h_{t+1})}{1-q_l(h_t,h_{t+1})} \equiv \varphi(h_t,h_{t+1}). \tag{3.26}$$

Therefore, plugging (3.25) into (3.26) characterizes human capital dynamics. To simplify these two expressions let z_t be $h_t/(1-h_t)$. This variable transformation allows us to write equations (3.25) and (3.26) as follows:

$$q^*_{l,t} = \frac{\beta(1-\alpha)}{(1+\beta)en_l}z_t - \frac{(1-\alpha)z_{t+1}}{(1+\beta)\alpha(1+z_{t+1}) - z_{t+1}} \equiv q_l(z_t, z_{t+1})$$

and

$$z_{t+1} = \frac{n}{1-q_l(z_t,z_{t+1})}z_t + \frac{q_l(z_t,z_{t+1})}{1-q_l(z_t,z_{t+1})} \equiv \varphi(z_t,z_{t+1}).$$

Moreover, the properties of the dynamical system are not modified by this transformation. The following proposition describes the long-run convergence of human capital in the *formality* regime:

Proposition 7 (Long-run convergence in the *formality* regime)**.** *The dynamical system characterized by (3.25) and (3.26) displays a stable steady state $h^{stst}_s > 0$ and an unstable steady state $h^{stst}_u = 0$ in $h \in [0,1]$ if and only if parameters satisfy that $(1+\alpha\beta)en_l < \alpha((1-\alpha)\beta + (1+\beta)nen_l)$.*

Proof. The proof is divided into three steps.

Step 1: there exists a function ψ that determines z_{t+1} given z_t and its slope is positive for all $z_t \geq 0$, i.e., $z_{t+1} = \psi(z_t)$ and $\psi'(z_t) > 0$.
Let F be a function $F : \Re^2 \to \Re$ such that $F(z_t, z_{t+1}) = \varphi(z_t, z_{t+1}) - z_{t+1}$. The vectors (z_t, z_{t+1}) such that $F(z_t, z_{t+1}) = 0$ characterize human capital dynamics. Taking partial derivatives we obtain the Jacobian

$$\begin{aligned} DF(z_t, z_{t+1}) &= \left[\frac{\partial\varphi(z_t,z_{t+1})}{\partial z_t}, \frac{\partial\varphi(z_t,z_{t+1})}{\partial z_{t+1}} - 1\right] = [DF_1, DF_2] \\ &= \frac{1}{(1-q_l)^2}\left[n(1-q_l) + q_1(1+nz_t), q_2(1+nz_t) - (1-q_l)^2\right], \end{aligned}$$

where $q_l = q_l(z_t, z_{t+1})$, $q_1 = \partial q_l(z_t, z_{t+1})/\partial z_t = \beta(1-\alpha)/(en_l(1+\beta)) > 0$, and $q_2 = \partial q_l(z_t, z_{t+1})/\partial z_{t+1} = -\alpha(1-\alpha)/\left((1+\beta)(\alpha(1+z_{t+1})-z_{t+1})^2\right) < 0$ for all z_t. Since $DF_2 < 0$, by the Implicit Function Theorem there exists a function $z_{t+1}(z_t) = \psi(z_t)$ in a neighborhood of z_t (for all z_t) and

$$z'_{t+1}(z_t) = \psi'(z_t) = -\frac{n(1-q_l) + q_1(1+nz_t)}{q_2(1+nz_t) - (1-q_l)^2}.$$

Moreover, ψ is increasing for all $z_t \geq 0$, i.e., $\psi'(z_t) > 0$, because the numerator is strictly positive if $z_t \geq 0$, while the denominator is negative.

Step 2: the dynamical system displays two steady state values in $z \geq 0$: 0 and $z_+ > 0$, and they are the only ones.
The steady state values are the vectors (z_t, z_{t+1}) such that $z_t = z_{t+1}$, or the values of z such that $F(z,z) = 0$. Note that (3.25) and (3.26) become

$$q_l(z,z) = z\frac{1-\alpha}{1+\beta}\left(\frac{\beta}{en_l} - \frac{1}{\alpha(1+z)-z}\right) \tag{3.27}$$

and

$$z = \frac{nz + q_l(z,z)}{1 - q_l(z,z)} \tag{3.28}$$

respectively. Plugging (3.27) into (3.28) and rearranging terms we obtain

$$F(z,z) = q_l(z,z)z + q_l(z,z) - (1-n)z.$$

Clearly, $z = 0$ satisfies $F(0,0) = 0$ because $q_l(0,0) = 0$. Since we are interested in the remaining solutions to the problem $F(z,z) = 0$, we substitute q_l, divide by z, and equalize to 0. The solutions to the resulting equation can be rewritten as the roots of the following grade 2 polynomial of z:

$$a_2 z^2 + a_1 z + a_0 = 0,$$

where

$$\begin{aligned}
a_0 &= -\left(1 + \alpha\beta + \frac{(1+\beta)(1-n)en_l\alpha}{1-\alpha}\right), \\
a_1 &= (1+\beta)(1-n) - (1-\beta+2\alpha\beta), \\
a_2 &= (1-\alpha)\beta.
\end{aligned}$$

Since $a_0 < 0$ and $a_2 > 0$, the roots of the polynomial are $z_- < 0$ and $z_+ > 0$. Hence, the steady state values of the dynamical system are z_-, 0, and z_+.

Step 3: $\lim_{z\to+\infty}\psi'(z)=0$.
Rewrite $\psi'(z_t)$ as

$$\psi'(z_t)=\frac{\frac{n(1-q_l)}{1+nz_t}+q_1}{-q_2+\frac{(1-q_l)^2}{1+nz_t}}$$

and note that the denominator goes to infinity when z_t goes to infinity whereas the numerator goes to 0 or to a constant because q_1 is a constant,

$$-\infty<\lim_{z_t\to+\infty}\frac{n(1-q_l(z_t,\psi(z_t)))}{1+nz_t}=-n\frac{\beta}{1+\beta}\frac{1-\alpha}{en_l}<+\infty,$$

$$0\leq\lim_{z_t\to+\infty}-q_2=<+\infty,$$

and

$$\lim_{z_t\to+\infty}\frac{(1-q_l)^2}{1+nz_t}=+\infty.$$

From *Steps 1* and *2* we know that the system is well defined and displays two different steady state values in $z\geq 0$: 0 and $z_+>0$. A necessary and sufficient condition for the instability of the 0 steady state is $\psi'(0)>1$, which is equivalent to $(1+\alpha\beta)en_l<\alpha((1-\alpha)\beta+(1+\beta)nen_l)$. Moreover, *Step 3* ensures that $z_{t+1}=\psi(z_t)<z_t$ for all $z_t>z_+$, and we can conclude that z_+ is stable because necessarily $0<\psi'(z_+)<1$. □

In the *informality* regime, i.e., $h_t<1/(1+\gamma)$ or $z_t<1/\gamma$, we have $q^*_{l,t}=\overline{q}_l$ if $h_{t+1}<1/(1+\gamma)$, which is satisfied if $1-n\geq\overline{q}_l(1+\gamma)$.[11] This condition is satisfied if the fertility ratio n is low enough, and both the relative productivity $\widetilde{\gamma}$ of the informal sector and the education cost $\tilde{e}$ are sufficiently high. In such case the dynamics of the skill ratio z_t are governed by

$$z_{t+1}=\frac{n}{1-\overline{q}_l}z_t+\frac{\overline{q}_l}{1-\overline{q}_l}\equiv\phi(z_t),\tag{3.29}$$

where $\phi(z_t)$ is a linear function of z_t with $\phi(0)>0$ and a slope smaller than one if $n<1-\overline{q}_l$.

Proposition 8 (Long-run effects of informality)**.** *There are poverty traps in the informality regime if and only if* $1-n\geq\overline{q}_l(1+\gamma)$.

Proof. Human capital dynamics are determined by (3.29). Thus, a stable poverty trap with informality emerges if and only if $\phi(1/\gamma)\leq 1/\gamma$, and $n/(1-\overline{q}_l)<1$. The former condition is equivalent to $1-n\geq\overline{q}_l(1+\gamma)$. In

[11] Follows from equation (3.14).

addition, this condition ensures that $1 > \overline{q}_l + n$. Hence, the former condition is sufficient for the latter condition to be satisfied. Therefore, there exists a steady state level of human capital such that $h^{ss} < 1/(1+\gamma)$ if and only if $1 - n \geq \overline{q}_l(1+\gamma)$. □

The previous two propositions characterize the equilibrium path of the skill ratio. Figure 3.2(a) shows the dynamics with and without informality. The solid line corresponds to an economy with informality if the skill ratio is lower than $z_0 = 1/\gamma$, while the dashed line corresponds to one without informality. For high enough levels of human capital there is not an informal sector and both lines coincide. As predicted by Proposition 7, without the informal sector the skill ratio converges to the point A_1 as long as the initial skill ratio is larger than 0. However, if the informal sector is at work, Proposition 8 states that there can be poverty traps as the one presented in Figure 3.2(a). The linear part of the solid line crosses the 45° line and the skill ratio converges to the point A_2 if the initial skill ratio is lower than z_0.

Figure 3.2(b) presents three different possibilities of skill ratio dynamics with informality. In all cases there is a jump from the *formality* to the *informality* regime due to child labor in the informal sector. Dynamic B is a possible situation without poverty traps. It might happen if, for example, the education cost $\tilde{e}$ is low enough. Dynamic A is a case with a poverty trap in the *informality* regime, and convergence to a high proportion of high-skilled workers in the *formality* regime. Whereas Dynamic C corresponds to a case where parameters satisfy that the stable steady state is 0 in the *formality* regime, or a case where there is a stable steady state greater than 0 but lower than $1/\gamma$. Because of the existence of the informal sector the poverty trap makes the economy to converge to the point C, which is characterized by a low proportion of high-skilled workers in the economy.

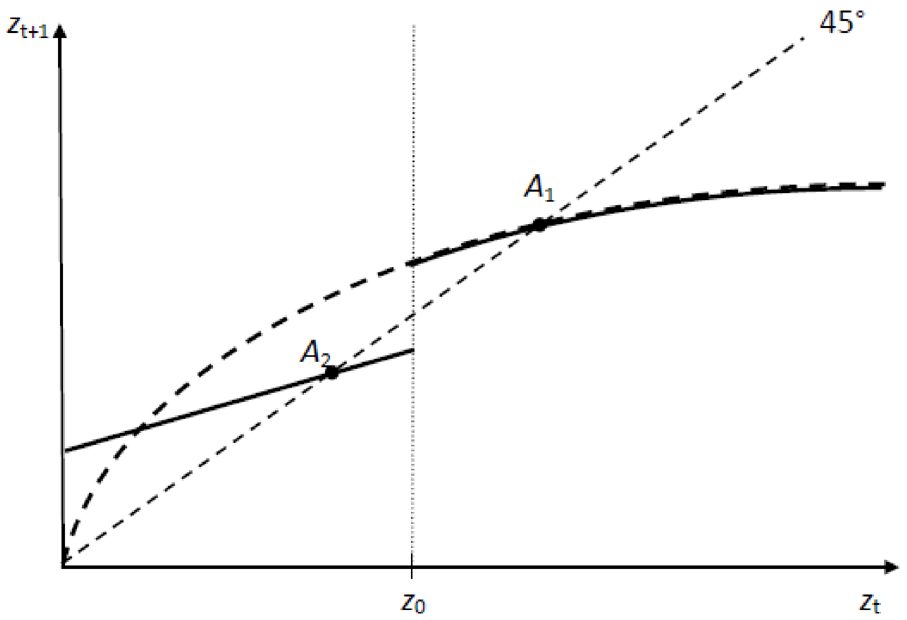
3.4 Quantitative assessment

We have shown that informality may slow down income convergence across countries or be the source of a poverty trap depending on the fact that the model exhibit multiple equilibria or uniqueness. In this section, we confront the theory with the data, calibrate the model, and discriminate between these two hypotheses. We use the stylized facts presented in the introduction and other consensual parameters found in the empirical literature.

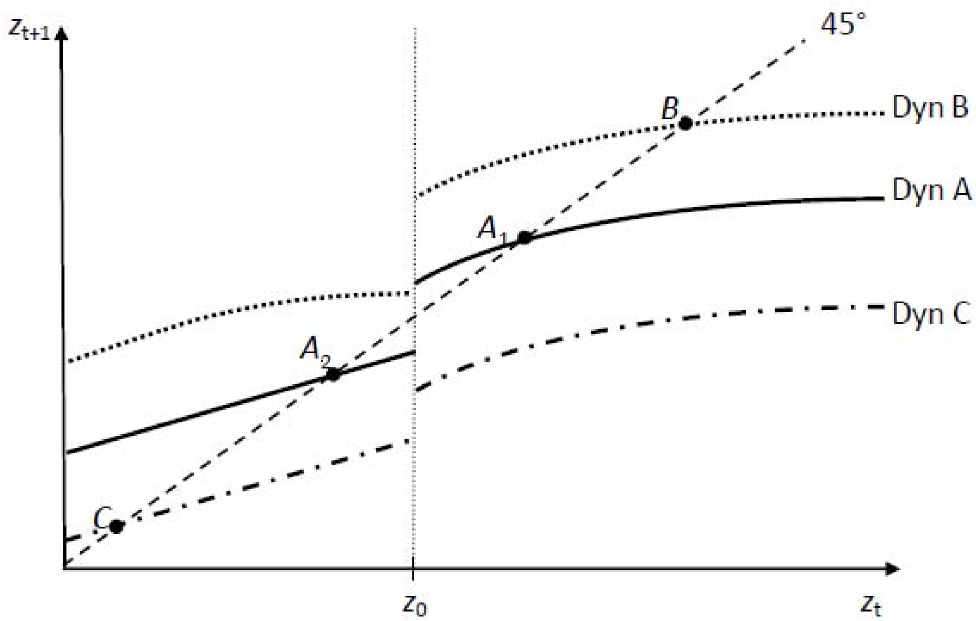
3.4.1 Parametrization

Our parametrization strategy is based on the following principles:

Figure 3.2: Dynamics of human capital accumulation with informality



(a) Dynamics with and without informality



(b) Different configurations with informality

- Parameters are calibrated so as to be compatible with developed and developing economies' observations. In particular, we require our calibrated model to be compatible with the stylized facts described in the introduction.
- The United States situation is considered as a possible steady state without *poverty-induced* informality.
- Least developed countries might be out of steady state and are characterized by the *informality* regime.
- Developing countries and the United States share the same exogenous characteristics: A_0, e, η, α, γ, β, n_h and n_l.

The model is calibrated under the assumption that one period (or generation) represents 30 years, and that individuals become high-skilled after 10 years of education. As far as production is concerned, we use recent data on skill premium from Hendricks (2004). The average return to one year of college for France, Germany, Japan, the United Kingdom and the United States males was around 6 percent in the late nineties. Assuming that high-skilled workers have 10 years of education, we have $\sigma_t^{US} = 0.79$ for the United States economy. According to Barro and Lee data, the United States proportion of workers with at least one year of college was around 50 percent in 2000, so $h^{US} = 0.5$. Barro and Lee report that between 1950 and 2000 the percentage of population with some college studies increased from 13 percent to $48,5$ percent, whereas from 2000 to 2010 it just increased to 51.8 percent. It seems reasonable to assume a steady state value of 50 percent of high-skilled workers in developed countries. Using (3.20), we obtain the parameter α to be 0.64. In addition, Hendricks (2004) reports a return to schooling of 15 percent in the least developed countries, or equivalently $\sigma_t^{Poor} = 3.04$. From (3.23), this requires γ to be equal to 2.27, which implies that $\widetilde{\gamma}$ is 0.16 and the threshold proportion of college graduates below which the *informality* regime is observed is $30,6$ percent.

So far, we have obtained the main parameters from the production side, now we turn to the parameters that affect household's decisions. The fertility ratio n of high- to low-skilled workers is set to 0.57 from Kremer and Chen (1995). They show that n does not vary that much with the level of development, it is stable across countries and over time. Moreover, as we can observe in the United States and other developed economies, we assume no population growth, which implies $n_h = 0.73$ and $n_l = 1.27$.[12] Haveman and Wolfe

[12] We assume these parameters to be constant because we depart from fertility decisions although poor countries have higher fertility rates than developed ones in the data so as

(1995) and Knowles (1999) suggest the education cost is around 15 percent of time endowment of parents while children live with parents. This implies that if children live with their parents for 15 years, then $e = 0.048$.[13] The remaining parameters are the weight of children's income on utility and child labor productivity. Assuming that United States economy is in the steady state, from (3.13) and (3.14) we obtain $\beta = 0.26$. And the relative productivity η of children compared to parents is 0.37 to match the empirical evidence presented by Horrell and Humphries (1995) who claim that 25 percent of family income comes from child labor.[14]

3.4.2 The case for multiplicity

Figure 3.3(a) shows human capital dynamics with parameter values obtained in the parametrization subsection. As predicted by Proposition 8, a poverty trap emerges in the presence of informality because the informal sector does not allow high-skilled wages to increase enough so as to encourage education. Moreover, the existence of informality opens the door to child labor.

As can be seen in Figure 3.3(b), human capital dynamics are driven by the proportion q_l of children of low-skilled parents. In the *informality* regime a constant share of children is educated. While in the *formality*, regime this share increases up to a point where parents do not find it profitable to educate so many children, and the proportion of children who provide education by low-skilled adults decreases.

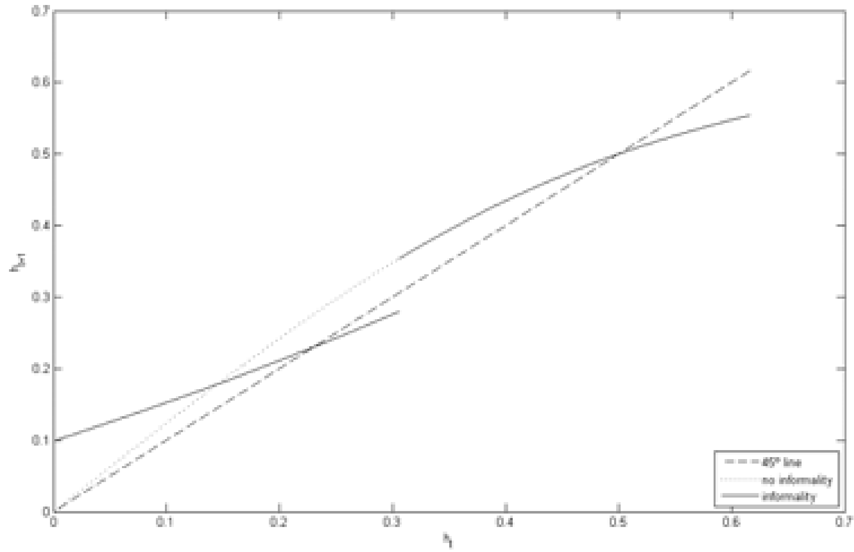
These two Figures explain why the poverty trap emerges. The lack of returns to education and the opportunity cost of sending children to school make the proportion q_l to be lower in the *informality* regime than in the *formality* regime for proportions h_t of high-skilled workers between 15% and 30%. Because agents do not internalize the externality of education on TFP and the low number of highly educated children, the proportion of high-skilled workers remains low and stable over time.

in the model.

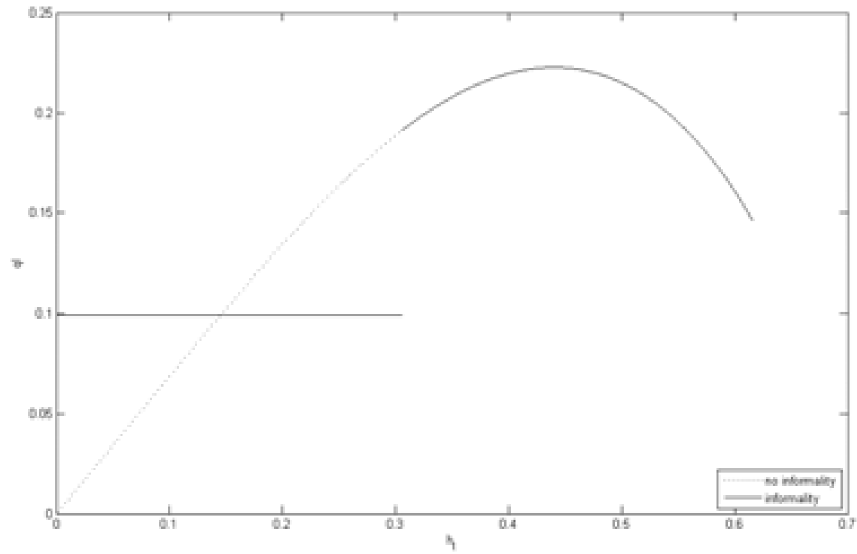
[13]For example, de la Croix and Doepke (2003) assume that children live 15 out of 30 years with parents.

[14]We obtain a relative productivity of children compared to parents higher than Doepke and Zilibotti (2005) who obtain 0.1 to match the same empirical fact. However, Goldin and Sokoloff (1984) claim that that the relative productivity of children and females compared to males rose from around 0.3 in the North (.58 in the South) to .5 from 1820 to 1850, which is in line with our value.

Figure 3.3: Human capital dynamics $h_{t+1}(h_t)$ and proportion $q_l(h_t)$ of educated children with and without informal sector



(a) $h_{t+1}(h_t)$



(b) $q_l(h_t)$

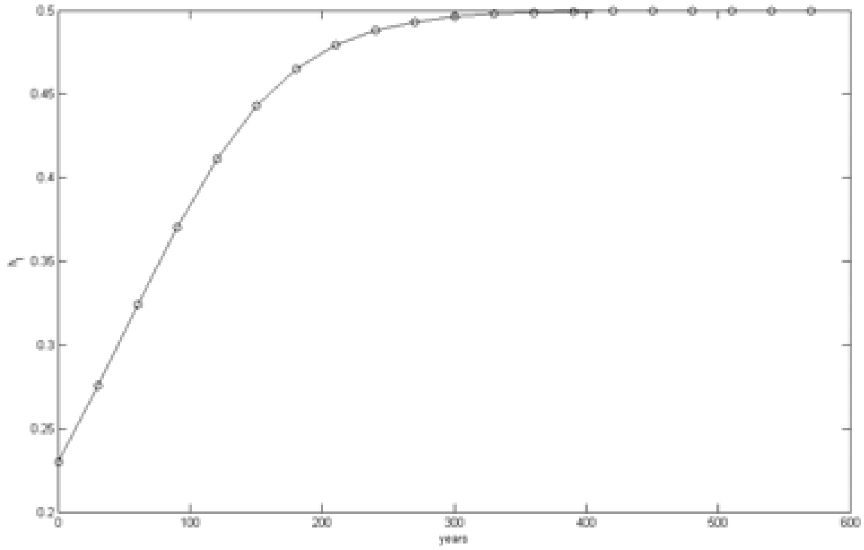
3.4.3 Removing informality

In this section we look at the transition from the low steady state to the high steady state if informality is removed. Figure 3.4(a) shows the transition from the low steady state to the high steady state if we keep the same parameter values obtained in the previous section, but do not allow for the existence of informality. We can observe that the transition would last around 300 years (or 10 periods) to achieve the new steady state. At the same time, we can also observe that after 3 periods the proportion of high-skilled individuals is higher than the threshold value that defines the *informality* regime.

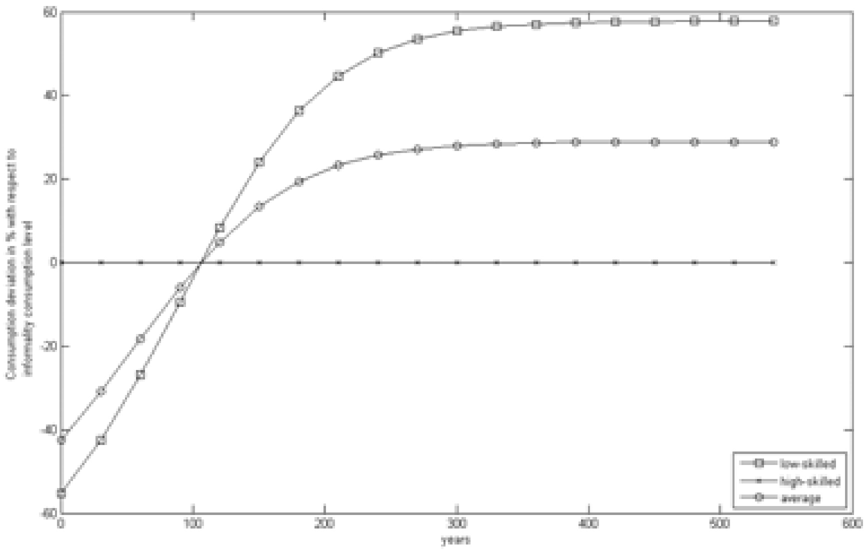
The question that follows is how removing informality would affect individuals in this economy. In Figure 3.4(b) we compute the welfare loss as the percentage of consumption deviation with respect to the consumption level observed in the steady state with informality. To be more precise, in every period t, we compute the percentage deviation as $100*(c_{k,t}-c_k^{inf})/c_k^{inf}$, where c_k^{inf} is the consumption level of a type-k worker in the steady state with informality, and $c_{k,t}$ is the consumption level in period t. As expected, the consumption level of high-skilled workers is not modified. Their wage is constant along the evolutionary path. For the low-skilled workers, however, consumption falls by more than 50% in the first period that informality vanishes. As time passes, the wage and the consumption level of low-skilled workers increase to overcome the consumption level observed in the steady state with informality. Another pattern that we can depict from Figure 3.4(b) is with regard to the average consumption deviation. In the initial periods, it is relatively closer to the consumption deviation of low-skilled workers than the consumption deviation of high-skilled workers, but it approaches the consumption deviation of the highly skilled over time. Hence, as the economy evolves, the proportion of high-skilled workers increases, and the weight of the low-skilled workers on the average deviation diminishes.

Clearly, the first period welfare loss is due to the flow of low-skilled workers from the informal to the formal sector. As the economy evolves, the higher proportion of high-skilled adults in the formal sector increases wages of low-skilled workers. Two different effects produce the raise of wages and consumption over time, the complementarity between high- and low-skilled workers on the one hand, and the increase in TFP on the other hand. Informality protected poor and less educated adults from a sharp wage cut in the short run but prevented the accumulation of human capital necessary to observe economic growth in the long run.

Figure 3.4: Transition from informality to formality



(a) Transition from low to high steady states



(b) Welfare loss due to transition

3.4.4 Implications for development policy

In the previous sections we established the result that the existence of an informal sector combined with human capital externalities can generate a poverty trap. We also showed that if informal activities were rendered illegal, low-skilled workers would suffer initially a quite dramatic drop in wages. In this section, we analyze policies that could help the economy to escape the poverty trap and converge towards the high-income steady state. We examine the cost-efficiency of such policies under the constraint that wage losses during the transition should be avoided.

We consider the situation of a developing country trapped in the low-income steady state and assume that it will obtain a windfall gain (which might come from different sources, e.g. foreign aid or the discovery of natural resources).[15] How can such a windfall gain be used in the most efficient way in order to escape the poverty trap? To answer this question, we analyze different policy instruments that address the human capital externality and the child labor trap, and compare their discounted costs. We first consider each instrument in isolation and then examine whether a combination of two instruments may be a cheaper alternative.

Alternative policies. On one hand, we consider the introduction of *education subsidies* that are either paid unconditionally to all families or targeted to low-skill parents. The latter policy can be interpreted as the education component of existing conditional cash transfers.[16] On the other hand, we analyze *wage subsidies* for jobs in the formal sector, allowing for different subsidy rates for low-skill and high-skill jobs.[17] To sum up, we introduce the following policy variables in the model:

- an education subsidy at rate s_t^e (paid to all families or targeted to low-skilled parents);
- a wage subsidy for low-skilled workers in the formal sector at rate s_t^l ;

[15]In the case of a resource-rich country, it would have to be assumed that the natural resource sector operates independently from the rest of the economy, excluding thereby Dutch disease effects.

[16]E.g., the *Oportunidades/Progresa* program in Mexico or the *Bolsa Familia* scheme in Brazil. These programs are targeted towards low-income families and provide grants for children conditional on school attendance.

[17]Equivalently the government could implement a combination of an output subsidy in the formal sector and a (progressive) tax on income from the formal sector. An output subsidy has the same effect as subsidizing high-skilled and low-skilled workers in the formal sector at the same rate. Adding a progressive income tax would be equivalent to differentiating the effective subsidy rates received by high and low-skilled workers.

- a wage subsidy for high-skilled workers in the formal sector at rate s_t^h.

From the assumptions of the model it is immediately clear that it would be inefficient to pay education subsidies to high-skilled parents since they educate all their children even without receiving any subsidies. Hence, the general education subsidy is less cost-efficient than the targeted education subsidy. As we will show below, the wage subsidy for high-skilled workers has similar effects as an education subsidy to all parents. This type of wage subsidy is therefore also dominated by the targeted education subsidy.

Policy effects in the *informality* regime. In the *informality* regime, the introduction of subsidies does not change the income of low-skilled workers. Subsidizing low-skilled workers draws them into the formal sector but as long as the informal sector exists, the low-skill wage is determined by the (exogenous) productivity in the informal sector. By contrast, the income of high-skilled workers is increased one-by-one by the subsidy. Hence, wages (including subsidies) and the skill premium in the *informality* regime are

$$\begin{aligned}\widetilde{w}_{h,t} &= A_0\alpha(1+s_t^h),\\ \widetilde{w}_{f,t} = \widetilde{w}_{i,t} &= \frac{A_0(1-\alpha)}{\gamma},\\ \widetilde{\sigma}_t &= \frac{\alpha\gamma}{1-\alpha}(1+s_t^h)-1.\end{aligned}$$

The number of low-skilled adults working in the formal sector is given by $L_{f,t} = (1+s_t^l)\gamma H_t$. The informal sector disappears if marginal cost exceeds the price of its output, i.e., if $L_{f,t}/H_t \leq (1+s_t^l)\gamma$. This condition is equivalent to

$$h_t \geq \frac{1}{1+\gamma(1+s_t^l)}. \tag{3.30}$$

The role of the two types of wage subsidies in the formal sector can now be made clear. Subsidizing high-wage jobs increases the skill premium but has no effect on the allocation of workers between sectors. By contrast, a subsidy for low-wage jobs in the formal sector does not affect the skill premium but lowers the critical human capital level at which the economy leaves the *informality* regime.

In turn, the budget constraint of adults is modified by the introduction of an education subsidy as follows:

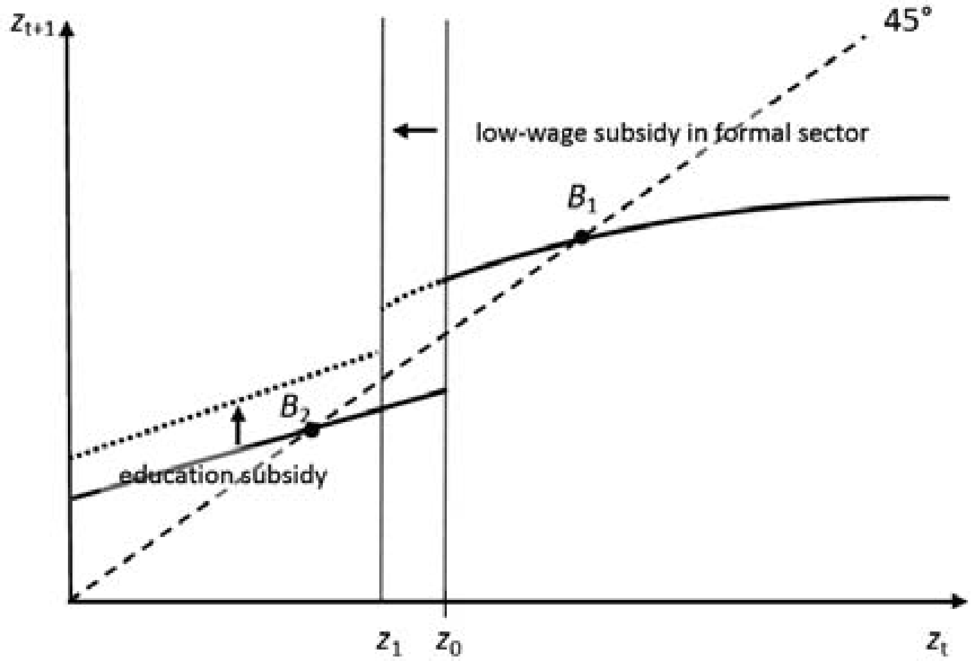
$$c_{k,t} = \widetilde{w}_{k,t} - n_k q_{k,t}\tilde{e}(1-s_t^e) + \eta n_k(1-q_{k,t})d_t\widetilde{w}_{i,t}.$$

The proportion of children who go to school is therefore equal to

$$q^*_{l,t} = \frac{\beta(1-\alpha)(1+n_l\eta)}{(1+\beta)\left[e\gamma(1-s^e_t)+\eta(1-\alpha)\right]n_l} - \frac{1-\alpha}{(1+\beta)\left[\alpha(1+\gamma(1+s^h_{t+1}))-1\right]}.$$

Subsidizing high-skilled workers in the next generation $(t+1)$ has similar qualitative effects as subsidizing education for the current generation t. Obviously, an expected rise in the future skill premium increases the incentive to send children to school. There is however a decisive difference between the two types of subsidies: an education subsidy can be targeted towards low-skilled parents and is therefore more cost-effective (since high-skilled parents educate all their children even without subsidies). Moreover, subsidizing the wages of relatively rich workers rather than the education of poor children seems politically less feasible.

Figure 3.5: Dynamics of human capital accumulation: the role of policies



The preceding results enable us to highlight the different (and possibly complementary) roles of the two most promising policies: targeted education subsidies and wage subsidies for low-skilled workers in the formal sector (see Figure 3.5). If the economy is initially stuck in the inferior steady-state (B_2), the introduction of targeted education subsidies increases the incentive of low-skilled parents to invest in their children's education and the informal sector schedule shifts upwards in Figure 3.5. If the subsidy rate is sufficiently high, the country can escape the poverty trap with the help of this single policy instrument; the new situation of the economy could then be described by Dynamic A in Figure 3.2(b).

By contrast, the subsidy for low-skilled workers in the formal sector pulls workers out of the informal sector and decreases the critical skill ratio from z_0 to z_1 in Figure 3.5 without changing the *informality* schedule. It is clear that such a low-wage subsidy has no effect on human capital accumulation if it is too small or if the economy is too far below the critical skill ratio; the subsidy rate must be sufficiently high to eliminate informal sector employment entirely. Wage subsidies should therefore only be used as a temporary policy allowing the transition to the *formality* regime to accelerate.

As the two types of subsidies address different aspects of the transition to the high income equilibrium, they can be implemented jointly and their combined use might possibly reduce the overall cost of escaping the poverty trap. This issue will be taken up below in the simulations. In any case, we assume that subsidies are abolished as soon as the economy reaches the *formality* regime.[18]

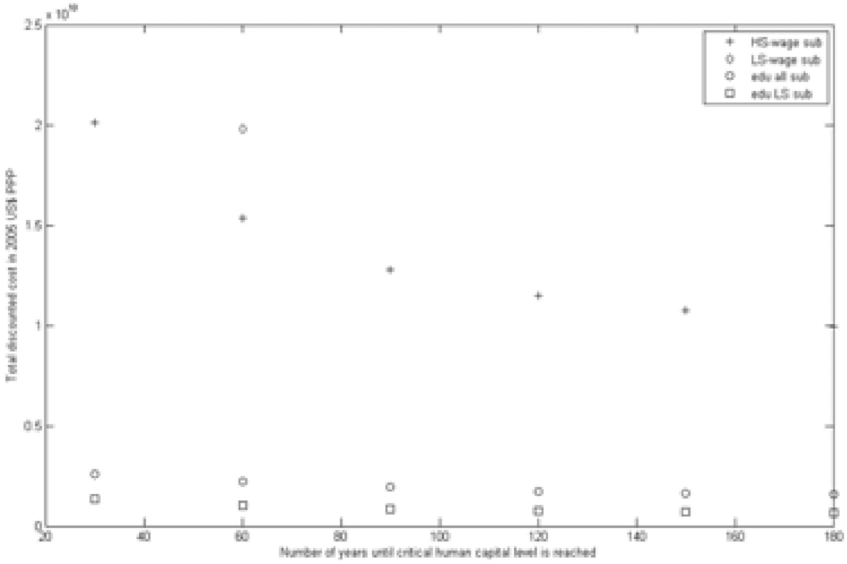
Cost-efficient policies. The calibrated model can now be used to calculate, for each policy, the minimum windfall gain necessary to enable the country to escape the poverty trap. This windfall gain (or discounted cost of policy) depends on the time horizon within which the economy leaves the *informality* regime. Consider a constant subsidy of each type, $s_t^k \equiv s^k$ for $k \in \{e, l, h\}$. The horizontal axis of Figure 3.6 indicates the time needed to achieve a level of human capital that ensures convergence to the high steady state, or equivalently, the time T needed to achieve a proportion of high-skilled workers higher than the threshold value $h_T > 1/(1+\gamma)$ delimiting the two regimes. The vertical axis of Figure 3.6 shows the total discounted cost of the policy for a country with an initial population of 20 million inhabitants, an initial TFP A_0 of 70000, and a discount factor equal to $.99^{120} \cong 0.2994$.[19]

[18]To avoid clutter, Figure 3.5 does not depict the policy-induced change in the dynamics of the *formality* regime. The two policies have different effects on the *formality* schedule. Whereas education subsidies shift the *formality* schedule unambiguously upwards, the introduction of low-skill wage subsidies has ambiguous effects: a positive income effect (low-skill parents receive a higher income which is partly spent on education of their children) and a negative substitution effect (low-wage subsidies decrease the future wage differential, diminishing the incentive for education). These changes do not seem to have a decisive influence on the transition from the informality to the *formality* regime.

[19]The parameter A_0 is set to 70000 to obtain that GDP per capita in the United States is 35000 in 2005 US$ PPP adjusted, which is close to the value in PWT 7.0. The discount factor is obtained from the literature taking into account that a period lasts 30 years and

As expected, targeted education subsidies are more cost-efficient than unconditional education subsidies or high-skill wage subsidies at any time horizon. A windfall of 1 to 1.5 billion 2005 US$ (PPP adjusted) is needed to help a country of around 20 million inhabitants escape from the poverty trap within one or two generations (30 or 60 years). As the initial skill ratio of this economy is far below the critical level, low-skill wage subsidies are very inefficient if they are used as a single policy instrument.

Figure 3.6: Total discounted cost of policies and time necessary to achieve $h_t > 1/(1+\gamma)$



Moreover, as Figure 3.6 makes clear, policies that take more time to leave the *informality* regime have lower discounted costs. Consider for example education subsidies targeted to low-skilled parents. The total discounted cost of attaining the critical human capital ratio is lower if the policy is implemented over several generations using a low subsidy rate (by opposition to a high-subsidy policy which operates within one generation). The reason for this result is twofold. Firstly, within a generation the marginal cost of subsidizing education increases with the proportion of children that are educated. Secondly, targeted education subsidies have a cumulative impact over time: in each generation, they provide an incentive to low-skilled adults to educate a larger proportion of their children. In the following generation, these high-skilled children will provide education to all their offspring although they do

the discount factor of a quarter of year is .99.

not receive the (targeted) education subsidy.

A similar result holds for low-skill wage subsidies in the formal sector: a marginal increase in formal employment of (low-skilled) workers is obtained at the cost of paying higher subsidies to *all* (low-skilled) formal sector workers, including the infra-marginal workers. Thus, the cost of eliminating the informal sector is a quadratic function of low-skilled employment. Indeed, the informal sector disappears if the low-skill wage subsidy is fixed at the rate $s_t^l = (1/\gamma)(\overline{L}_t/\overline{H}_t)$.[20] Therefore, the cost of eliminating the informal sector within the current generation is $s_t^l \overline{L}_t = (1/\gamma)(\overline{L}_t^2/\overline{H}_t)$.

Combination of policies. The preceding results leave scope for a cost-reducing combination of policies. As the marginal cost of a single policy increases with its rate, it might be more cost-efficient to combine two instruments using lower rates. We explore this possibility by combining targeted education subsidies and low-skill wage subsidies. As we have argued above, the latter should only be used as a transitory measure. In the simulations reported in Figure 3.7, low-skill wage subsidies are only used if they enable the economy to reach the *formality* regime within the next generation.[21] Therefore low-skill wage subsidies are only implemented during one generation and the subsidy rate is set such as to make the informal sector disappear within this generation. Figure 3.7 shows the minimum cost of reaching the higher income equilibrium either by using only targeted education subsidies or by combining the two policy instruments. The combination of the two policies is cheaper than the single instrument for time horizons that exceed three generations (90 years). Note that for slightly richer countries (that are closer to the critical skill ratio), a combination of the two policy instruments is likely to be more cost-efficient even for shorter time horizons.

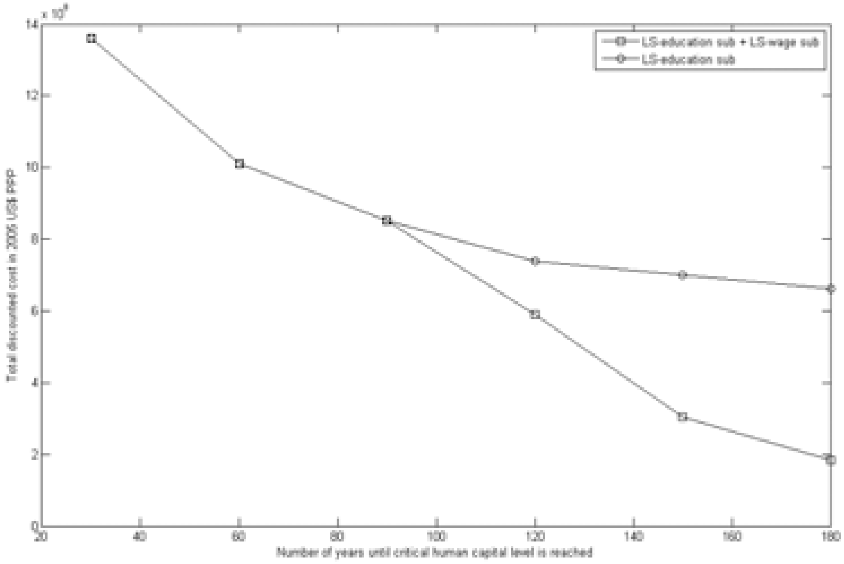
Our policy findings can be summarized as follows. First, among four possible education and wage subsidy schemes, two policies dominate the others in terms of cost efficiency: education subsidies to low-income families and wage subsidies for low-skill jobs in the formal sector. Second, these two policies play distinct and possibly complementary roles in the transition to the high-income equilibrium. Whereas the education subsidy speeds up the accumulation of human capital, the low-skill wage subsidy reduces the threshold at which the informal sector disappears. Third, targeted education subsidies are the cheapest single policy but for medium time horizons a combination

[20] It is the subsidy rate that makes condition (3.30) binding.

[21] Alternatively, one could assume that the subsidy is phased out gradually if it takes several generations to attain the critical human capital level. This possibility is disregarded in our search of the cheapest policy combination.

of the two policies turns out to be the most cost-efficient choice.

Figure 3.7: Cheapest combination of policies and time necessary to achieve $h_t > 1/(1 + \gamma)$



3.5 Conclusion

This paper establishes a theoretical relation between education, child labor and the informal sector. In the data we observe a direct relation between informality and education, countries with high proportions of tertiary educated workers tend to show lower levels of informality than countries with low proportions. Moreover, child labor is part of the informal sector, and the data shows that countries with more informality have more children involved in production activities. With these facts in mind we construct an overlapping generations model that is able to reproduce these relations in line with previous findings of other authors.

The model is able to explain, or to give a complementary view of, the documented fact that low-developed countries present higher levels of inequality than developed countries but much lower than standard models predict. The introduction of the informal sector in a model with complementarity between high- and low-skilled workers makes the skill premium lower and constant than without informality. In other words, we view informality as a possible channel to reduce the skill premium in developing countries.

The reduction in inequality due to informality generates several effects in the short and the long run. On one hand, low-skilled workers may obtain a higher salary with the existence of an informal sector than in its absence, because there is an alternative sector where they can supply their working hours. However, this sector is not controlled by state agencies and enables children to use their time to work and generate an extra source of income for the household. Hence, the model is able to replicate the relations between informality and education, and between child labor and education in line with the data, high-skilled workers are negatively correlated with informality, and informality is positively correlated with child labor. On the other hand, the model has several predictions on the long run. The trade-off between child income and future education of children is taken into account and is key to generate poverty traps due to informality and child labor. The "low" inequality observed in developing countries and the opportunity cost of sending children to work instead of going to school can make a pernicious effect on parents. They may not provide enough education for their children so as to increase the aggregate proportion of educated workers in the labor force. Parents do not internalize the positive externality of aggregate education on firms productivity. Therefore, the informal sector can make the economy not to develop as it should in the absence of informality.

The model is calibrated to reproduce several facts throughout the data. The model is also calibrated to evaluate different policies considered to reduce the size and effects of informality. The calibration exercise reveals that the case for the poverty-trap hypothesis is strong: although informality serves to protect low-skilled workers from extreme poverty in the short-run, it prevents income convergence between developed and developing nations. Sudden elimination of informality would induce severe welfare losses for poor people on the transition path.

Hence, we analyze policies that could help the economy to escape the poverty trap and converge towards the high-income steady state. We analyze the cost-efficiency of such policies under the constraint that wage losses during the transition should be avoided. Assuming that an inflow of resources is provided to a developing country, for example in the form of foreign aid, we analyze the effects of different subsidies. One possible way to reduce informality may come from reducing education costs or making the formal sector more attractive, as for example, increasing formal firms wages. Then, we consider four possible subsidies on education and formal firms wages. Subsidizing education is the most cost-effective policy, and it can be targeted towards low-skilled parents to reduce costs. Subsidizing high-wage jobs increases the skill premium but has no effect on the allocation of workers between sectors. Moreover, the increase in the skill premium gives similar incentives to parents on children's education than reducing education costs. By contrast, a subsidy for low-wage jobs in the formal sector does not affect the skill premium but lowers the critical human capital level necessary to skip the poverty trap. Because of the possible complementary effect of different subsidies, we turn to analyze the cost-efficiency of a combination of subsidies on education to low-income parents and low-skilled formal firms wages. Although targeted education subsidies are the cheapest single policy, for medium time horizons a combination of the two policies is found to be the most cost-efficient choice.

References

Alesina, A., and D. Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*, 109(2), pp.465-490.

Angrist, J.D. 1995. "The Economic Returns to Schooling in the West Bank and Gaza Strip." *American Economic Review*, 85(5), pp. 1065-1087.

Barro, R., and J-W. Lee. 2010. "A New Data Set of Educational Attainment in the World 1950-2010." NBER Working Paper No. 15902.

Basu, K., and P. Van. 1998. "The Economics of Child Labor." *American Economic Review*, 88, pp. 412-427.

Becker, G. 1965. "A Theory of the Allocation of Time." *Economic Journal*, 75, pp. 493-517.

Borjas, G., and L. Katz. 2007. "The Evolution of the Mexican-Born Workforce in the United States." *Mexican Immigration to the United States National Bureau of Economic Research Conference Report* (ed. George Borjas), Cambridge, MA.

Buehn, A., Montenegro, C.E., and F. Schneider. 2010. "Shadow Economies All over the World. New Estimates for 162 Countries from 1999 to 2007." Policy Research Working Paper Series 5356, The World Bank.

Calderon, C., Moral-Benito, E., and L. Serven. 2011. "Is Infrastructure Capital Productive? A Dynamic Heterogeneous Approach." Policy Research Working Paper Series 5682, The World Bank.

Card, D., and Th. Lemieux. 2001. "Can Falling Supply Explain the Rising Returns to College for Younger Men? A Cohort Based Analysis." *Quarterly Journal of Economics*, 116(2), pp. 705-746.

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

de La Croix, D., and M. Doepke. 2003. "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review*, 93(4), pp. 1091-1113.

de La Croix, D., and F. Docquier. 2012. "Do Brain Drain and Poverty Result from Coordination Failures?" *Journal of Economic Growth*, 17(1), pp. 1-26.

De Paula, A., and J.A. Scheinkman. 2010. "Value-Added Taxes, Chain Effects, and Informality." *American Economic Journal: Macroeconomics*, American Economic Association, 2(4), pp. 195-221.

De Paula, A., and J.A. Scheinkman. 2011. "The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil." *Review of Income and Wealth*, 57, pp. S8-S26, 05.

Doepke, M., and F. Zilibotti. 2005. "The Macroeconomics of Child Labor Regulation." *American Economic Review*, 95(5), pp. 1492-1524.

Edmonds, E. 2007. "Child Labor." *Handbook of Development Economics*(eds. T.P. Schultz and J. Strauss), 1(4), ch.57, pp. 3607-3710.

Galor, O., and J. Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60, pp. 35-52.

Goldin, C., and K. Sokoloff. 1984. "The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850." *Quarterly Journal of Economics*, 99(3), pp. 461-487.

Hendricks, L. 2004. "A Database of Mincerian Earnings Regressions." www.lhendricks.org/Mincer.htm.

Haveman, R., and B. Wolfe. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33(4), pp. 1829-1878.

Heston, A., Summers, R., and B. Aten. 2011. "Penn World Table Version 7.0." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Horrell, S., and J. Humphries. 1995. "The Exploitation of Little Children": Child Labor and the Family Economy in the Industrial Revolution." *Explorations in Economic History*, 32(4), pp. 485-516.

Katz, L., and K.M. Murphy. 1992. "Changes in Relative Wages 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1), pp. 35-78.

King, L. 1995. "Quantitative Theory and Econometrics." *Federal Reserve Bank of Richmond Economic Quarterly*, 81(3), pp. 53-105.

Knowles, J. 1999. "Can Parental Decisions Explain U.S. Income Inequality?" Mimeo, University of Pennsylvania.

Kremer, P., and D. Chen. 2002. "Income Distribution Dynamics with Endogenous Fertility." *Journal of Economic Growth*, 7, pp. 227-258.

Krugman, P. 1991. "History versus Expectations." *Quarterly Journal of Economics*, 106(2), pp. 651-667.

La Porta, R., and A. Shleifer. 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity*, 39(2), pp. 275-363.

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

Murphy, K.M., Shleifer, A., and R.W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy*, 97(5), pp. 1003-1026.

Perry, G.E., Maloney, W.F., Arias, O.S., Fajnzybler, P., Mason, A.D., and J. Saavedra-Chandeuvi. 2007. *Informality: Exit and Exclusion.* World Bank Latin American and Caribbean Studies 40008: Washington DC.

Rebelo, S. 1991. "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy*, 99, pp. 500-521.

Rodrik, D. 2011. "The future of Economic Convergence." NBER Working Paper No. 17400.

Romer, P. M. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy*, 94, pp. 1002-1037.

Rosenstein-Rodan, P.N. 1943. "The Problems of Industrialization of Eastern and South-Eastern Europe." *Economic Journal*, 53(210/11), pp. 202-211.

Rosenzweig, M., and R. Evenson. 1977. "Fertility, Schooling, and the Economic Contribution of Children in the Rural India: An Econometric Analysis." *Econometrica*, 45, pp. 1065-1079.

Schneider, F. 2005. "Shadow Economies around the World: What do we really know?" *European Journal of Political Economy*, 21(3), pp. 598-642.

Schultz, T. W. 1960. "Capital Formation by Education." *Journal of Political Economy*, 68, pp. 571-583.

World Bank. 2012. Child labor data from World Development Indicators Online (WDI) database.

Zenou, Y. 2008. "Job Search and Mobility in Developing Countries. Theory and policy implications." *Journal of Development Economics*, 86, pp. 336-355.

Chapter 4

Latent Growth in Sub-Saharan Africa

Written jointly with Ayşegül Kayaoğlu.

4.1 Introduction

Growth is one of the main concerns of the international agenda. The case of Africa is of particular interest because it is the region with the lowest levels of GDP per capita. As probably we all know, this low level of wealth and income translates into desnutrition, lack of health facilities, child labor and many other outcomes that affect both life expectancy so as its quality. Although these outcomes are of extreme importance, the case of Africa is of particular interest not only because of these reasons, but also because in the last century, despite the large increases in literacy and urban population growth has not been materialized. Africa is the region with the highest increase in human capital and urban population but has the slowest economic growth. As we will see below this fact is surprising because this is just the contrary to what we have observed in the process of development of the now called developed countries. We name this phenomenon the African paradox.

In developed countries, the Industrial Revolution changed the accumulation of factors of production. Before the Industrial Revolution, agriculture was the leading sector and it progress allowed food surpluses and storage. Moreover, geographical barriers and climate differences prevented the free diffusion of technologies among different regions of the world. All of these factors made agricultural activities to be the source of economic development when food production was the leading sector. With the Industrial Revolution, technological progress allowed economies to adopt a less labor intensive agricultural sector. Moreover, the decrease in the transportation costs and thereby the reduced role of geographical barriers made the diffusion of ideas and new technologies between regions and countries easier than before. This revolution bring importance to the accumulation of human capital because it determines a country's capacity to invent and adopt new technologies. The outcome was a massive migration from rural areas to urban areas. Urban areas are characterized by a higher concentration of people that facilitate the interaction among them and reduce the cost of translating knowledge among them.

Since Lucas (1988) and Azariadis and Drazen (1990), human capital disparities play a central role in the analysis of growth and development. Historical analyzes confirm that the transition from economic stagnation to growth was preceded (Cipolla, 1969) and then accompanied (Maddison, 1995) by enormous increases in literacy and average level of schooling. In addition, it has long been argued that there is a close connection between urbanization and economic growth. Lucas (1988) discussed the leading effects of cities and urban development on national economic growth. Cities are places where most high-skilled workers are located, interacting with each other, innovating and adopting modern technologies. However, urban areas are also places where access to schooling is better. Bertinelli (2003) argues that "urbanization plays a non-negligible role in spending human capital accumulation. Closeness between people favors interactions, which may be at root of spillovers from human capital. In return, incentives to invest in education are reinforced, leading hence to higher levels of education." Lucas (2009) also analyzes with the link between urbanization and human capital accumulation. He builds a model with two sectors (rural and urban) and assumes that more educated workers reside in the urban sector. Hence, according to his model, countries with a large share of their population working in rural sector (due to agriculture still being traditional) have a low ability to absorb technology from leading economies. It implies that migration out of traditional agriculture is crucial for growth and that countries with low initial endowment in human capital, or high proportion of rural workers, will have a late take-off. Moreover, he focuses on the reasons for cross-country spillover effects which play an important role for the growth behavior of economies. After the calibration analysis, he argues that among several economic forces which contribute to the cross-country growth effects, migration from rural to urban areas seems to be a very important factor for the convergence mechanism. This paper also argues that human capital accumulation and urbanization play a central role in the analysis of growth and development.

In such direction, we find in the Millennium Declaration of the United Nations that member states and international organizations agreed to achieve eight human development goals (United Nations, 2008). Achieving 100 percent of enrollment in primary education, reducing illiteracy rates and gender discriminations in the access to schooling are among the top priorities. In addition, the World Bank is committed to promoting sustainable cities and towns that fulfill the promise of development for their inhabitants. The effectiveness of such policies depends on the intensity of causal links between urbanization, education and growth, as well as on the timing of development process.

However, many economist are concerned with the fact that investment in education in different regions of the world has not lead to growth in these regions. Pritchett (2001) documented the negative association between educational investment and output growth rates and asked: *where has all education gone?* Other empirical studies bring into question the existence and the magnitude of the causal impact of education on development.[1] How can this be reconciled with the strong cross-country or historical associations between

[1]See, among others, Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Parente and Prescott (2000), Bils and Klenow (2000), Caselli (2005).

literacy, schooling and development? It is obvious that the causal impact of human capital on TFP growth may require long delays. There are some traditional explanations for this African growth paradox. One of the arguments is that quality of education does not follow the quantity of investment in SSA. Therefore, educational investments cannot be efficiently transferred to productivity gains. Manuelli and Seshadri (2007) show that effective human capital has a strong impact on economic performances when corrected for differences in the quality of education. Another straightforward claim is the impact of congestion costs as negative externalities. Jones (2009) shows that despite rises in educational attainment, technology adoption is slower when knowledge traps are at work: poor countries invest too much in "generalist" education and not enough in "specialist" education, given the coordination cost imposed by a "specialist" economy. In the same vein, Vandenbussche et al. (2006) show that different types of human capital are needed at various stages of development. Pritchett (2001) argues that the institutional environment in poor countries has been sufficiently perverse that the accumulated human capital has been applied to activities that served to reduce economic growth. Furthermore, Easterly and Levine (1997) point out the importance of public policies in the growth processes and argue that "Africa's growth tragedy is associated with low schooling, political instability, underdeveloped financial systems, distorted foreign exchange markets, high government deficits, and insufficient infrastructure."

However, these explanations are not very convincing and also very pessimistic because they lead to the conclusion that investments in Africa so far were deadweight losses. There is, however, a more optimistic explanation for the situation of Africa. In this paper we put forward the hypothesis that rapid urbanization and human capital development in SSA have not yet given rise to high economic growth rates due to temporary urbanization costs. However, they create a latent growth potential which will materialize when urban population and human capital growth rates will be slower. The deceleration will automatically come as the rate of urbanization and proportion of college graduates increase, or could come sooner if development policies become less generous after the redemption date of the Millennium Declaration ($1990 - 2015$).

This paper focuses on the nexus between human capital accumulation and urbanization, and argues that low or negative social return to education observed in the short run might be due to transitory adjustment or urbanization costs. High-skilled workers mainly operate in cities and more education increases labor demand for low-skilled employees in urban areas. Urbanization makes access to schooling easier, increases schooling level, accentuates the urbanization process (virtuous circle), but generates adjustment and congestion costs. Increasing cities' size, number of firms and urban employment requires enormous public infrastructure investments which affect urban quality of life, in particular, health, safety, commuting, and congestion costs. Moreover, rapid urbanization has often occurred in the face of low or negative economic growth over some decades, and over- or under-concentration can be very costly in terms of productivity growth.

Thus, the aim of this paper is to calibrate a model with endogenized human capital accumulation, urbanization and economic development. An adjustment-cost hypothesis is put forward to explain the African paradox. According to this hypothesis, low or negative social return to education in the short run might be due to the transitory adjustment or urbanization costs. We then confront data to theory and calibrate the parameters of our model using panel regressions and identification strategies. Hence, we can analyze the dynamics of the model, which enables us to make predictions about growth and convergence. The calibrated model does an excellent job in fitting historical data on the proportion of college graduates, share of urban population and growth in GDP per capita. Therefore, the model can be seen as a good source to predict the evolution of the key variables for the next decades.

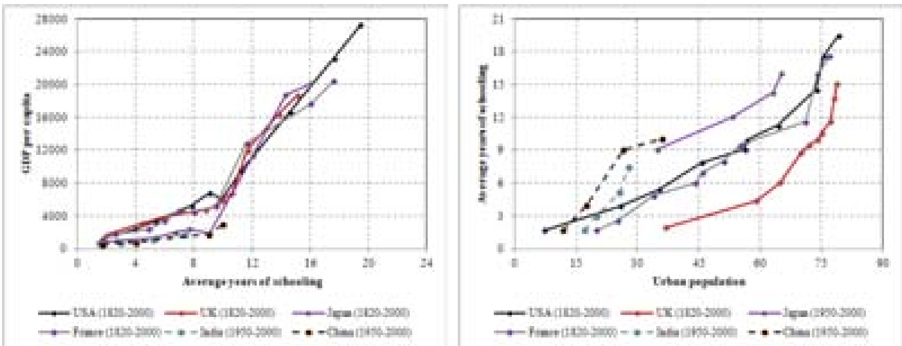
Simulation exercises show that countries in each region converge towards a specific steady state. In other words, only conditional convergence can be obtained in the long-run and developing countries do not catch up the leading region (high-income countries). Besides, we find sub-Saharan Africa to be a distinct case. Comparatively higher levels of educational investment in that region could not be realized in the short-run in terms of economic growth. In fact, many explanations such as institutional structure, political instability, low investments in infrastructure and so on have make an attempt to explain this paradox. This paper argues that these investments in education are not a deadweight loss even though they have not generated growth in the short-run but instead they hide a latent growth potential, and this time lag of realization is due to temporary adjustment costs of urbanization on human capital. However, apart from the optimistic prediction of the African latent growth, our analysis foresees also a pessimistic result, which is the no convergence in income levels in sub-Saharan Africa with respect to the leading regions.

The remainder of the paper is organized as follows. Section 4.2 provides some stylized facts on urbanization, education and development. Section 4.3 describes the model. A numerical analysis is carried out in Section 4.4. Finally, Section 4.5 concludes.

4.2 Stylized facts

Cipolla (1969) documents that the spread of literacy started between the 17th and 19th century, 5,000 years after the first rudimentary appearance of writing. Before that period, the arts of writing and reading remained the monopoly of small elites. But it is mainly in the 19th century that the advance of literacy and the development of education occurred in the west, and it was invariably connected with the condition of urbanization, the emergence of public schools, and the industrial revolution.

Figure 4.1: Historical association between urbanization, education and development



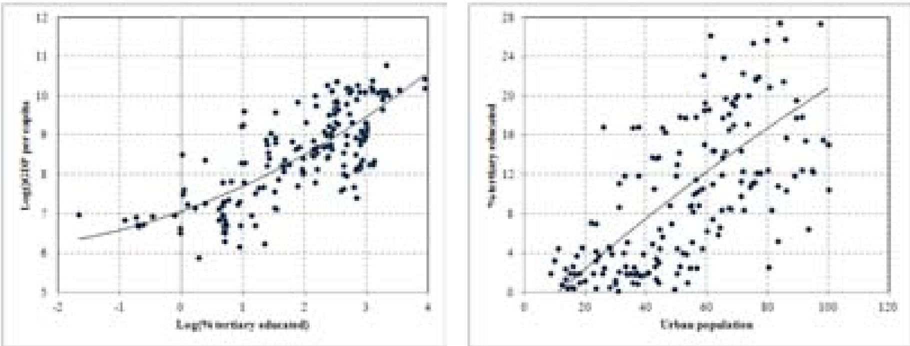
(a) Historical data on GDP per capita and education

(b) Historical data on education and urban population

Data Sources: GDP per capita in PPP is taken from Maddison (1995). The rest of the data are obtained from census. Observations are available for 1820, 1870, 1890, 1913, 1929, 1938, 1950, 1973, 1990, 1998.

Causation is obviously hard to establish. On the one hand, urbanization facilitated access to schooling and the industrial revolution drastically sparked the demand for skills and human capital. On the other hand, education increased workers' capacity of adaptation to new technologies and faculty to innovate. Figures 4.1(a) and 4.1(b) illustrate historical associations between education and GDP per capita, urbanization and education. We report data from 1870 to 2000 for four industrialized countries (the US, the UK, Italy and France) and data again from 1870 to 2000 for Japan, India and Chile. Figure 4.1(a) shows that the rise in years of schooling preceded the rise in income. Figure 4.1(a) also shows that accumulation of human capital stimulates the GDP per capita and therefore the growth of GDP per capita becomes much higher when countries have better educational prospects. Figure 4.1(b) shows that urbanization increased at a faster pace than education in the early stage of development. This conclusion cannot be applied to developing countries such as India. England and the US were clearly a pioneer in the processes of urbanization and adult literacy. Country-specific factors such as the population density and religion (e.g. protestantism) may have facilitated the takeoff.

Figure 4.2: Cross-country association between urbanization, education and development in 2000



(a) Cross-country data on education and GDP per capita

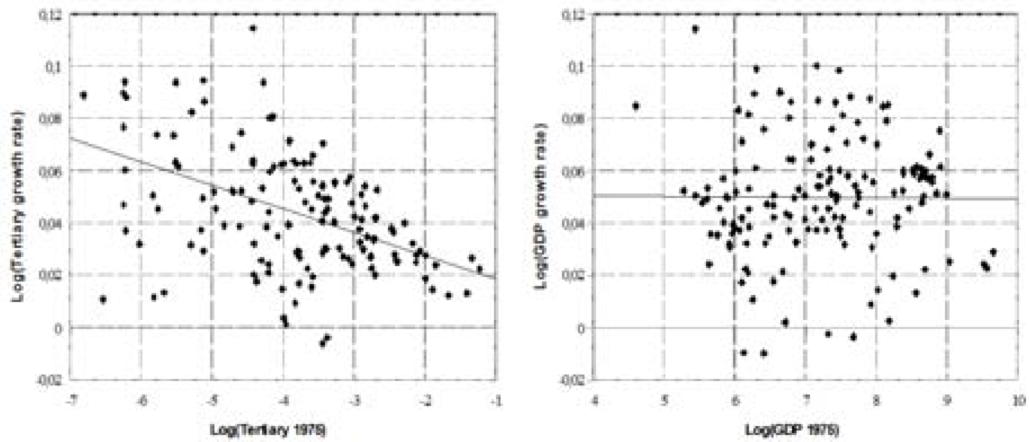
(b) Cross-country data on education and urban population

Data Sources: GDP per capita in PPP is taken from Penn World Tables; proportion of college graduates in the population aged 25 and over is taken from Docquier et al. (2007); and urban population as a percentage of total population is taken from World Bank World Development Indicators (2008).

The same patterns arise when using a cross-country perspective based on recent data. Figure 4.2(a) shows the association between the proportion of college graduates in the population aged 25 and over and the level of GDP per capita in 2000 (both in logs). In our sample of 177 countries the coefficient of correlation is close to 0.60. Turning to the association between the urbanization rate (proportion of population living in urban areas) and the proportion of college graduates, Figure 4.2(b) shows a coefficient of correlation that amounts to 0.40. From these stylized facts, we can conclude that promoting education and urbanization should help developing countries to take off.

A quick look at the data also reveals that the links between urbanization, education and growth are more complex. Figure 4.3 studies convergence in income and human capital. In Figure 4.3(a), we regress the average annual growth rate 1975 – 2000 of the proportion of college graduates in the adult population on its level of 1975 (in logs). The slope is clearly negative and the speed of convergence of human capital is equal to one percent per year. Figure 4.3(b) does the same exercise on GDP per capita. There is no sign of convergence: the 1975−2000 growth rate is independent on the level observed in 1975. We could add that urbanization (proportion of population in urban areas) tripled in sub-Saharan Africa between 1950 and 2000 (from 11 to 33 percent), while it was multiplied by 1.4 in high-income countries (from 53 to 73 percent), 1.9 in South-Central Asia (from 16 to 30 percent), and 1.1 in Latin America (from 73 to 80 percent).

Figure 4.3: Convergence in income and human capital

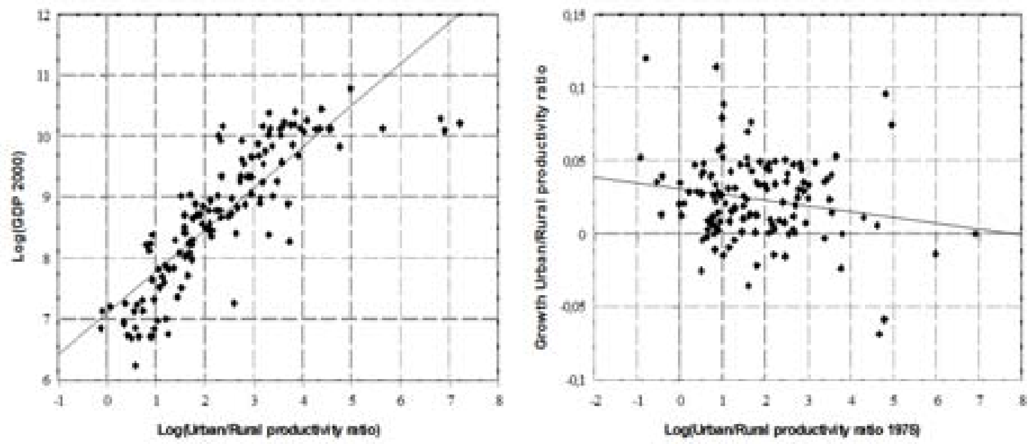


(a) Human capital (b) GDP per capita

Notes: On Figure 4.3(a), we plot the log of the proportion of college graduates in the population aged 25 and over as a function of the log of their 1975 initial value. Similarly, on Figure 4.3(b), we plot the log of the average annual rate of growth of GDP per capita from 1975 to 2000 as a function of the log of their 1975 initial value (beta-convergence analysis). A negative slope (resp. positive slope) reflects convergence (resp. divergence).

We also look at the stylized facts on the development and the urban-to-rural productivity ratio. First of all, we check the association between the urban-to-rural productivity ratio and the level of GDP per capita. As we can observe in Figure 4.4(a), the slope is clearly positive. Moreover, in Figure 4.4(b), we regress the average annual growth rate between the years 1975 and 2000 of the urban-to-rural productivity ratio on its level in 1975 (in logs). To construct the urban-to-rural productivity ratio, we divide total GDP in PPP in the urban sector in 2000 by the total GDP in PPP in rural sector at the same year. Moreover, total GDP in PPP in each sector is calculated by the product of the total GDP in PPP in a sector with the value added of that specific sector to GDP. Data on agricultural value added (as % of GDP) is obtained from WDI (2008). As it can be seen from the graph, the slope is negative and, thus, there is convergence of urban-to-rural productivity ratios among countries.

Figure 4.4: Development and urban-to-rural productivity ratio

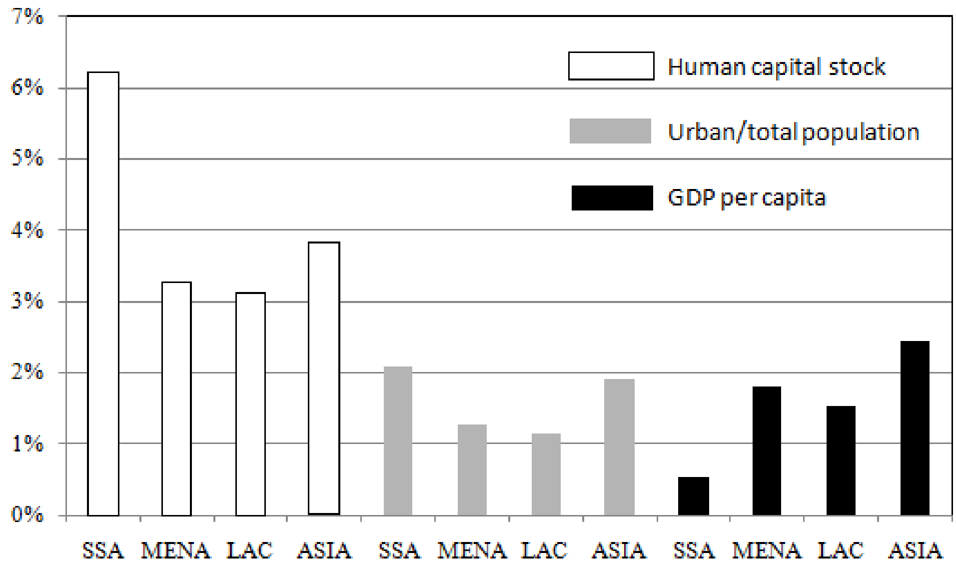


(a) Development and Productivity Ratio (b) Convergence in Productivity Ratio

Notes: On Figure 4.4(a), we plot GDP per capita in 2000 (in logs) against the urban-to-rural productivity ratio in 2000 (in logs). And on Figure 4.4(b), we plot the growth of the productivity ratio between 1975 and 2000 as a function of the log of their 1975 initial value (beta-convergence analysis). We also provide the linear fitted curve. A negative slope (resp. positive slope) reflects convergence (resp. divergence).

The stylized facts described above demonstrate that economic development is connected with urbanization and education. However, Figures 4.3 and 4.4 show that there is no convergence in GDP even though we observe a convergence pattern in urban/rural productivity ratio and educational investment in tertiary level which have a positive impact on GDP. Hence, it is also a paradox that, in the second half of the past century, even the highest progress in schooling and urbanization was observed in sub-Saharan Africa (henceforth referred to as SSA), it is still the region with the lowest GDP per capita growth rates. As shown on Figure 4.5 below, the proportion of college graduates grew by 6.2 percent a year in SSA between 1955 and 2000, to be compared with 3.2 percent in the Middle East and Northern Africa (MENA), 3.1 percent in Latin America and the Caribbean (LAC) and 3.8 percent in South Asia (ASIA). Over the same period, the urban-to-total population increased by 2.1 percent in SSA, to be compared with 1.3 percent in MENA, 1.1 percent in LAC and 1.9 percent in ASIA. Surprisingly, the annual GDP growth rates were 0.5 percent in SSA, 1.8 in MENA, 1.5 percent in LAC and 2.4 in Asia.

Figure 4.5: Average annual growth rate of human capital, urban population and GDP per capita in developing regions (1955-2000)



Definitions and data sources: Human capital = proportion of college graduates in the population aged 25 and over (Barro and Lee, 2009); Urban/total population = urban population / total population (World Urban Population Prospects, United Nations); GDP per capita in constant 2000 US$ (World Development Indicators); SSA = sub-Saharan Africa; MENA = Middle East and Northern Africa; LAC = Latin America and the Caribbean; ASIA = South Asia.

4.3 Model

In this section, we describe a theoretical model of endogenous human capital formation, urbanization and development. Then, in the next section we confront the theory to the data to calibrate the general parameters of the model.

Consider an economy with an agricultural and an urban sector producing a single homogeneous good. The price of the homogeneous good is the numeraire. There are two types of individuals in this economy, the highly educated and the less educated. We use h_t to denote the proportion of highly educated workers at time t. Assume that highly skilled people work in the urban sector, whereas the remaining $1 - h_t$ less educated workers can freely choose between the two sectors. Hence, less educated workers either work in the agricultural sector in rural areas or in low-skilled jobs in urban areas.

Since we mainly focus on the role of urbanization, we formalize human capital accumulation using the following predetermined process:

$$h_{t+1} = a_t h_t^{1-\beta} H_t^{\beta} \phi(u_t), \tag{4.1}$$

where H_t is the proportion of highly educated workers in the leading countries, β is the speed of convergence towards the long-run equilibrium, $\phi(u_t)$ is an increasing function of the variable u_t, which measures the degree of urban concentration of less educated workers, and a_t is a scale factor representing the quality and quantity of education infrastructure in the country. Looking at the urban concentration of less educated workers is equivalent, in our model, to look at the urban concentration of overall people in the country because all high-skilled workers are assumed to work in the urban sector. Hence, we consider the flow of ideas among people in two dimensions. On the one hand, we take into account the transmission of knowledge from leading countries to developing countries. And on the other hand, we consider the concentration of people in urban areas as a mean to transfer knowledge among people out of the traditional agricultural sector.

Equation (4.1) is compatible with the stylized facts sketched above: there are convergence forces guiding the dynamics of human capital, and urbanization facilitates the access to schooling. Lucas (2009) uses a similar hypothesis. The econometric calibration exercise will guide the functional form $\phi(u_t)$ for the urbanization effect.

At any time t, highly skilled workers, in proportion h_t, are entrepreneurs operating in the urban area. Each of them hires ℓ_t less educated workers to produce $y_{ut} = A_{ut}\ell_t^{\alpha}$ units of output, where A_{ut} is the total factor productivity in the urban sector, and $\alpha \in [0, 1]$ is a parameter of decreasing marginal productivity of labor in the urban sector. The labor market for less educated workers is competitive and the urban wage rate equals the marginal productivity of labor. Moreover, when the proportion of entrepreneurs increases, each entrepreneur incurs a congestion cost per firm c_t, which is proportional to the change in the number $h_t - h_{t-1}$ of new firms and divided by the number h_t of entrepreneurs,

$$c_t = qA_{ut}\frac{h_t - h_{t-1}}{h_t}.$$

Note that the introduction of A_{ut} in the cost function reflects the higher costs associated to produce in a more productive and specialized urban sector. Furthermore, note that c_t is equal to zero at any steady state.

The profit function of each entrepreneur is

$$\pi_{u,t} = A_{ut}\ell_t^{\alpha} - w_{ut}\ell_t - c_t,$$

provided that $\pi_{u,t} \geq 0$, otherwise entrepreneurs do not have incentives to produce. Congestion costs are also deducted from the earnings so as to reach the profit per firm. These congestion costs can be interpreted as either the opportunity cost of time that is spent in training the immigrants from rural areas, so as the cost of reduction in market share due to increase in firm number in cities.

Maximizing $\pi_{u,t}$ with respect to ℓ_t determines the equilibrium wage rate

$$w_{ut} = \alpha A_{ut}\ell_t^{\alpha-1}. \tag{4.2}$$

Then, the profit rate in the urban sector is

$$\pi_{u,t} = (1-\alpha)A_{ut}\ell_t^{\alpha} - c_t.$$

We can observe that in the steady state ($c_t = 0$), income inequality between highly educated and less educated workers is

$$\pi_{u,t}/w_{u,t} = \frac{\ell_t(1-\alpha)}{\alpha}, \tag{4.3}$$

which is increasing in ℓ_t.

Less educated can work either in the urban sector, described above, so as in the rural sector, where productivity of each worker is $w_{ft} = A_{ft}$. We assume that less educated workers are freely mobile between sectors and are allocated so as to equalize net wages. Thus, the equilibrium number of less educated employees per entrepreneur is

$$\ell_t = \left(\frac{\alpha A_{ut}}{A_{ft}}\right)^{1/(1-\alpha)}. \tag{4.4}$$

Note that the number of less educated workers is bounded by $(1-h_t)$. Hence, in case that the total demand $\ell_t^D = h_t\ell_t$ of workers in the urban sector is higher than $(1-h_t)$, the equilibrium number ℓ_t of less educated workers per entrepreneur is $(1-h_t)/h_t$, wages in the urban sector are given by (4.2) with $\ell_t = (1-h_t)/h_t$, and the rural sector disappears. This case would happen if the wage in the urban sector for low skilled workers is higher than the one in the rural sector. We disregard this case because it requires no rural population, which only happens in a very limited number of countries such as Singapore or Hong Kong, which are small in land area terms and highly industrialized.

From the previous expressions, we can obtain income per capita to be

$$y_t = h_t \pi_{u,t} + (1 - h_t) w_{ut},$$

and the degree of urban concentration of less educated workers

$$u_t = \frac{h_t \ell_t}{1 - h_t}, \tag{4.5}$$

and measures the proportion of less educated workers in the urban sector.

The levels of total factor productivity in the urban and rural sectors are endogenous. Following Lucas (1988), Azariadis and Drazen (1990), or Benhabib and Spiegel (2005), we assume that they are determined by the proportion of highly educated workers in the country, i.e.

$$\begin{aligned} A_{ut} &= \gamma^t A_u (h_t), \\ A_{ft} &= \gamma^t A_f (h_t). \end{aligned}$$

In line with the stylized facts above, the derivative of A_f with respect to h_t is assumed to be larger than the derivative of A_u. In other words, the productivity gap between urban and rural sectors decreases with the level of development.

From (4.4), this implies

$$\ell_t \equiv \ell(h_t) = \left(\frac{\alpha A_u(h_t)}{A_f(h_t)} \right)^{1/(1-\alpha)} \tag{4.6}$$

with $\partial \ell_t / \partial h_t < 0$. It follows that income inequality between high-skilled and less educated workers $(\pi_{u,t}/w_{u,t})$ decreases with the level of development.

Plugging (4.1) into (4.5), we can rewrite the dynamics of the economy as follows:

$$h_{t+1} = a_t \times h_t^{1-\beta} \times H_t^{\beta} \times \phi \left[\frac{h_t \ell(h_t)}{1 - h_t} \right], \tag{4.7}$$

where a_t denotes the efficiency of the education system. Hence, along the transition path, income per capita is given by

$$y_t = h_t A_{ut} \left[(1-\alpha) \left[\ell(h_t) \right]^{\alpha} - q \frac{h_t - h_{t-1}}{h_t} \right] + (1 - h_t) \alpha A_{ut} \left[\ell(h_t) \right]^{\alpha - 1}.$$

It clearly appears that a rise in h_t increases the levels of total factor productivity A_{ut} and A_{ft}; increases the number of entrepreneurs in the urban sector; but reduces the number of employees per entrepreneur ℓ_t and induces congestion costs c_t.

4.4 Numerical analysis

In this section, we confront data to theory to calibrate the parameters of equations (4.1) and (4.6) using panel data. Then, other parameters are calibrated so as to match certain data moments. Finally, the dynamics of human capital accumulation, urbanization and GDP per capita for high-income countries and developing regions are depicted.

First, we look at the impact of urbanization on human capital accumulation. With this aim, we estimate an empirical convergence model in line with a logarithmic transformation of equation (4.1). In particular, we use estimate

$$\ln\left(\frac{h_{i,t+1}}{h_{i,t}}\right) = a_0 + \beta \ln\left(\frac{H_t}{h_{i,t}}\right) + \delta u_{i,t} + a_i + a_t + \varepsilon^h_{i,t}, \qquad (4.8)$$

where H_t stands for the proportion of highly skilled workers in the leader economy (the U.S. in our case), $h_{i,t}$ is the percentage of highly skilled workers in the resident population of country i at time t, $u_{i,t}$ is the proportion of less educated workers living in cities, as defined in (4.5), a_0 is the general intercept, a_i is the country fixed effects, a_t is time fixed effects which capture common time-dependent shocks, β is a parameter that captures the speed of convergence to the level in long-run equilibrium (the higher it is, the faster human capital level of the country i converges to the human capital level of the U.S.), and δ measures the effect of urbanization on human capital accumulation. Note that we assume the following functional form: $\phi(u_{i,t}) = \exp(\delta u_{i,t})$.

We use data from Docquier, Lowell and Marfouk (2007) (henceforth referred to as DLM (2007)) for the highly skilled workers H_t and $h_{i,t}$. The proportion of high skilled workers corresponds to people with tertiary education and 25 percent of the total secondary educated population as a percentage of total population. DLM (2007) construct human capital indicators from De La Fuente and Domenech (2002) for OECD countries and from Barro and Lee (2001) for non-OECD countries. In addition, for countries where Barro and Lee measures are missing, they predict the proportion of educated using Cohen-Soto's measures (see Cohen and Soto, 2007). Urbanization $u_{i,t}$ is defined as a function of $h_{i,t}$ and $\ell_{i,t}$. The data for the low-skilled labor in urban areas is calculated as the difference of high-skilled population from the urban population and the data on urban population is again obtained from WDI (2008). Urban population is defined by WDI (2008) as "the midyear population of areas defined as urban in each country and reported to the United Nations". Data is available for each five-year time period between 1975 to 2000. Number of countries in the data is 136. (Total number of countries in DLM (2007) is 195. Countries in conflict (19), with insufficient statistics

(16) and newly created countries (24) are excluded from the DLM (2007) dataset.)

Table 4.1: Beta-convergence model without urbanization

	OLS1	OLS2	IV1	IV2	D GMM	S GMM
$\ln(H/h)$	.080***	.504***	.065***	.512**	.391***	.225***
Constant	.050**	-.557***	.061***	-.409***	-.183**	
Country FE	no	yes	no	yes	no	no
Year FE	no	yes	no	yes	yes	yes
Nb obs.	679	679	543	543	543	679
Nb cties	136	136	136	136	136	136
R^2/Wald st$^+$	.137	.486	.11	.57	429.74$^+$	443.12$^+$

Notes: * p<0.05; ** p<0.01; *** p<0.001. Wald chi2 statistic is given for GMM estimations. Instrument in IV method: a lagged value of the deviation from the frontier is used as an external instrument. Results of the first stage regressions confirm the validity and relevance of the instrument. Diff GMM and System GMM: two step GMM and Windmeijer finite-sample correction is done for the calculation of standard errors. Hansen-J statistic does not support the validity of the instruments although there is no specification problem according to Arellano-Bond test statistic.

To supplement the stylized facts depicted on Figure 4.3 and check whether there is absolute or conditional convergence in human capital, the model without urbanization effects is estimated at the first step. Table 4.1 presents the results. In the model without fixed effects, we obtain a significant convergence rate of 8 percent a year to a common steady state (column 1). This is in line with Figure 4.3(a). However, this simple regression suffers from two important problems: *i*) unobserved heterogeneity, and *ii*) a Nickell bias (Nickell, 1981) due to the presence of $\ln(h_{i,t})$ on both sides of equation (4.8).

We can solve the first problem by adding country and time fixed effects. Islam (1995) argues that the importance of capturing unobserved individual effects in studying the convergence (since they are positively correlated with the initial level of human capital) and ignoring them will result to a biased convergence coefficient β. In the second colum of Table 4.1, we show that the convergence speed increases from 8 to 50 percent; country fixed effects are highly significant. This means that the model generates conditional convergence: each country converges to a specific steady state, increasing with the level of human capital of the leader.

To solve the Nickell bias, we use IV and GMM techniques. In IV estimations, we use the lagged value of the suspected endogenous variable as

Table 4.2: Beta-convergence model with urbanization

	IV1	IV2	IV3
Urbanization	.136***	.591**	627**
$\ln(H/h)$	.087***	.534***	.565***
Constant	-.037	-.594***	-.638***
Country FE	no	yes	yes
Year FE	no	yes	yes
Nb obs.	543	543	543
Nb countries	136	136	136
R^2	.552	.837	.838

Notes: * p<0.05; ** p<0.01; *** p<0.001. Instruments in IV methods: the lagged value of the deviation from the frontier and the lagged value of urbanization variable are used as an external instrument for IV_FE (1). In IV_FE (3), only urbanization variable is instrumented again with its lagged value. Results of all the first stage regressions confirm the validity and relevance of the instrument. First-stage F statistics are well above the rule of thumb value of 10. Under-identification null hypothesis is rejected at the 1% level. Hansen-J statistic does support the validity of the instruments.

instrument. The choice of instruments, however, is different in GMM estimations and depends on which GMM technique we prefer. In difference-GMM (Arellano and Bond, 1991) we instrument the differences (or orthogonal deviations) with the levels of the suspected endogenous variables. In system-GMM (Blundell and Bond, 1998), the level of suspected endogenous variable is instrumented with their differenced values (current value minus the lagged value of the same variable). Combining these techniques with fixed effects does not change the conclusions.

In the second step, we introduce the urbanization effect. Among the several specifications with a linear, logarithmic or exponential functional forms we have tried in the calibration, we have found that the linear specification defined in (4.5) provides the best fit.

This is similar to Lucas (2009) who introduced urbanization externalities and a convergence effect whose strength depends on the width of the gap between the level of human capital in the U.S. and in the country under study. Table 4.2 provides the results. The columns IV1, IV2 and IV3 show that urbanization has a positive and significant impact on human capital accumulation. Again, the model predicts conditional convergence. As urbanization is likely to be an endogenous process, we prefer the IV specification based on internal instrumentation. In order to determine the parameter values for the simulation exercises we preferred to use the parameter values from IV2 since that specification performs internal instrumentation for both the urbanization and the deviation of human capital from the leader country.

To explain urban employment in (4.6), we need to specify and estimate the TFP function in the urban and rural sectors. We assume the following functional forms: $A_{ut} = A_{u0} h_t^z$ and $A_{ft} = A_{f0} h_t^v$. From (4.6), this gives $\ell_t^{1-\alpha} = \frac{\alpha A_{u0}}{A_{f0}} (h_t)^{z-v}$, or equivalently

$$\ln \ell_t = \ln \rho_0 + \epsilon \ln h_t + \ln \rho_i + \ln \rho_t + \varepsilon_{i,t}^l, \qquad (4.9)$$

where $\rho_0 = \left(\frac{\alpha A_{u0}}{A_{f0}}\right)^{1/(1-\alpha)}$, $\epsilon = (z-v)/(1-\alpha)$, ρ_i is the country fixed effects, and ρ_t stands for time fixed effects.

To estimate (4.9), we use data on the proportion h_t of highly educated workers and the number ℓ_t of less educated in cities for each country in our dataset from 1975 to 2000. Data sources for h_t and l_t are the ones defined in the previous section.

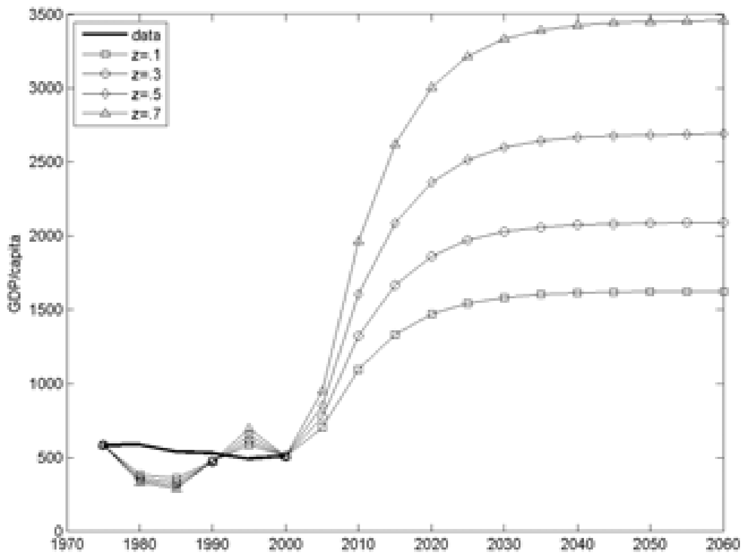
Table 4.3: Modeling low-skilled workers education choices. Dependent = $\ln l_t$

	OLS (2000 data)	OLS (1975-2000 data)
ϵ	-.747***	-.969***
$\ln \rho_0$	-.645***	-1.091***
Country fixed effects	no	yes
Time fixed effects	no	yes
Nb observations	136	815
R^2	.491	.964
$\ln \rho$ - High-income countries[a]	-.512	-.792
$\ln \rho$ - LAC[a]	-.542	-.953
$\ln \rho$ - MENA[a]	-.208	-.645
$\ln \rho$ - South Asia[a]	-1.214	-1.744
$\ln \rho$ - Sub-Saharan Africa[a]	-.606	-1.248

Notes: * p<0.05; ** p<0.01; *** p<0.001. [a]Regional intercepts are obtained by substracting $\epsilon.\ln(h2000)$ from $\ln l_t$.

Table 4.3 shows the results. Firstly, the result of a cross-country regression on 2000 data is presented; then the panel estimate results with fixed effects are provided. The results show that there is a significant convergence rate of 9 percent without fixed effects which rises to above 50 percent when we control for the country and year fixed effects. Thus, we can argue that there is conditional convergence even after controlling for the impact of urbanization, which is significantly positive.

Figure 4.6: Effects of parameter z



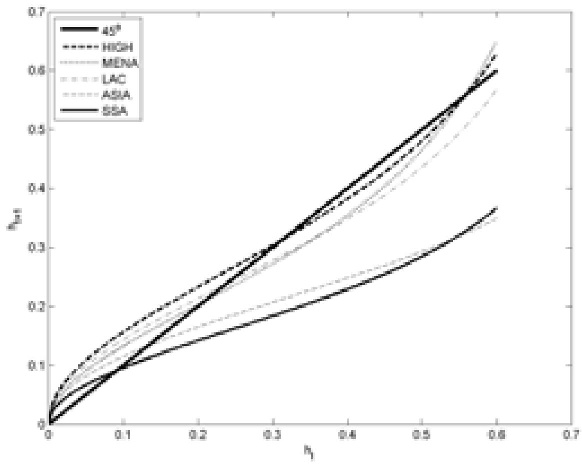
The remainder of parameters used in the numerical simulation do not come from panel data estimation techniques. From Hendricks (2004) we obtain that the income ratio between high- and low-skilled workers is of 2.13 in the US in 2000 (assuming that high-skilled workers have 10 years of education more than low-skilled workers). This fact combined with equation (4.3) and $l_{2000}^{USA} = 0.42$ implies that α is around 0.2.[2]

Parameter z is set equal to 0.3. Note that the magnitude of predictions for GDP per capita are sensitive to this parameter. Contrary to the rest of parameters we fix its value. Hence, predicted values should be taken as suggestive. Figure 4.6 shows the effects of different possible values of parameter z on the predictions for sub-Saharan Africa. This parameter specification does not affect human capital nor urban population simulations. If the effect of human capital is higher, i.e., z is higher than in the benchmark case, we would expect higher long-run values in the long run.

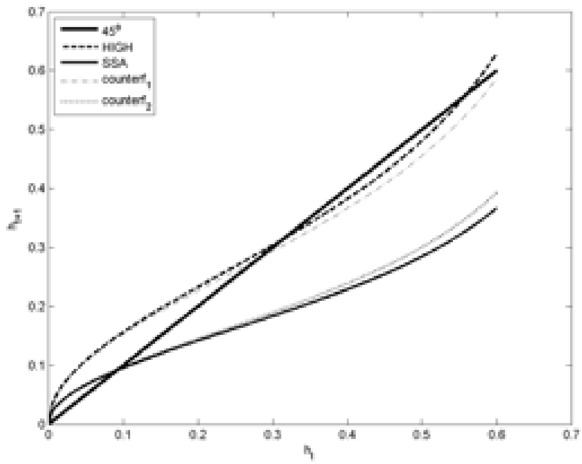
Finally, for every country or region i, we set the scale parameter A_{u0}^{i} in the urban TFP function so as to match the level of GDP per capita observed in 1975 and the adjustment scale parameter q^{i} so as to match the level of GDP per capita observed in 2000.

[2]Hendricks (2004) provides the most recent data on Mincerian returns. We use the most recent year in the data for the USA (1995) as a proxy for the Mincerian return in 2000.

Figure 4.7: Dynamics of human capital



(a) Dynamics per region



(b) Counterfactual dynamics for Africa

Note: On Figure 4.7(a), we plot the proportion of high-skilled workers in t+5 as a function of the proportion in t, using the fixed effects obtained for different regions. An intersection with the 45 degree line is a long-run steady state. On Figure 4.7(b), we focus on sub-Saharan Africa and compare the dynamics obtained with regional fixed effects (SSA), with high-income fixed effects for the human capital equation ($counterf_1$), and with the high-income fixed effect for the urban-to-rural productivity ratio ($counterf_2$). We also represent the high-income dynamics for a matter of comparison.

4.4.1 Evolution of human capital

Once all parameters are calibrated, they can be included in the dynamic equation (4.7). Figure 4.7(a) depicts the dynamics of human capital in each developing region and in high-income countries. Regional fixed effects are weighted averages of country fixed effects in each region. Given those region-specific parameters, the steady state differs across regions. The long-run human capital stock equals 0.32 in high-income countries, 0.25 in LAC, 0.22 in the MENA, 0.13 in South Asia and .07 in sub-Saharan Africa. Those steady states are locally stable. Unbounded growth could be obtained if countries would start with a proportion of college graduates above 60 percent. When we compare these simulation results with the observed human capital levels in each region, it seems that LAC and MENA have the highest distance to their steady-state levels compared to other regions. Moreover, as it will be discussed more in detail in Section 4.4, SSA needs about 20 years to reach its steady-state human capital level. Furthermore, it is predicted that there is no room for the absolute convergence among regions since the efficiency of the education system in each region is different which is an important factor of the human capital dynamics. This is exemplified in Figure 4.7(b) for SSA.

Figure 4.7(b) is a counterfactual exercise in which we substitute the sub-Saharan fixed effect by those observed in high-income countries. The simulation $counterf_2$ shows that substituting the value for ρ_i in equation (4.9) does not modify the steady state that much. On the contrary, substituting the value for a_i in equation (4.8) has a major impact on human capital accumulation. The counterfactual steady state in simulation $counterf_1$ would be almost identical to that obtained in high-income countries. This shows that the technological function of human capital formation plays a key role in the determination of long-run performances of developing countries.

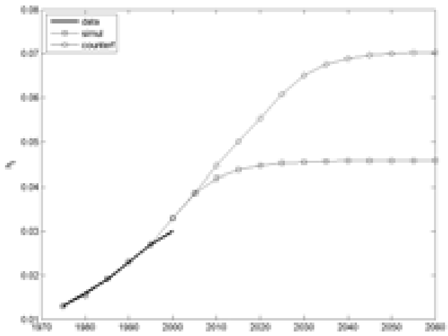
4.4.2 Sub-Saharan Africa

In this section, the results of the dynamic simulations for sub-Saharan Africa are discussed and interpreted in detail.We use the observed levels of human capital in 1975 are used as the starting point. Then, the model is employed for the simulation from 1975 to 2060 of human capital, urbanization and GDP per capita.

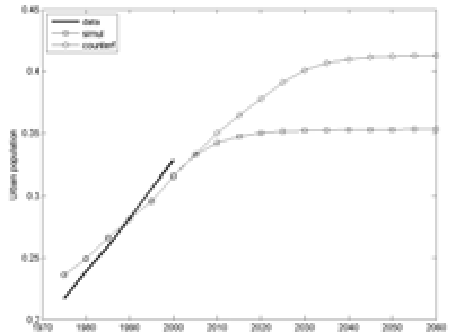
Figure 4.8 presents the simulated transition paths of human capital (Fig 4.8a), urbanization (Fig 4.8b) and GDP per capita (Fig 4.8c) for sub-Saharan Africa. The bold lines represent observations from 1975 to 2000. The lines with squares depict a simulation with constant time fixed effect after 2000, and the lines with circles assume that the time fixed effect will grow up till

2030, i.e., in case that we consider an expansionay education policy until 2030. Before 2000 we use the time fixed effects obtained from the estimation. We first notice that our calibration strategy does an excellent job in matching the data. Figure 4.9 shows that the model also provides an excellent fit for the other developing regions. Note that urbanization is the less precise because we use the cross-country estimation for the concentration of low skilled workers in the urban sector in year 2000.[3]

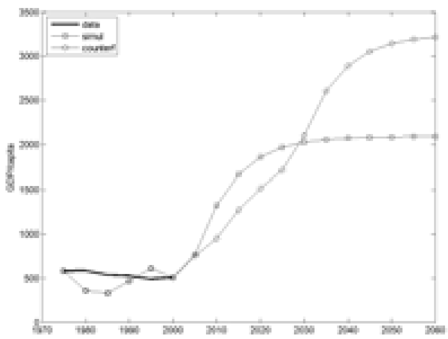
Figure 4.8: Latent growth in Sub-saharan Africa



(a) Trajectory of human capital



(b) Trajectory of urban population



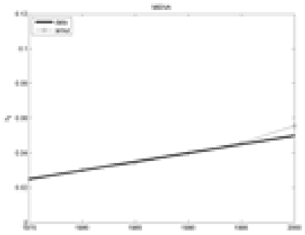
(c) Trajectory of GDP per capita

Notes: Observed (bold line) and simulated (line with squares) trajectories of the proportion of college graduates (6a), share of urban population (6b) and GDP per capita (6c). The line with circles gives the simulation with a continued expansionary education policy.

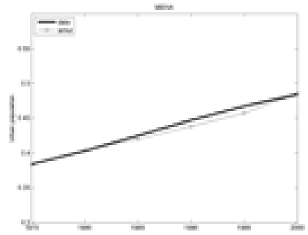
[3]We use this parametrization because long-run outcomes are not modified by changes in parameter ρ, as shown on Figure 4.7(b), and it enables to decrease the number of parameters used in the simulation exercise.

In Figures 4.8(a) and 4.8(b), it can be seen that a virtuous circle of urbanization and human capital accumulation will be such that the proportion of college graduates will reach 7 percent in the long-run, and the share of urban population will reach 34 percent in the pessimistic scenario. Basically, those long-run values will be attained around 2020. As education and urbanization growth rates decline, the economic takeoff takes place and GDP per capita will be multiplied by 4 in the long-run, as shown on Figure 4.8(c). It takes about 30 years for GDP per capita to reach its long-run value. As in Lucas (2009), the growth in human capital leads to higher income in the long-run, but induces temporary costs in the medium term.

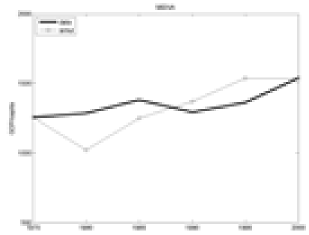
Figure 4.9: Model validation on other developing regions



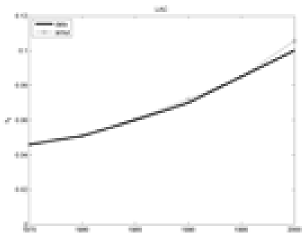
(a) Human capital - MENA.



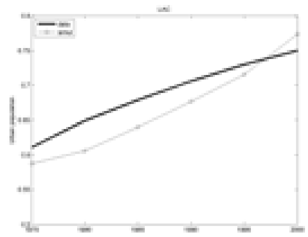
(b) Urban population - MENA.



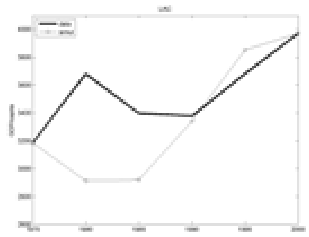
(c) GDP per capita - MENA.



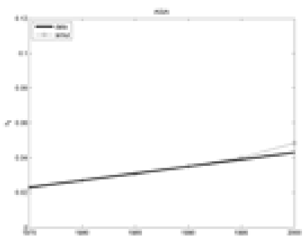
(d) Human capital - LAC.



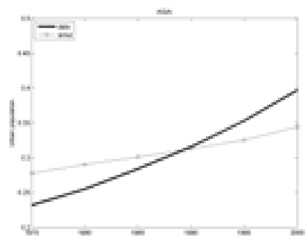
(e) Urban population - LAC.



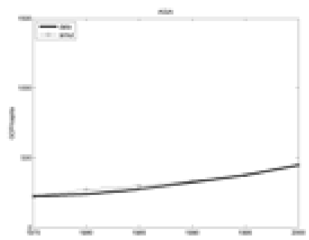
(f) GDP per capita - LAC.



(g) Human capital - ASIA.



(h) Urban population - ASIA.



(i) GDP per capita - ASIA.

Note: Observed (bold line) and simulated (line with circles).

In the more optimistic scenario, which can be depicted by the lines with circles, the trend in human capital keeps on increasing until 2030, as well as the urbanization rate. The takeoff of Africa is then delayed by a few decades or so, but the long-run effect will be stronger since GDP per capita will be multiplied by 6. In that scenario, it is also interesting to note that the increase in human capital is higher than the increase in urbanization. Thus, it can be argued that temporary urbanization costs are to some extent overwhelmed by the positive externality of urbanization on human capital, which causes a higher GDP per capita growth in the long-run. Therefore, one can conclude that educational investments in Africa cannot be seen as a deadweight loss but rather continued progress in education should be promoted in the region to have higher latent growth potential in the future.

4.5 Conclusion

This paper analyzes the impact of urbanization on economic development and growth through its impact on human capital accumulation. We build a theoretical model of endogenous human capital accumulation and calibrate it to account for the effects of urbanization on human capital dynamics.

The theoretical model includes urbanization in human capital dynamics, in such a way that urbanization externalities are taken into account in the convergence process. Besides, negative externalities of urbanization, namely congestion costs, are also considered in the generation process of economic growth. In the model, a rise in human capital increases the level of productivity but also induces congestion costs that may delay the positive effects of the increase in human capital. The positive externality associated to the higher human capital accumulation may be considered as a latent growth potential rather than a deadweight loss. In other words, countries that have experienced a rise in human capital but have not experienced as much growth as expected, according to our theory, will grow when urbanization and human capital accumulation become adjusted. Moreover, the theoretical analysis reveals a negative relation between income inequality (defined as the income ratio between high- and low-skilled workers) and the level of development. Note that our model does not allow for population size changes, we consider a constant labor force.

The calibration exercise for the human capital dynamics per region indicates that there exists conditional convergence between regions: high-income countries, sub-Saharan Africa, Middle East and Northern Africa, Latin America and the Caribbean, and South Asia. Hence, there is no absolute converge, each regions converges to a different steady state and poor countries do not achieve the income levels of rich countries. Moreover, it is found that the technological function of human capital formation plays a key role in the determination of long-run performances of developing countries. Furthermore, the counterfactual analysis provides an optimistic view on the growth tragedy of Africa and shows that the reason behind the fact that educational investments in sub-Saharan Africa do not generate an economic growth is the temporary adjustment costs due to urbanization. We perform an exercise that considers the case of absence of these urbanization costs, and the rise in educational investments would trigger a direct increase in GDP. Our calibrated model predicts that GDP per capita in SSA could be multiplied by 4 within about 30 years. The region would reach the income level of current middle income countries. In another scenario where progress in schooling keeps on until 2030, the African takeoff will be delayed but the long-run GDP per capita will be multiplied by 6 in comparison with the current level. Hence, in the calibration exercise reinforces the idea of a latent growth potential in Africa.

References

Arellano, M., and S. Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The Review of Economic Studies*, 58(2), pp. 277-297.

Azariadis, C., and A. Drazen. 1990. "Threshold Externalities in Economic Development." *Quarterly Journal of Economics*, 105(2), pp. 501-526.

Barro, R.J., and J.W. Lee. 2001. "International data on educational attainment: Updates and implications." *Oxford Economic Papers*, 53, pp. 541-563.

Benhabib, J., and M.M. Spiegel. 2005. "Human Capital and Technology Diffusion". *Handbook of Economic Growth* (eds. P. Aghion and S. Durlauf), 1(1), ch. 13, pp. 935-966.

Bertinelli L. 2003. "Does urbanization always foster human capital accumulation?" CORE discussion papers 2003040.

Bils, M., and P. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review*, 90(5), pp. 1160-1183.

Blundell, R., and S. Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, 87, pp. 115-143.

Black D., and V. Henderson. 1999. “A theory of urban growth.” *Journal of Political Economy*, 107, pp. 252-284

Caselli, F. 2005. “Accounting for Cross-Country Income Differences.” *Handbook of Economic Growth* (eds. P. Aghion and S. Durlauf), 1(1), ch. 9, pp. 679-741.

Cipolla, C.M. 1969. *Literacy and Development in the West*. Pelican: Baltimore.

Cohen, D., and M. Soto. 2007. “Growth and human capital: Good data, good results.” *Journal of Economic Growth*, 12(1), pp. 51-76.

De la Fuente, A. and R. Domenech. 2006. “Human capital in growth regressions: How much difference does data quality make? An update and further results.” *Journal of the European Economic Association*, 4(1), pp. 1-36.

Docquier, F., B.L. Lowell, and A. Marfouk. 2007. “A gendered assessment of highly skilled emigration.” *Population and Development Review*, 35(2), pp. 297-321.

Easterly, W., and R. Levine. 1997. “Africa’s growth tragedy: Policies and ethnic divisions.” *Quarterly Journal of Economics*, 112(4), pp. 1203-1250.

Hall, R.E., and Ch. I. Jones. 1999. “Why Do Some Countries Produce So Much More Output Per Worker Than Others?” *The Quarterly Journal of Economics*, 114(1), pp. 83-116.

Heston, A., R. Summers, and B. Aten. 2006. Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Islam, N. 1995. “Growth Empirics: A Panel Data Approach.” *Quarterly Journal of Economics*, 110(4), pp. 1127-70.

Jones, B.F. 2009. “The Knowledge Trap: Human Capital and Development Reconsidered.” NBER Working Papers n. 14138.

King, L. 1995. “Quantitative Theory and Econometrics.” *Federal Reserve Bank of Richmond Economic Quarterly*, 81(3), 53–105.

Klenow, P., and A. Rodriguez-Clare. 1997. “The neoclassical revival in growth economics: Has it gone too far?” *NBER Macroeconomics Annual*(eds. Ben S. Bernanke and Julio J. Rotemberg). Cambridge, MA: MIT Press, 73–103.

Lucas R. 1988. "On the mechanics of economic development." *Journal of Monetary Economics*, 22, pp. 3-42.

Lucas R. 2009. "Trade and the diffusion of the industrial revolution." *American Economic Journal: Macroeconomics"*, 1(1), pp. 1-25.

Maddison, A. 1995. *Monitoring the world economy 1820-1992*, Paris: OECD.

Manuelli, R.E., and A. Seshadri. 2007. "Human capital and the Wealth of Nations." University of Wisconsin-Madison, mimeo.

Morrisson, C., and F. Murtin. 2009. "The Century of Education", Centre for the Economics of Education Discussion papers, London School of Economics and Political Science, UK.

Nickell, S. 1981. "Biases in dynamic models with fixed effects." *Econometrica*, 49(6), pp. 1417-1426.

Parente, S.L., and E.C. Prescott. 2000. *Barriers to Riches*. Cambridge, MA: MIT Press.

Pritchett, L.A. 2001. "Where has all education gone?" *World Bank Economic Review*, 15(3), pp. 367-391.

Vandenbussche, J., Ph. Aghion, and C. Meghir. 2006. "Growth, distance to frontier and composition of human capital." *Journal of Economic Growth*, 11(2), pp. 97-127.

United Nations. 2008. *The millennium development goals report 2008*. United Nations, New York.