Brain Drain, Inequality and Growth*

Fabio Mariani†

October 19, 2004

Abstract

This paper provides an additional channel through which inequality may influence growth, when labor migration is taken into account. In fact, we show that human capital distribution is crucial to determine whether allowing migration of the most skilled workers from a developing country may be beneficial for growth, from the perspective of the source economy. The net effect linked to a brain drain is more likely to be negative in the short run if human capital is very unequally distributed. In addition, we find that econometric analysis supports our theoretical claims: the estimation of different growth equations in a cross-section of developing countries, based on a brand new dataset on skilled migration (Docquier and Marfouk, 2004) shows that a brain drain can have a positive impact only when it is associated with low inequality (in income or schooling).

JEL classification: F22; J24; J61; I20.

Keywords: High-skilled migration; Inequality; Education.

*I wish to thank François Bourguignon and David de la Croix for their precious guidance. Comments from Frederic Docquier on an earlier draft were extremely important to improve the paper. I would also like to thank Yannis Vailakis as well as other seminar participants at UCL, Louvain-la-Neuve and UPF, Barcelona for lively and useful discussion. All remaining errors are, of course, my own responsibility. Financial support by the French Speaking Community of Belgium in the framework of the ARC Project "New Macroeconomic Perspectives on Development" is gratefully acknowledged.

†IRES, Université catholique de Louvain - Place Montesquieu 3, B-1348 Louvain-la-Neuve (Belgium). E-mail address: mariani@ires.ucl.ac.be .

1 Introduction

By "brain drain" we mean the migration of (part of) the most skilled workers in a population, from less developed countries to more developed and richer ones[1].

The relevance of high-skilled migration in the real world is undoubtable and growing (for an empirical assessment, see Carrington and Detragiache, 1998, as well as Docquier and Marfouk, 2004); to some extent, it concerns not only LDC's, but also industrialized countries.

Since Bhagwati and Hamada (1974), development economists have worried about the possible effects that this "human capital flight" could exert on welfare and growth in the source economy (and on convergence/divergence with respect to the destination country). The recognition of human capital formation as a crucial engine of growth, in a consistent strand of literature inspired by Lucas (1988), has even added new interest and life to the debate.

All along the 90's, there has been a flourishing of endogenous growth models all assessing the negative effects of brain drain on growth: in particular we refer to Miyagiwa (1991), Haque and Kim (1995), Reichlin and Rustichini (1999) and Wong and Yip (1999). All these contributions share the view that the migration of the most skilled individuals, through a decrease in the average human capital (and the consequently negative externality effects), would be bad for growth and imply diverging growth trajectories between rich and poor countries.

More recently, some models have been published arguing that a brain drain might even be good for growth and welfare in the developing economy. In fact Mountford (1997), Beine et al. (2001), Stark and Wang (2002) put forward the idea that a positive chance of migration may foster human capital accumulation in the source country, since it entrains an incentive effect linked to the higher wages available abroad. In this framework, people may choose

[1]This one is a "broad" definition. A "narrow" definition of the brain drain would refer to the migration of scientists, engineers, faculty members ..., rather than to the outflow of educated people in general.

to educate more in order to become eligible for migration, and then to have access to the higher returns of the foreign labor market. The gains in total human capital formation produced by this incentive effects may outweigh, in the end, the human capital loss represented by the actual brain drain. However, even in this context, the "unpleasant" result of divergence holds.

In the present paper we consider this incentive argument, and explore its consequences in term of growth, *when inequality is concerned*. To be more precise, we want to see if inequality affects the interplay between skilled migration and growth. We try to accomplish this task, by adopting an OLG setup in the fashion of Azariadis and Drazen's (1990) analysis of the trade-off between studying and working.

By doing this, we will be able to show that human capital distribution is crucial to determine whether the brain *gain* would be strong enough to prevail over the brain *drain*[2] (at least in early periods). We will also show how allowing migration modifies human capital distribution in the long run, and that in the long run even the (possible) net brain gain experienced at the beginning, is outweighed by the persisting brain drain.

Our model contributes to the literature on high-skilled migration and growth, filling the interesting space which lies between Beine et al. (2001) and Mountford (1997). In fact, the first paper assumes inequality (of innate learning abilities) as renewing itself randomly in every period, and in the end does not care too much about the dynamics of human capital distribution, while we think that in developing economies the inequality of chances (and its persistence over time) matters a lot. On the other side Mountford (1997), who concentrates on the "long run" of income distribution in its analysis of the relation between brain drain and growth, cares more about the consequences *on* inequality, than about the consequences *of* inequality. In addition, our analysis offers a contribution to the literature on the relationship between inequality and growth. In fact, we retrieve an additional channel through which inequality (of abilities, income, education ...) may af-

[2]We borrow these terms of "brain drain" vs "brain gain" from Stark et al. (1997).

fect economic growth. Of course this channel works, in developing economies, only when labor is (at least partially) internationally mobile. And it has to be underlined that, similarly to other papers (see for instance de la Croix and Doepke, 2003), in our model the growth-inequality link runs through human capital and education.

Moreover, we are able to provide some empirical evidence in support of our theoretical claims. To this scope we exploit both the well-known brain drain data provided by Carrington and Detragiache (1998) for selected developing countries, and a brand-new data set built by Docquier and Marfouk (2004) that revise the previous one and adds industrialized countries to the sample. Estimating growth equations (enhanced with terms that account for inequality and high-skilled migration) in a cross-section of developing countries, we find that a brain drain can positively affect income growth only if schooling and/or income are not too unequally distributed across classes.

The paper is thus organized as follows. Section 2 presents the simple model upon which our analysis is built. Migration is explicitly introduced in Section 3, which then analyzes consequences in terms of growth and inequality. Section 4 is devoted to the presentation of our econometric findings. Finally, Section 5 provides a short concluding discussion.

2 The model

Our model is inspired, as pointed out before, by Azariadis and Drazen (1990), who develop a framework of analysis to study the trade-off between studying and working.

The source economy (developing country) is populated by overlapping generations of utility maximizing individuals, who live for two periods and are heterogeneous only with respect to their parental human capital, which is distributed according to the density function $f(h_t)$ over the interval $(\underline{h}, \overline{h})$.

In the first period of their life, agents can devote a fraction τ_t of their time to education, building up human capital for the next period; in the remainder of the time $(1 - \tau_t)$, they can earn a wage w_t from their part of inherited human capital.

We assume the "inheritance" function to be:

$$j(h_t) = h_t^\delta \tag{1}$$

while the production function of human capital through education (schooling) writes as:

$$h_{t+1} = a\tau_t^\sigma h_t^\gamma \tag{2}$$

where σ, γ and δ should all belong to the open interval $(0, 1)$.

Agents maximize life-time income evaluated at time t, i.e.:

$$\Omega_t(h_t, \tau_t) = (1 - \tau_t)h_t^\delta w_t + a\tau_t^\sigma h_t^\gamma \frac{w_{t+1}}{R_{t+1}} \tag{3}$$

where w_{t+1} is the wage at time $t + 1$, and $R_{t+1} = 1 + r_{t+1}$ accounts for the interest rate[3].

It's worth noting that, by writing (3), we implicitly assume that the "inherited" part of human capital (the one that can provide wages in the first period) does not generate income in the second period; we may think to this part of human capital as "physical strength" or "manual skills", which are likely to decay more rapidly than "intellectual skills" (acquired through schooling and education in general). In other words, schooling not only enhances human capital, but makes it productive for a longer time.

From:

$$\frac{\partial \Omega_t}{\partial \tau_t} = -h_t^\delta w_t + a\sigma\tau_t^{\sigma-1} h_t^\gamma \frac{w_{t+1}}{R_{t+1}} = 0, \tag{4}$$

we can get the optimal choice for education:

$$\tau_t^* = \left(a\sigma \frac{w_{t+1}}{w_t} \frac{1}{R_{t+1}}\right)^{\frac{1}{1-\sigma}} h_t^{\frac{\gamma-\delta}{1-\sigma}}. \tag{5}$$

It can be easily seen that $\frac{\partial \tau_t^*}{\partial h_t} > 0$, if $\gamma > \delta$.

[3] We may even assume that $w_{t+1} = w_t = w$, in such a way that no wage dynamics have to be modelled; alternatively one may introduce exogenous dynamics.

This means that the demanded (or better: "desired") amount of schooling τ_t^* will be increasing in h_t, if (and only if) human capital "matters more" in the educational process than it does in the inheritance process. This hypothesis may be justified by claiming that the cultural environment matters more than genetics, in the transmission of skills and knowledge. In fact, it does not seem too unrealistic to say that the human capital of the parent has a relevant importance in determining the school performance of the children and their ability to take profit from their studies.

It it easy to check that $\partial^2 \tau_t^* / \partial h_t^2 < 0$, in the case of $\gamma > \delta$. The function $\tau_t^*(h_t)$ is thus upward sloping and concave. Not surprisingly, the desired amount of schooling also turns out to depend positively on the expected wage dynamics w_{t+1}/w_t, and negatively on the discount rate R_{t+1}.

At this stage we introduce an assumption which is useful to simplify the further developments of the model, without causing any significant loss of realism. In fact we assume that, although demanded continuously by the agents, education is offered as a discrete variable (let's say by the national education system). We simply think to different "packages" of schooling years, that in the real world could correspond for example to primary school, secondary school, high school, university, post-graduate studies, and so on.

For ease of representation, we start by considering the discrete supply of education as assuming three distinct discrete values, namely: $\tau_1 < \tau_2 < \tau_3$. To know how educational choices are effectively made (and by whom), we simply need to compare the values that the function $\Omega_t(h_t, \tau_t)$ assumes for $\tau = \tau_1$, for $\tau = \tau_2$, and finally for $\tau = \tau_3$. Then, according to which one of the three functions $\Omega_t(h_t, \tau_1)$, $\Omega_t(h_t, \tau_2)$, or $\Omega_t(h_t, \tau_3)$ attains the highest level, people will chose the underlying value of τ_t.

We claim what follows:

Proposition 1 *Given τ_1, τ_2 and τ_3, such that*
$\sigma^{\frac{1}{1-\sigma}} < \tau_1 < \tau_2 < \tau_3 < \left(\frac{a\sigma\gamma}{\delta}\frac{w_{t+1}}{w_t R_{t+1}}\right)^{\frac{1}{1-\sigma}} \underline{h}^{\frac{\gamma-\delta}{1-\sigma}}$,
then there exist threshold values $h_{1,2}$, $h_{2,3}$ and $h_{1,3}$ such that:

- *for* $\underline{h} < h < h_{1,2} \Rightarrow \Omega(\tau_1) > \Omega(\tau_2) > \Omega(\tau_3)$
- *for* $h_{1,2} < h < h_{1,3} \Rightarrow \Omega(\tau_2) > \Omega(\tau_1) > \Omega(\tau_3)$
- *for* $h_{1,3} < h < h_{2,3} \Rightarrow \Omega(\tau_2) > \Omega(\tau_3) > \Omega(\tau_1)$
- *for* $h_{2,3} < h < \overline{h} \Rightarrow \Omega(\tau_3) > \Omega(\tau_2) > \Omega(\tau_1)$.

(The Proof is given in Appendix A.)

Assuming, as required by the condition stated above, that $\tau_1 > 0$ (so that the case of "no schooling" is excluded, and even the least endowed spend some time in formal education), things are as explained in Fig.1: people with parental human capital such that $\underline{h} < h_i < h_{1,2}$ will have the lowest education; agents characterized by $h_{1,2} < h_i < h_{2,3}$ will educate at the intermediate level τ_2; and the segment with $h_{2,3} < h_i < \overline{h}$ will buy the highest possible amount of schooling τ_3.

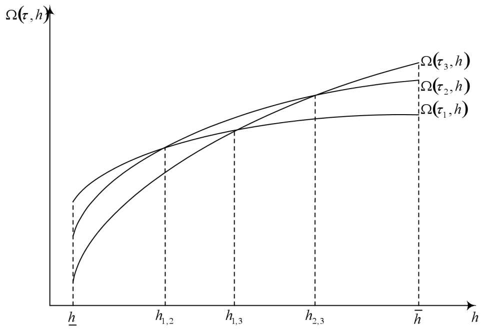

Figure 1: Evaluating life-time incomes: the choice of educational levels

Reintroducing time in our notation, and given these choices, we can represent in Fig. 2 the relationship $h_{t+1} = G(h_t)$ as a discontinuous function, which in fact results from the different values of $\tau^*_{t,i}(h_{t,i})$.

All the eventual crossings between $G(h_t)$ and the 45° line represent different (and all stable) educational steady-states. As time goes to infinity,

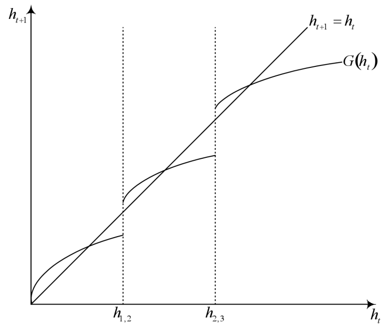

Figure 2: Educational steady states

people tend to converge toward multiple fixed levels of human capital. Different cases may arise. For instance, it may happen that there is only one crossing, at the highest level. In such a case, in the very long run, inequality disappears, and all individuals converge to the same level of human capital, determined by the highest degree of education[4]. On the opposite side, nothing excludes that everyone converges to a lower educational steady-state.

Another qualification is now in order: when considering the function $h_{t+1} = G(h_t)$ we need to introduce a further assumption in order to prevent eventual (future) values of h_{t+i} from falling below $\underline{h}$ (thus redefining, over time, the interval for which Proposition 1 holds). This assumption is simply that:

$$\tau_1 > \left(\frac{\underline{h}^{1-\gamma}}{a} \right)^{\frac{1}{\sigma}};$$

it is derived from $\underline{h} < a\tau_t^{\sigma}\underline{h}^{\gamma}$, which requires $\underline{h}$ to be lower than the lowest (stable) steady-state value of h_t.

We can now summarize the condition for all our dynamic model to be "well behaved" as being simply:

[4]It has to be said that our model allows for social "convergence", but does not encompass the possibility of two-way social mobility.

$$\max\left[\sigma^{\frac{1}{1-\sigma}},\left(\frac{\underline{h}^{1-\gamma}}{a}\right)^{\frac{1}{\sigma}}\right]<\tau_1<\tau_2<\tau_3<\min\left[\left(\frac{a\sigma\gamma}{\delta}\frac{w_{t+1}}{w_t R_{t+1}}\right)^{\frac{1}{1-\sigma}}\underline{h}^{\frac{\gamma-\delta}{1-\sigma}},1\right],$$

that indeed does not turn out to be a heavy restriction[5].

Until now we have described the behavior of the source economy in "autarky". Let's now see how the picture changes once we allow for migration.

3 Migration

3.1 Introducing migration

First of all, let's make clear that in our framework international migration (or better: the will to migrate) is motivated only by the fact that the foreign country can offer, for whatever exogenous reason[6], a unit wage $w^F > w^H$ to the prospective immigrant[7].

We model a brain drain as follows: migration arises as the interaction of the agents' will to migrate with two factors (policies). The immigration policy of the destination country consists in a minimal educational requirement $\widehat{\tau}$, that the prospective immigrant needs to meet if she wants to be accepted. On the other side, the emigration policy fixed by the source country is entirely described by the parameter m, i.e. the fraction of individuals allowed to migrate, among the ones with $\tau_i \geq \widehat{\tau}$. This parameter m thus becomes simply the probability to migrate, from the point of view of "home" workers.

With no educational requirements, we would have

$$\tau_t^M=\left\{\alpha\sigma\frac{[w_{t+1}+m(w_{t+1}^F-w_{t+1})]}{w_t}\frac{1}{R_{t+1}}\right\}^{\frac{1}{1-\sigma}}h_t^{\frac{\gamma-\delta}{1-\sigma}}>\tau_t^* \qquad (6)$$

[5]Numerical examples are available upon request.

[6]It is quite natural to put forward technological reasons; on this subject see for instance Domingues Dos Santos and Postel-Vinay (2001).

[7]The value of w^F may be implicitly discounted (thus, lower than its "real" value) to take into account that workers normally prefer, other things being equal, to live in their birth country.

so that the resulting function $\tau_t^M(h_t)$ is simply an "expansion" of $\tau_t^*(h_t)$.

In this case of "general" migration, which is indeed quite unrealistic, the incentive effect causes both an increase of the enrollment in highest education and a decrease in the number of people who will choose the minimal education.

3.2 Consequences of a brain drain on education choices and human capital distribution in the long run

In the more interesting case of a brain drain (which is modeled setting $\widehat{\tau} = \tau_3$ and represented in Fig. 3) the incentive works only for the highest stage of education: the curve $\Omega(\tau_3)$ is the only one to shift upwards. Opening frontiers moves the threshold value $h_{2,3}$ to the left: $h_{2,3}^M < h_{2,3}^A$; this means that the marginal individual (who is indifferent between τ_2 and $\tau_3 \equiv \widehat{\tau}$) will now be an agent with a lower level of parental human capital. The higher return for superior education ($w_{t+1} + m(w_{t+1}^F - w_{t+1}) > w_{t+1}$) induces more people (let's say from the middle class) to educate at the highest level.

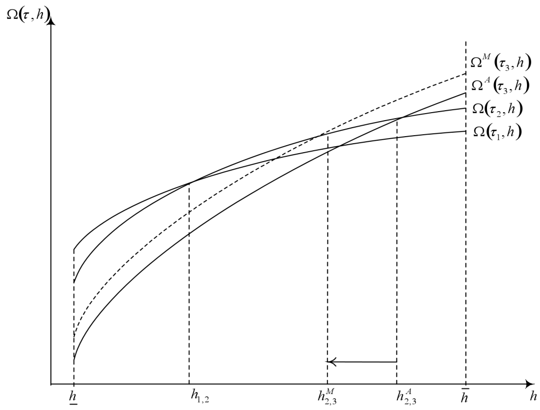

Figure 3: The educational incentive of a brain drain

A first consequence of this incentive effect can be seen in Fig. 4a, that shows how this shift to the left of the marginal individual may make attainable a high educational steady state.

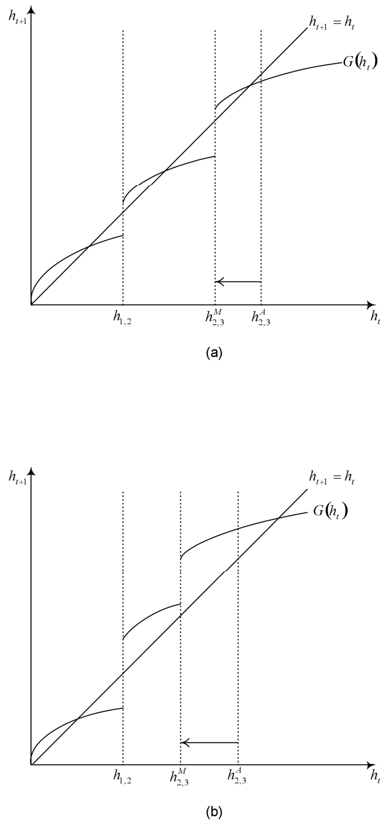

Figure 4: Long run consequences of a brain drain on human capital distribution

Of course, it may also cause an intermediate steady-state to disappear (Fig.4b); if this is the case, the brain drain will turn out to be inequality-worsening, in fact (what we may call) a middle class will disappear, and the higher class, after being initially enlarged by the incentive effect, will progressively shrink by effect of repeated migration outflows.

But what happens in terms of growth?

3.3 Growth effects of a brain drain

Since we assume total output to be given by:

$$Y_t = A_t H_t \tag{7}$$

the growth rate of the economy, keeping A_t constant, will be given by the growth rate of average human capital[8]:

$$g_t = \frac{\widetilde{h}_{t+1}}{\widetilde{h}_t} \tag{8}$$

Assuming that frontiers are opened starting from time $t+1$, and that all the agents correctly anticipate this policy at period t, the possible merits of migration can be simply assessed by evaluating g_t^M (M for "migration"), and eventally comparing it with g_t^A (A stands for "autarky").

The growth rate g_t^M will turn out to depend crucially on $f(h_t)$, the density function of human capital over the interval $(\underline{h}, \overline{h})$, that means on inequality.

In fact, we can write, in general terms:

$$g_t = \frac{\int_{\underline{h}_t}^{\overline{h}_t} \{a[\tau_t(h_t)]^\sigma h_t^\gamma\} f(h_t) dh_t}{\int_{\underline{h}_t}^{\overline{h}_t} h_t f(h_t) dh_t}. \tag{9}$$

Let's now consider our simple "discrete" setting with three educational levels (τ_1 , τ_2 and τ_3). We have:

$$g_t^A = \frac{\int_{\underline{h}_t}^{h_{1,2}} \{a\tau_1^\sigma h_t^\gamma\} f(h_t) dh_t + \int_{h_{1,2}}^{h_{2,3}^A} \{a\tau_2^\sigma h_t^\gamma\} f(h_t) dh_t + \int_{h_{2,3}^A}^{\overline{h}_t} \{a\tau_3^\sigma h_t^\gamma\} f(h_t) dh_t}{\int_{\underline{h}_t}^{\overline{h}_t} h_t f(h_t) dh_t} \tag{10}$$

and

$$g_t^M = \frac{\dfrac{\int_{\underline{h}_t}^{h_{1,2}} \{a\tau_1^\sigma h_t^\gamma\} f(h_t) dh_t + \int_{h_{1,2}}^{h_{2,3}^M} \{a\tau_2^\sigma h_t^\gamma\} f(h_t) dh_t + (1-m) \int_{h_{2,3}^M}^{\overline{h}_t} \{a\tau_3^\sigma h_t^\gamma\} f(h_t) dh_t}{1 - m \cdot \int_{h_{2,3}^M}^{\overline{h}_t} f(h_t)}}{\int_{\underline{h}_t}^{\overline{h}_t} h_t f(h_t) dh_t}. \tag{11}$$

[8]To take into account the fact that young individuals work when they do not study, we could have written something like $g_t = f(\frac{\widetilde{h}_{t+1}}{\widetilde{h}_t})$

When compared with g_t^A, the formulation for g_t^M in (11) displays some key differences. First, it takes into account the fact that the "threshold individual" moves from $h_{2,3}^A$ to the lower level $h_{2,3}^M$, by effect of the stronger incentive linked to the migration opportunity. Second, it considers that, after actual migration, only a fraction $(1-m)$ of the most educated remains in the home country. Third, it averages the post-migration human capital over a mass which falls short from 1 by the quantity $m \cdot \int_{h_{2,3}^M}^{\overline{h}_t} f(h_t)$, which represents the proportion of actual migrants in the *total* population.

Then it is clear how $f(h_t)$ can play a decisive role. In particular, the more people is endowed with a human capital comprised between the two values $h_{2,3}^M$ and $h_{2,3}^A$, the more ample will be the possible gain from migration (provided that the actual number of successful migrants is not too large). It'a also worth noting that, for each $f(h_t)$, there will exist a value of m which maximizes g_t^M[9]. If, for instance, $\int_{h_{2,3}^M}^{h_{2,3}^A} f(h_t)dh_t = 0$ (that means that nobody will be touched by the incentive), the optimal value for m will be zero.

Another specification is in order: while a raise in m will extend highest education to agents with less and less human capital, the m migrants will be selected randomly (in the luckiest case[10]) among the prospective migrants. This implies that, *ceteris paribus*, there is a negative "composition" effect of migration.

The general conclusion we can draw is that allowing a (limited) brain drain may be growth enhancing, at least in a short run perspective, if there is a numerous enough middle class which could be interested in higher education; while it is likely to be harmful if human capital is extremely unevenly distributed[11] so that, at the limit, nobody is motivated to shift to a higher

[9] The determination of the optimal value to choose for m is a central issue in Beine et al. (2001) and in Stark and Wang (2002), but both these papers do not deal with the problem of inequality.

[10] In fact, the successful migrants could be selected as the best m among all the eligible, and not randomly.

[11] All over the paper, we have related the notion of inequality to the existence (and the size) of a middle class. However, we need to underline that, in general terms, keeping fixed

educational level.

These findings allow us to establish a link with the literature on the relationship between inequality and growth[12]. In fact, we have shown that, when (skilled) labor is (at least partially) internationally mobile, an additional channel through which inequality may significantly influence growth is turned on, in developing economies. Since in many LDC's the size of skilled migration is important and growing, and since increasing globalization calls for frontiers to be more and more open, we believe that this channel is not negligible.

It may be worth saying that most of the available literature on inequality and growth identifies the accumulation of *physical* capital as the key factor through which this relationship runs. In our setting inequality affects growth through migration and education. This makes this paper closer to the strand of literature which focuses on the accumulation of *human* capital, when looking for a mechanism through which inequality may affect growth; an example in this direction is represented by de la Croix and Doepke (2003, 2004), who link inequality and growth through human capital accumulation and differential fertility.

3.4 The simple case of a uniform distribution

Here we want to show, by means of a very simple example, how we can evaluate and compare the different effects of a positive migration chance, when differently egalitarian distributions of parental human capital are going to be considered.

In particular, we turn our view to the class of uniform distributions that have the same mean $\widetilde{h}$, but are defined on different intervals. In this case, the width L of the interval of definition fully characterizes the distribution, the density function being $1/L$ and the extremes respectively $\widetilde{h} - L/2$ and $\widetilde{h} + L/2$. Therefore, since $var(L) = L^2/12$, it is clear that the smaller is L

the size of the middle class, inequality may nonetheless vary.

[12]See Benabou (1996) for a comprehensive discussion.

the more equally parental human capital is distributed, according to what we meant in the previous section.

For sake of simplicity and ease of computations, in the remainder of this Section we will consider two educational levels instead of three. The quality and the interpretation of the results we will obtain are unaffected by this change.

3.4.1 Analytical results

Suppose that, at time t, a developing country opens its frontiers, or that it experiences some shift in m (increased mobility of high skilled workers). We are first interested in establishing the sign of $\partial(g_t^M)/\partial L$.

In other words we are trying to understand how the possible gains from allowing (more) migration depend on human capital distribution (that in this example is entirely described, as we were saying, by the parameter L).

We claim what follows:

Proposition 2 *If m is not too low, then $\frac{\partial(g_t^M)}{\partial L} < 0$.*

(The Proof is given in Appendix B.)

The result is clear-cut: if frontiers are opened and skilled migration is likely to occur, then more inequality will mean less growth.

In addition, we can prove that:

Proposition 3 *Under appropriate values of the parameters, there exists a strictly positive $m^* = argmax(g_t^M)$.*

(The Proof is given in Appendix C.)

That means that, if frontiers are opened, there may exist a strictly positive value of the migration rate which maximizes growth. The fact that this m^* can be possibly different from 0 means that we can identify conditions (depending on inequality, see the Appendix) for a brain drain to be beneficial; in other words, we can rule out the possibility that a brain drain is necessarily growth-worsening.

3.4.2 A numerical example

Here we build a numerical example (i.e. we fix the parameter values and solve our model), in order to provide some results that cannot be proved analytically.

In particular we want to show that, once we define:
- $\Delta g \equiv g_t^M - g_t^A$
- m^* as being that value of m which maximizes the possible gains from migration ($m^* = argmax(\Delta g)$), and
- m° as the highest possible value for which opening frontiers do not provoke a loss ($\Delta g(m^\circ) = 0$, with $m^\circ \neq 0$),

then the following results hold:
(i) $\frac{\partial m^*}{\partial L} < 0$, and
(ii) $\frac{\partial m^\circ}{\partial L} < 0$.

The meaning of this double result is easy to explain. A more egalitarian country would be able to tolerate relatively higher values of m, without experiencing any growth losses (result (ii)). In addition, a more egalitarian human capital distribution would push a developing country to optimally choose a relatively larger value of m^*, when m is under its policy control (result (i)).

Before choosing a particular configuration of the parameters, we make explicit a requirement we ask them to meet, namely that there will be always (i.e. with or without migration) someone who chose to educate at the higher level and someone who opt for lower education. This requirement, that is made only for sake of realism and that is not necessary for our results to hold, translates into: $\widetilde{h} - L/2 < h_{1,2}^M < h_{1,2}^A < \widetilde{h} + L/2$.

The latter leads to a condition on $\gamma - \delta$, that must hold for every m:

$$\frac{\log\left\{\frac{w_t R_{t+1}(\tau_2 - \tau 1)}{a[w_{t+1}(\tau_2^\sigma - \tau_1^\sigma) + m(w_{t+1}^F - w_{t+1})\tau_2^\sigma]}\right\}}{\log(\widetilde{h} - \frac{L}{2})} < \gamma - \delta < \frac{\log\left\{\frac{w_t R_{t+1}(\tau_2 - \tau 1)}{a[w_{t+1}(\tau_2^\sigma - \tau_1^\sigma) + m(w_{t+1}^F - w_{t+1})\tau_2^\sigma]}\right\}}{\log(\widetilde{h} + \frac{L}{2})}.$$

Once we take into account the above restriction, we can compose the list

L	7.4	7.6	7.8	8.0	8.2	8.4	8.6	8.8
m^*	0.275	0.251	0.233	0.219	0.207	0.197	0.187	0.179
m°	0.995	0.903	0.825	0.757	0.698	0.647	0.602	0.561

Table 1: Simulation results

of the parameter values as follows: $\tau_1 = 0.45$, $\tau_2 = 1.9 \cdot \tau_1$, $w_t = 1$, $w_{t+1} = 1.7$, $w^F_{t+1} = 2.1$, $R_{t+1} = 1.6$, $a = 1.1$, $\delta = 0.35$, $\sigma = 0.3$, $\gamma = 0.7$ and $\widetilde{h} = 4.4$.

The output of this simulation is reported in Table 1 and relates the computed values of m^* and m° to the different, exogenously fixed, values of L. It can be seen that what was claimed above is reproduced by our simulation. A graphic representation is provided in Figure 5, where f^X is associated with a smaller L (when compared to f^Y).

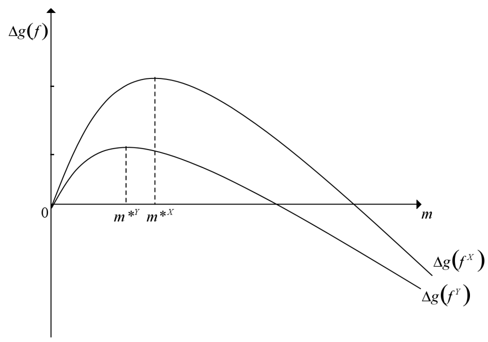

Figure 5: Growth-maximizing brain drain with alternative human capital distributions

Let us also underline that changing the values of the parameters inside the admissible range defined in Sections 2 and 3 would not affect the *quality* of the results[13].

To conclude, it's maybe worth saying that we are able to reproduce the same kind of results if, instead of uniform distributions, we work with exponential distributions belonging to the class described by the density function

[13] Some sensitivity analysis is available upon request.

$f(h) = \lambda e^{-\lambda h}$, where if $\lambda > 0$ increases, inequality decreases[14]. In particular we get that $\frac{\partial m^*}{\partial \lambda} > 0$.

3.5 What about long-run growth?

A peculiar feature of this kind of model is that, keeping m positive but fixed from time $t+1$ on, the "threshold" individual would not move further. However, if we compare this situation with the *status quo* (closed frontiers), we see that the motivation to educate more remains in place, and all the "post-opening" generations will educate more because of this policy change. But in every period, there will be a constant loss of a fraction m of workers from the most educated class, without any other addition to this class (provided that divergence holds, keeping alive the attractiveness of getting a job abroad). At the limit, for $t \to \infty$, there will be strictly no one to hold the highest degree of study in the developing economy. Along time, the net effect of these two forces on the growth rate is ambiguous, as it was in the case of period 1. What is sure is that, going toward infinity and by effect of successive migration waves, the higher educational steady state will be attained by a negligible share of the population, and the whole economy will register an average level of human capital lower than the level it would have reached keeping its frontiers closed. So, if the long-run of such a model is of some interest, we can say that a brain drain will be unambiguously harmful for long-run income, and it will be inequality increasing all along the transition path.

But, as we have already pointed out, the relevance and the interest of the long-run in an OLG model with migration policies remain somewhat questionable. However, it could be interesting to study how a government could play "optimally" with m along time.

To some extent, it would also be attractive to consider m as exogenously growing, as a consequence of the increasing globalization of the world econ-

[14] With this exponential form also the mean depends on λ. More precisely we have that $\mu = \frac{1}{\lambda}$ and $\sigma^2 = (\frac{1}{\lambda})^2$

omy. In this case, our short-run results would continuously replicate over time[15].

It may also be interesting to underline that, in our model, the emigration policy parameter m is not the only policy instrument available for the developing economy. In fact, also educational policies (the configuration of τ's offered by the state to its citizens) matter to the growth perspectives of the country, and can be used in combination with migration policies.

4 Empirical evidence

Here we want to present some evidence that econometric analysis do not reject the predictions suggested by our model. To sum up, the theoretical part of our paper argued that *ceteris paribus* a brain drain can be good for growth if there exists a large enough "middle class" which can benefit from the educational incentive derived from an increased migration probability.

To test this hypothesis we proceed to cross-country estimations of the following equation:

$$ygr = \alpha_0 + \alpha_1 \cdot \frac{I}{Y} + \alpha_2 \cdot \log(y0) + \alpha_3 \cdot dumAFR + \alpha_4 \cdot BD + \alpha_5 \cdot (BD \cdot MID), \quad (12)$$

where ygr is the annual average growth rate of GDP per capita, I/Y is the average ratio of investment to GDP, $\log(y0)$ is the logarithm of initial GDP per capita (this term accounts for convergence effects: we expect α_2 to be negative), $dumAFR$ is a dummy variable for sub-Saharian Africa, BD is the brain drain, and MID is the 'middle class' size variable.

As it will be extensively explained later on, we will take different possible measures of MID. Moreover BD, which is defined as the migration rate of

[15]This scenario of progressive *globalization*, that implies an anticipated and permanent(ly increasing) trajectory for m, can be contrasted with the opposite case of a *revolution* (like the Khomeinist coup in Iran), which involves a non-anticipated and temporary shift of m. In the latter situation the incentive effect does not exsist, and there isn't any brain gain to compensate for the brain drain.

people with tertiary education or more, can be measured either with reference to the U.S. or to the whole OECD area. In any case, it is important to underline that the predictions of our model require α_5 to be positive. Roughly speaking, it means that a brain drain can exert a positive effect on growth when it is associated with a large enough middle class (a high MID); in general it can be said that we look for non-linearities in the relation between brain drain and growth.

4.1 Data

4.1.1 Brain drain

Doing empirical work on the economics of migration has usually undergone severe limitations due to the lack of extensive and reliable datasets on this issue. However, the data provided by Carrington and Detragiache (1998) for a sample of developing countries can be used for statistical inference, once we combine them with well known data on educational attainment and income inequality.

It has to be said that these data (CD henceforth) have been for several years the unique reference for the empirical literature on the brain drain. However, their reliability is not unquestioned: in fact the CD estimates of the emigration rates are obtained starting from three main statistical sources (US Census data on the skill composition of immigration, OECD data on migration inflows by sending country, and the Barro-Lee data on educational attainment in source countries), but relying for the rest on quite strong assumptions[16].

That's why Docquier and Marfouk (2004) have built a new database (DM throughout the rest of the paper) on the brain drain, that aims to improve over CD in two respects. First, they integrate data on the skill distribution of immigrants for the vast majority of the OECD countries, refining the

[16] Like, for example, that the skill structure of US immigration applies also to the other receiving countries.

quality of the estimates. Second, they expand the dimension of the sample providing data about developing countries that were not covered in CD, but also supplying estimates about the North-North skilled migration. To have an idea, the CD dataset conveys reliable data on the brain drain evaluated at the year 1990 for a maximum of 52 developing countries[17], while the DM dataset, for the same year, contains complete information about 170 countries.

For purposes of comparison, we report in Table 2 both CD and MD data on brain drain estimates and general migration rates, for all those countries covered by Carrington and Detragiache (1998). The differences between the two studies are not negligible. As Docquier and Marfouk (2004) explain in details, the CD data led to a general underestimation of high-skilled migration outflows from developing countries [18].

Our strategy is to exploit both data sets to perform our estimations. We will proceed in three stages. First, we will work with brain drain estimates coming from CD. Then, we will use data from DM for the same countries covered by CD. Finally we will employ all the data in DM, thus introducing also the North-North brain drain.

4.1.2 Other variables

To account for equality and/or middle class size (MID) we rely on six different measures. The first four are taken from a popular dataset built by Barro and Lee (2000); $MIDT$ is computed as the ratio between the percentage of the population (aged over 25) that has been enrolled at most in secondary school and the sum of the percentages that have been enrolled at most in

[17]Since CD is organized on data about U.S. immigrants, we decided to keep only those countries for which U.S. emigrants are more than 30%.

[18]With some notable exceptions for countries like Algeria, Morocco or Turkey whose emigration is massively directed to non-US countries like France or Germany, and for which the brain drain was overestimated because of the erroneous assumption of identical skill distribution of immigrants (with respect to the US). However, data on these countries were considered as being non-reliable according to the 30% criterion we adopted, and then do not appear in Table 1.

	Carrington and Detragiache (1998)					Docquier and Marfouk (2004)		
Country	*BDO*	*BDUS*	*MR0*	*MRUS*	*US/OECD migr.*	*BD*	*MR*	*reliability rate*
Argentina	2.7	1.9	0.6	0.4	72.3	3.3	0.9	94.8
Benin	0.4	0.4	na	na	100.0	6.1	0.2	91.5
Bolivia	4.2	4.2	0.7	0.7	100.0	5.9	1.0	95.4
Brazil	1.4	0.6	0.2	0.1	44.0	1.7	0.3	62.5
Cameroon	3.2	3.2	na	na	100.0	15.2	0.4	93.6
Central African Republic	1.7	1.7	na	na	100.0	4.4	0.2	96.8
Chile	6.0	3.3	1.1	0.6	54.3	6.3	1.7	82.5
China	3.0	1.4	0.1	na	51.5	3.1	0.2	79.4
Colombia	5.8	5.6	1.1	1.1	96.9	9.2	1.8	97.0
Costa Rica	7.1	7	2.4	2.4	100.0	7.7	2.6	99.0
Dominican Republic	14.7	14.2	6.5	6.3	96.7	17.9	7.9	97.5
Ecuador	3.8	3.8	1.9	1.9	100.0	5.4	2.7	98.0
Egypt	5.0	2.5	0.5	0.3	50.6	5.3	0.7	86.4
El Salvador	26.1	26.1	11.3	11.4	100.0	32.9	14.4	99.3
Fiji	21.3	21.3	3.6	3.6	100.0	63.6	15.5	99.9
Gambia	61.4	59.1	0.2	0.2	100.0	76.0	1.3	81.4
Ghana	25.7	15.1	0.4	0.2	53.3	33.7	1.2	76.4
Guatemala	13.5	13.5	3.4	3.4	100.0	18.2	4.3	99.7
Guyana	77.5	77.3	14.5	14.5	100.0	89.2	28.0	98.8
Honduras	15.7	15.7	3.0	3.0	100.0	21.1	4.0	99.6
India	2.6	1.1	0.2	na	44.1	2.6	0.2	96.6
Indonesia	1.5	1.4	na	na	90.5	6.2	0.3	35.3
Jamaica	77.4	67.3	20.3	13.4	61.0	84.1	25.6	99.8
Kenya	10.0	9.9	0.1	0.1	100.0	26.9	0.5	96.6
Lesotho	2.9	2.9	na	na	100.0	6.2	0.1	92.5
Malawi	2.0	2.0	na	na	100.0	7.5	0.1	95.7
Mali	0.9	0.9	na	na	100.0	6.6	0.7	99.2
Mauritius	7.2	7.2	0.2	0.2	100.0	37.2	5.3	91.2
Mexico	10.3	10.3	7.7	7.7	100.0	10.4	7.4	99.9
Mozambique	8.6	8.6	na	na	100.0	18.2	0.8	99.3
Nicaragua	18.8	18.7	4.7	4.7	100.0	29.0	7.7	99.7
Pakistan	6.7	2.4	0.3	na	35.2	6.1	0.4	85.9
Panama	19.6	19.5	6.7	6.7	100.0	21.7	7.7	99.6
Papua New Guinea	2.2	2.2	na	na	100.0	35.2	0.8	99.4
Paraguay	2.0	1.9	0.2	0.2	100.0	3.2	0.6	96.8
Peru	3.4	3.0	1.0	0.9	87.1	5.6	1.6	85.3
Philippines	9.0	6.6	3.1	2.2	71.6	12.8	4.1	91.9
Rwanda	2.2	2.2	na	na	100.0	9.4	0.1	87.7
Sierra Leone	24.3	24.1	0.3	0.3	100.0	31.0	0.5	94.1
South Korea	14.9	5.7	4.2	1.6	36.0	20.2	4.8	40.0
Sudan	1.8	1.7	na	na	100.0	5.0	0.1	86.4
Syria	3.1	3.1	0.7	0.7	100.0	6.9	1.7	90.2
Thailand	1.5	1.2	0.2	0.2	87.6	2.4	0.4	86.5
Togo	1.3	1.3	na	na	100.0	8.9	0.5	90.3
Trinidad & Tobago	57.8	57.2	9.5	9.4	100.0	77.2	18.9	99.7
Uganda	15.5	15.4	0.1	0.1	100.0	29.9	0.4	95.6
Uruguay	3.8	3.7	1.1	1.1	100.0	6.1	1.9	96.3
Venezuela	2.1	1.6	0.4	0.3	77.4	3.9	0.8	96.9
Zambia	5.0	5.0	0.1	na	100.0	12.2	0.2	92.5
Zimbabwe	4.7	4.6	0.1	0.1	100.0	5.1	0.5	97.3
South Africa	7.9	2.6	0.4	0.1	32.4	7.2	0.5	96.2

Sources: Carrington and Detragiache (1998), Docquier and Marfouk (2004).
Notes: all data are in % and refer to 1990; by 'brain drain' we mean the migration rate of people with at least tertiary education.
Definitions of variables:
- BDO, BD: brain drain to OECD countries
- BDUS: brain drain to the U.S.
- MRO, MR: total migration rate to OECD countries
- MRUS: total migration rate to the U.S.

Table 2: The size of the brain drain: comparing data on migration from selected developing countries.

either first or post-secondary; $MIDC$ is obtained in the same way, but looking at completed educational levels instead of simple enrollment; $MIDdumT$ and $MIDdumC$ are dummy variables derived from $MIDT$ and $MIDC$ (they are given a value equal to 1 when the base variable is larger than 0.35 and 0.45 respectively, and 0 otherwise). These four measures focus on *schooling* inequality, and we think that they are very appropriate to capture the size of that middle class which could react to eventual educational incentive derived from migration. The last two measures focus on *income* inequality and are taken from Deininger and Squire's (1996) data set; $MIDUMQ$ is a dummy variables which assumes value 1 if the sum of the income shares of third and fourth quintiles exceeds 0.34; $MIDUMG$ is a dummy variable as well, which takes value 1 if the Gini coefficient in income distribution is lower than 45. Obviously, for all these variables we take 1990 values.

Data on income variables (ygr and I/Y) come from the Penn World Tables version 6.1 (updated October 2002)[19], and are averaged on a yearly base over the period 1990-2000; initial GDP per capita ($y0$) is obviously that of 1990.

4.2 Econometric results

Since we are employing two alternative data sets for the brain drain estimates, we start by showing in Table 3 the correlation matrix of the different measures of both high-skilled and general migration. What we are mostly interested in is the couple of variables that will enter the equation we want to estimate: BDO (from CD) and BD (from DM). Despite the notable differences in Table 1, their correlation is indeed quite high, even if it is lower than, for instance, the correlation between BD and $BDUS$ or between MRO and MR.

[19]This data set is available on www.pwt.econ.upenn.edu. The complete reference is: Heston, A., R. Summers and B. Aten (2002): "Penn World Table version 6.1", Center for International Comparisons at the University of Pennsylvania (CICUP). For more information on the variables, see also Summers and Heston (1991).

	BD	BDO	$BDUS$	MR	MRO	$MRUS$
BD	1.000	0.941	0.942	0.794	0.650	0.640
BDO		1.000	0.990	0.799	0.750	0.725
$BDUS$			1.000	0.808	0.742	0.741
MR				1.000	0.926	0.930
MRO					1.000	0.969
$MRUS$						1.000

Table 3: Measures of brain drain and general migration: correlation matrix

The first part of our econometric analysis is based on brain drain estimates taken from the CD data set. The estimation results are presented in Table 4: the average growth rate of per capita GDP(ygr) is the dependent variable. A benchmark case without the migration/equality issue is considered to allow for a comparison with Barro (2000) and de la Croix and Doepke (2003).

The regression equations are all estimated by means of the Generalized Method of Moments (GMM).

In Table 4, looking at Regression (1) (the benchmark case) we can see that the signs of the coefficients are as expected: the investment/GDP ratio has a positive effect, while the African dummy and initial GDP (accounting for convergence) have both negative coefficients. These findings reproduce the standard results of the empirical growth literature. In particular, the estimated coefficient of I/Y is remarkably close to de la Croix and Doepke's (2003) estimates.

Regressions (2)-(7) add the brain drain and the cross-term $BD * MID$ to the benchmark equation, employing different measures for MID. In the majority of cases, the two terms BD and $BD * MID$ (in its different declinations) are strongly significant and have got the expected sign. Moreover, the coefficient of $BD * MID$ turns out to be, in absolute value, larger than the coefficient of BD. This is true for both type of (in)equality measures (schooling and income).

In general the J-test, that tells us if the residuals are reasonably close to being orthogonal to the instruments we are using, never rejects the overiden-

Independent	Regressions						
variables	(1: benchmark)	(2: MIDT)	(3: MIDdumT)	(4: MIDC)	(5: MIDdumC)	(6: MIDdumQ)	(7: MIDdumG)
constant	12.01* (6.60)	7.71 (6.88)	13.19* (6.96)	9.52 (6.52)	11.04 (6.92)	10.80 (6.81)	12.40* (7.34)
I/Y	0.13** (0.03)	0.13** (0.03)	0.14** (0.03)	0.12** (0.03)	0.16** (0.03)	0.16** (0.04)	0.12** (0.03)
$log(y0)$	-1.46* (0.79)	-0.94 (0.83)	-1.55* (0.84)	-1.16 (0.77)	-1.42* (0.85)	-1.33 (0.89)	-1.38 (0.90)
$dumAFR$	-3.92** (1.10)	-3.37** (1.12)	-3.92** (1.22)	-2.96** (0.99)	-3.32** (1.11)	-2.52** (1.19)	-4.02** (1.42)
BDO		-0.09* (0.05)	-0.20* (0.11)	-0.17** (0.07)	-0.05 (0.03)	-0.18* (0.10)	-0.25** (0.12)
$BDO * MID$		0.18** (0.08)	0.23** (0.10)	0.37** (0.13)	0.13** (0.05)	0.27** (0.10)	0.32** (0.11)
n. obs.	50	50	50	50	50	35	42
J-test	7.76	3.73	2.78	2.28	4.22	5.80	5.97
χ-sq. 5%	14.07	11.07	11.07	11.07	11.07	11.07	11.07

Notes.
The dependent variable is the growth rate of real per capita GDP (ygr). Definitions of the independent variables are given in the text.
As instruments, we use: constant, log of initial per capita GDP, log of initial per capita GDP squared, initial investment over GDP, Africa dummy, initial life expectancy (from U.S. Bureau of Census), initial life expectancy squared, total enrollment (in %) in secondary school at year 1985, percentage of population with no schooling at year 1985, and the tropic and distance variables of Gallup et al. (1999).
The J-test is the one for overidentifying restrictions proposed by Hansen (1982): it asymptotically behaves as a χ^2 with n degrees of freedom, and n (the number of overidentifying restrictions) is given by the difference between the number of predetermined variables (instruments) and the number of estimated coefficients.
Standard errors are reported in parenthesis; $**$ and $*$ denote significance at the 10% and at the 5% respectively.

Table 4: GMM estimations (CD)

tifying restrictions at the 5% level.

It's worth saying that, with respect to most of the empirical studies on growth, we don't have G/Y among the explanatory variables. This choice allows us to gain degrees of freedom. However, we performed regression with G/Y as well, but the substance of the results did not change; nevertheless, it has to be said G/Y happened to be non-significant in a few cases, and the J-test produced less satisfactory results. We would also underline that the results reported in Table 4, which are obtained employing data on high-skilled migration to the whole group of OECD countries ($BD \equiv BDO$), hold essentially unchanged if instead we consider data on the brain drain directed to the U.S. ($BD \equiv BDUS$)[20].

The last remark is about the Africa dummy, which appears with a smaller coefficient in the specification with $MIDdumQ$. This result is due to the fact that $MIDdumQ$ is derived from Deininger and Squire's data on inequality: since data on the quintile distribution of income are often missing for many of the poorest African countries, these country are dropped from the sample, leading to a weaker dummy coefficient.

The second part of our exercise consists in re-estimating our equation for the same developing countries as before, but using DM data instead of CD data. This can be seen as a sort of test of robustness, and the estimation results are presented in Table 5.

We see that the results are fairly stable when compared to the ones presented in Tables 4. Changing the database has not altered the substance of our findings, apart from some gains (like in the equation with $MIDC$) or losses ($MIDdumQ$) in the significance of the explanatory variables.

It's also worth noting that in most cases the cross-term $BD \cdot MID$ tends to be more significant than BD, as it was before.

Finally, we have rerun our regression trying to use all the information contained in Docquier and Marfouk's data set. Unfortunately, out of 170 countries for which they provide brain drain estimates, only a maximum of 91

[20] All these complementary estimation results are available upon request.

Independent	Regressions						
variables	(1: benchmark)	(2: MIDT)	(3: MIDdumT)	(4: MIDC)	(5: MIDdumC)	(6: MIDdumQ)	(7: MIDdumG)
constant	12.01* (6.60)	8.02 (6.78)	12.60* (6.93)	19.06** (9.03)	10.77 (6.69)	2.37 (4.65)	10.86* (6.48)
I/Y	0.13** (0.03)	0.12** (0.03)	0.13** (0.04)	0.12** (0.03)	0.15** (0.03)	0.11** (0.04)	0.12** (0.03)
$log(y0)$	-1.46* (0.79)	-0.95 (0.82)	-1.44* (0.83)	-2.26** (1.07)	-1.34 (0.82)	-0.20 (0.63)	-1.23 (0.80)
$dumAFR$	-3.92** (1.10)	-3.36** (1.11)	-3.59** (1.21)	-4.59** (1.54)	-3.21** (1.06)	-0.30 (0.91)	-3.63** (1.28)
BD		-0.08* (0.04)	-0.13* (0.07)	-0.10** (0.05)	-0.06* (0.03)	-0.16* (0.08)	-0.17* (0.09)
$BD * MID$		0.14** (0.06)	0.15** (0.06)	0.21** (0.08)	0.11** (0.05)	0.19** (0.07)	0.24** (0.08)
n. obs.	50	50	50	50	50	35	42
J-test	7.76	3.59	3.07	1.69	4.34	7.03	6.27
χ-sq. 5%	14.07	11.07	11.07	11.07	11.07	11.07	11.07

Notes.
The dependent variable is the growth rate of real per capita GDP (ygr). Definitions of the independent variables are given in the text.
As instruments, we use: constant, log of initial per capita GDP, log of initial per capita GDP squared, initial investment over GDP, Africa dummy, initial life expectancy (from U.S. Bureau of Census), initial life expectancy squared, total enrollment (in %) in secondary school at year 1985, percentage of population with no schooling at year 1985, and the tropic and distance variables of Gallup et al. (1999).
The J-test is the one for overidentifying restrictions proposed by Hansen (1982): it asymptotically behaves as a χ^2 with n degrees of freedom, and n (the number of overidentifying restrictions) is given by the difference between the number of predetermined variables (instruments) and the number of estimated coefficients.
Standard errors are reported in parenthesis; $**$ and $*$ denote significance at the 10% and at the 5% respectively.

Table 5: GMM estimations (DM, developing countries)

have simultaneously available data for the other variables we need (with data on inequality representing the most severe limitation). Then, our sample is less than doubled, from 50 to 91 countries. However, it is interesting to verify how the inclusion of OECD countries in the sample, and thus the explicit consideration of a North-North brain drain, modifies the picture. Table 6 presents the estimation output.

Two things arise quite clearly. First, the quality of the *growth* regression improves a lot: all the standard explanatory variables are now always strongly significant (with the usual exception of the Africa dummy, when a number of African countries is dropped out of the sample). Second: while the interactive term $BD \cdot MID$ continues to be significant and to appear with the expected sign, the BD term, i.e. the brain drain *per se* loses its significance in most cases. Technically, this comes as a consequence of the introduction of industrialized countries in the sample. The intuition behind this fact could be that, apart from the incentive effect that still holds for developed economies, in those countries the brain drain in itself is less harmful, maybe because of the easier replacement of high skilled workers.

The same remark we made about public expenditure when commenting Table 3 applies. In fact, both the results in Table 4 and 5 would not have undergone any significant change after including G/Y in the set of independent variables.

To sum up our empirical findings, we would say that our econometric work confirms that the relation between brain drain and growth is highly non-linear, with this non-linearity being possibly linked to some inequality (or middle class size) measure. As it was suggested by our model, only the presence of a fairly numerous middle class can make really effective (in terms of overall growth) the incentive effect induced by an increased migration chance, thus possibly encouraging developing countries to set up quite "permissive" emigration policies.

Moreover, these results are quite robust to the change of the database (from CD to DM) if the estimations are limited to a cross-section of develop-

Independent	Regressions						
variables	(1: benchmark)	(2: MIDT)	(3: MIDdumT)	(4: MIDC)	(5: MIDdumC)	(6: MIDdumQ)	(7: MIDdumG)
$constant$	6.59** (2.40)	8.62** (2.58)	8.73** (2.76)	9.63** (2.69)	8.52** (2.22)	8.17** (3.40)	9.30** (3.54)
I/Y	0.15** (0.03)	0.15** (0.04)	0.14** (0.04)	0.16** (0.04)	0.19** (0.04)	0.10** (0.05)	0.13** (0.05)
$log(y0)$	-0.83** (0.30)	-1.07** (0.35)	-1.02** (0.36)	-1.19** (0.36)	-1.16** (0.31)	-0.83* (0.45)	-1.02** (0.45)
$dumAFR$	-2.77** (0.75)	-3.14** (0.82)	-2.65** (0.89)	-2.94** (0.79)	-2.90** (0.75)	-0.45 (0.87)	-2.03* (1.05)
BD		-0.05 (0.05)	-0.10 (0.08)	-0.06 (0.05)	-0.03 (0.04)	-0.24** (0.10)	-0.23 (0.15)
$BD * MID$		0.11** (0.05)	0.12* (0.07)	0.12** (0.05)	0.13** (0.04)	0.23** (0.10)	0.28** (0.14)
n. obs.	91	91	91	91	91	70	77
J-test	10.60	5.41	5.64	2.83	4.05	5.17	5.79
χ-sq. 5%	14.07	11.07	11.07	11.07	11.07	11.07	11.07

Notes.

The dependent variable is the growth rate of real per capita GDP (ygr). Definitions of the independent variables are given in the text.

As instruments, we use: constant, log of initial per capita GDP, log of initial per capita GDP squared, initial investment over GDP, Africa dummy, initial life expectancy (from U.S. Bureau of Census), initial life expectancy squared, total enrollment (in %) in secondary school at year 1985, percentage of population with no schooling at year 1985, and the tropic and distance variables of Gallup et al. (1999).

The J-test is the one for overidentifying restrictions proposed by Hansen (1982): it asymptotically behaves as a χ^2 with n degrees of freedom, and n (the number of overidentifying restrictions) is given by the difference between the number of predetermined variables (instruments) and the number of estimated coefficients.

Standard errors are reported in parenthesis; $**$ and $*$ denote significance at the 10% and at the 5% respectively.

Table 6: GMM estimations (DM, full sample)

ing countries. On the contrary, the inclusion of industrialized countries in the sample, while confirming the incentive effect argument, makes less clear the negative effect that the brain drain "per se" should exert on overall growth.

5 Conclusions

In this paper we have proposed an OLG model of human capital led growth, to show that allowing for emigration of the highest skilled individuals from a developing country is likely to encourage human capital formation (thus fostering domestic growth in the short run), only if human capital is quite equally distributed, i.e. if there exists a numerous enough "middle class" waiting for educational opportunities. However, in a long run perspective this incentive gets dispersed (in the sense that it will touch a progressively shrinking portion of the population), while the human capital flight lasts forever, depressing average income.

Our "short run" claim could be straightly translated into an econometrically testable equation: as a consequence, we have been able to perform some growth regressions using available data on production, inequality and high-skilled migration. Estimation outputs turned out to be in line with our theoretical findings, both when schooling and income inequality have been considered. These results have shown remarkable robustness both to the change in the dataset for a cross-section of developing countries and to the inclusion of developed countries in the sample.

Moreover, we have also proposed a deeper interpretation of our main result: in fact it entrains the claim that in a given backward economy, inequality is harmful to growth, when it is seen in a perspective of progressive opening of the frontiers. In other words our model would suggests that, as skilled workers become more and more mobile, a new channel through which more inequality may translate into slower growth for developing countries begins to operate and should deserve consideration.

To conclude, we would underline that our model could be extended at least along one direction. In fact, we did not consider the issue of social mobility (from one class to another), that is likely to be quite important in qualifying the dynamic relationship between inequality and growth.

References

[1] Azariadis, C. (1996): "The economics of poverty traps. Part one: complete markets", *Journal of Economic Growth* 1, 449-486.

[2] Azariadis, C. and A. Drazen (1990): "Threshold externalities in economic development", *Quarterly Journal of Economics* 105 (2), 501-526.

[3] Barro, R.J. (2000): "Inequality and growth in a panel of countries", *Journal of Economic Growth* 5, 5-32.

[4] Barro, R.J. and J.-W. Lee (2000): "International data on educational attainment: updates and implications", *CID Working Paper* no. 42. [Data set available on: www.cid.harvard.edu/ciddata/ciddata.html]

[5] Beine, M., F. Docquier and H. Rapoport (2001): "Brain drain and economic growth: theory and evidence", *Journal of Development Economics* 64, 275-289.

[6] Benabou, R. (1996): "Inequality and growth", in Bernanke, B.S. and J. Rotemberg (eds.): *NBER Macroeconomics Annual* 11, MIT Press.

[7] Bhagwati, J.N. and K. Hamada (1974): "The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis", *Journal of Development Economics* 1, 19-24.

[8] Carrington, W.J. and E. Detragiache (1998): "How Big is the Brain Drain?", *IMF Working Paper* 98/102.

[9] de la Croix, D. and M. Doepke (2003): "Inequality and growth: why differential fertility matters", *American Economic Review* 93 (4), 1091-1113.

[10] de la Croix, D. and M. Doepke (2004): "Public versus private education when differential fertility matters", *Journal of Development Economics* 73, 607-629.

[11] de la Croix, D. and P. Michel (2002): *A Theory of Economic Growth. Dynamics and Policy in Overlapping Generations*, Cambridge University Press.

[12] Deininger, K. and L. Squire (1996): "A new data set measuring income inequality", *The World Bank Economic Review* 10 (3), 565-591. [Data set available on: www.worldbank.org/research/growth/dddeisqu.html]

[13] Docquier, F. and A. Marfouk (2004): "Measuring the international mobility of skilled workers (1990-2000) - Release 1.0", *World Bank Working Paper* 3381.

[14] Domingues Dos Santos, M. and F. Postel-Vinay (2003): "Migration as a source of growth: the perspective of a developing country", *Journal of Population Economics* 16, 161-175.

[15] Gallup, J.L., J.D. Sachs and A. Mellinger (1999): "Geography and economic development", *CID Working Paper* no. 1. [Data set available on: www.cid.harvard.edu/ciddata/ciddata.html]

[16] Greene, W.H. (2000): *Econometric analysis*, Prentice Hall.

[17] Haque, N.U. and S.-J. Kim (1995): "Human Capital Flight: Impact of Migration on Income and Growth", *IMF Staff Papers* 42 , 577-607.

[18] Hansen, L.P. (1982): "Large sample properties of Generalized Method of Moments estimators", *Econometrica* 50 (4), 1029-1054.

[19] Kim, J. (1998): "Economic analysis of foreign education and students abroad", *Journal of Development Economics* 56, 337-365.

[20] Lucas, R.E. (1988): "On the mechanics of economic development", *Journal of Monetary Economics* 22 (1), 3-42.

[21] Miyagiwa, K. (1991): "Scale economies in education and the brain drain problem", *International Economic Review* 32, 743-759.

[22] Mountford, A. (1997): "Can a brain drain be good for growth in the source economy?", *Journal of Development Economics* 53, 287-303.

[23] Reichlin, G. and A. Rustichini (1998): "Diverging patterns with endogenous labor migration", *Journal of Economic Dynamics and Control* 22, 703-28.

[24] Stark, O., C. Helmenstein and A. Prskawetz (1997): "A brain drain with a brain gain", *Economic Letters* 55, 227-234.

[25] Stark, O. and Y. Wang (2002): "Inducing human capital formation: migration as a substitute for subsidies", *Journal of Public Economics* 86, 29-46.

[26] Summers, R. and A. Heston (1991): "The Penn World Table (Mark 5): an expanded set of international comparisons, 1950-1988", *Quarterly Journal of Economics* 106 (2), 327-368.

[27] U.S. Census Bureau (2004): *International Data Base (IDB).* [Data available on: www.census.gov/ipc/www/idbprint.html]

[28] Wong, K.-Y. and C.K. Yip (1999): "Education, economic growth, and brain drain", *Journal of Economic Dynamics and Control* 23, 699-726.

A Proof of Proposition 1

Step 0.
First, we claim that $\Omega_t(h_t, \tau_t) = (1-\tau_t)h_t^\delta w_t + a\tau_t^\sigma h_t^\gamma \frac{w_{t+1}}{R_{t+1}}$ is monotonically increasing in h_t.

The simple proof is given by taking its first partial derivative with respect to h_t:

$$\frac{\partial \Omega_t}{\partial h_t} = (1-\tau_t)\delta h_t^{\delta-1} w_t + a\tau_t^{\sigma}\gamma h_t^{\gamma-1}\frac{w_{t+1}}{R_{t+1}} \tag{13}$$

which is positive for every h_t.

Step 1.
Let's consider $\tau_1 < \tau_2$;
then: $\exists h_t = h_{1,2} : \Omega(\tau_1) = \Omega(\tau_2)$
and precisely it is given by:

$$h_{1,2} = \left[\frac{w_t R_{t+1}}{a w_{t+1}}\frac{(\tau_1 - \tau_2)}{(\tau_1^{\sigma} - \tau_2^{\sigma})}\right]^{\frac{1}{\gamma-\delta}} \tag{14}$$

Similarly, there do exist $h_{2,3}$ and $h_{1,3}$ as defined in Proposition 1.

Step 2.
We claim that, *for appropriate values of the parameters* ,
$\Omega(\tau_j) > (<)\Omega(\tau_i)$ for $h > (<)h_{i,j}$
where $i, j = 1, 2, 3$ and $i < j$.
In other words we want to show that the lifetime income curve associated with higher education crosses *from below* the one with lower schooling, meaning that higher education will be more convenient *for relatively higher values of parental human capital.*
To prove this, we simply have to find conditions on the parameters such that the following holds:

$$\frac{\partial^2 \Omega_t(\tau_t, h_t)}{\partial h_t \partial \tau_t} > 0 \tag{15}$$

and, since:

$$\frac{\partial^2 \Omega_t(\tau_t, h_t)}{\partial h_t \partial \tau_t} = -\delta h_t^{\delta-1} w_t + a\sigma\tau_t^{\sigma-1}\gamma h_t^{\gamma-1}\frac{w_{t+1}}{R_{t+1}},$$

inequality (15) is true for:

$$\tau_t < \left(\frac{a\sigma\gamma}{\delta}\frac{w_{t+1}}{w_t R_{t+1}}\right)^{\frac{1}{1-\sigma}} h_t^{\frac{\gamma-\delta}{1-\sigma}}. \tag{16}$$

As it is, (16) is not a simple condition on the parameters, since it involves h_t, an endogenous variable of our model. But, since it poses a lower bound on h_t and $h_t > \underline{h}$, we can ensure that it holds at time t if:

$$\tau_t < \left(\frac{a\sigma\gamma}{\delta} \frac{w_{t+1}}{w_t R_{t+1}} \right)^{\frac{1}{1-\sigma}} \underline{h}^{\frac{\gamma-\delta}{1-\sigma}}. \tag{17}$$

Step 3.

We need that:

$\tau_3 > \tau_2 \Rightarrow h_{2,3} > h_{1,2}$,

which is true for:

$$\tau > \sigma^{\frac{1}{1-\sigma}}. \tag{18}$$

In fact:

$$h_{2,3} = \left[\frac{w_t R_{t+1}}{w_{t+1}} \frac{1}{a} \frac{(\tau_2 - \tau_3)}{(\tau_2^\sigma - \tau_3^\sigma)} \right]^{\frac{1}{\gamma-\delta}}$$

and

$$h_{1,2} = \left[\frac{w_t R_{t+1}}{w_{t+1}} \frac{1}{a} \frac{(\tau_1 - \tau_2)}{(\tau_1^\sigma - \tau_2^\sigma)} \right]^{\frac{1}{\gamma-\delta}};$$

and therefore, $h_{2,3} > h_{1,2}$ if and only if:

$$\frac{(\tau_2 - \tau_3)}{(\tau_2^\sigma - \tau_3^\sigma)} > \frac{(\tau_1 - \tau_2)}{(\tau_1^\sigma - \tau_2^\sigma)}.$$

The latter holds when the slope of τ^σ is less than 1, which means for all the values of τ implied by (18).

Step 4.

To complete our proof, we now simply need to prove that $h_{1,2} < h_{1,3} < h_{2,3}$, i.e. that the crossing point between $\Omega(\tau_1)$ and $\Omega(\tau_3)$ lies between the other two.

Why should it be true?

It is true because both the inequalities $h_{1,3} > h_{2,3}$ and $h_{1,2} < h_{1,3}$ are verified for the same values of (τ, σ) which validate Step 3.

Then, $h_{1,2} < h_{1,3} < h_{2,3}$ necessarily holds, when also $1 > \sigma\tau^{\sigma-1}$ holds; that means for all those values of τ such that the slope of the function τ^{σ} is less than 1.

Putting all things together, for the whole proof to hold, we need to assume simply that:

$$\sigma^{\frac{1}{1-\sigma}} < \tau < \left(\frac{a\sigma\gamma}{\delta}\frac{w_{t+1}}{w_t R_{t+1}}\right)^{\frac{1}{1-\sigma}} \underline{h}^{\frac{\gamma-\delta}{1-\sigma}}. \tag{19}$$

We can try to provide some economic intuitions about the above inequality. In fact, we can see that the first part of the inequality becomes less and less binding as σ decreases; this means when parental skill is not so important in the reproduction of human capital (i.e., when time spent in school is much more important than parents), even with a very low starting level of τ_1. A possible deeper meaning, however, is not trivial and is not easy to find. In fact Step 3 tells us simply that, given $\tau_3 > \tau_2$, the individual who is indifferent between τ_2 and τ_3 has more parental human capital than the one who is indifferent between τ_1 and τ_3.

On the other hand, the second part of inequality (19), when satisfied, ensures that higher education is relatively more attractive for higher values of parental human capital (the crossing-from-below property of our curves). It becomes less and less binding (at the limit it is not a restriction any more) as human capital increases, and as the returns of education increase, while the effect of σ is ambiguous. In fact we can rewrite (19) as:

$$\sigma^{\frac{1}{1-\sigma}} < \tau < \sigma^{\frac{1}{1-\sigma}} \left(\frac{a\gamma}{\delta}\frac{w_{t+1}}{w_t R_{t+1}}\right)^{\frac{1}{1-\sigma}} \underline{h}^{\frac{\gamma-\delta}{1-\sigma}}. \tag{20}$$

Ceteris paribus, with too high values of "offered" τ's, the latter may not hold, and we would have a crossing-from-above situation, since in this case the weight of parental human capital is somewhat "minimized" (when compared with the large amount of schooling time employed in the production of new human capital), and thus education becomes relatively more attractive for the less skilled.

B Proof of Proposition 2

Recalling that we are dealing with the case with two possible educational levels, we can write:

$$g_t^M = \frac{\left\{\int_{\underline{h}_t}^{h_{1,2}^M} a\tau_1^\sigma h_t^\gamma f(h_t)dh_t + (1-m)\int_{h_{1,2}^M}^{\overline{h}_t} a\tau_2^\sigma h_t^\gamma f(h_t)dh_t\right\} \cdot \left[\frac{1}{1-m\cdot\int_{h_{1,2}^M}^{\overline{h}_t} f(h_t)}\right]}{\int_{\underline{h}_t}^{\overline{h}_t} h_t f(h_t)dh_t}. \tag{21}$$

Since we are in a short-run perspective, we can simplify assuming that $\gamma = 1$.

Then we have that:

$$\lim_{m\to 0}\frac{\partial g_t^M}{\partial L} = \frac{a(\tau_1^\sigma - \tau_2^\sigma)[(\widetilde{h}+\frac{L}{2})(\widetilde{h}-\frac{L}{2}) - (h_{1,2}^A)^2]}{2L^2\widetilde{h}}$$

and

$$\lim_{m\to 1}\frac{\partial g_t^M}{\partial L} = -\frac{a\tau_1^\sigma}{4\widetilde{h}_t}.$$

The limit for $m \to 1$ is always negative, while the limit for $\to 0$ is always positive if $h_{1,2}^A > \widetilde{h}$. The claim of Proposition 2 is thus established.

C Proof of Proposition 3

To prove Proposition 3 we need to focus on the two factors composing the numerator of the expression for g_t^M, since the denominator does not depend on m. These two factors are respectively:

$$A = \int_{\underline{h}_t}^{h_{1,2}^M} a\tau_1^\sigma h_t^\gamma f(h_t)dh_t + (1-m)\int_{h_{1,2}^M}^{\overline{h}_t} a\tau_2^\sigma h_t^\gamma f(h_t)dh_t$$

and

$$B = \frac{1}{1 - m \cdot \int_{h_{1,2}^M}^{\overline{h}_t} f(h_t)}.$$

It is easy to see that $\frac{\partial B}{\partial m}$ is always positive. Therefore, if A gets a maximum for $0 < m \leq 1$, the same can be said about g_t^M as a whole.

We need the following inequalities to hold simultaneously: $\lim_{m \to 0} \frac{\partial A}{\partial m} > 0$ and $\lim_{m \to 1} \frac{\partial A}{\partial m} < 0$.

As for the first limit, we have that:

$$\lim_{m \to 0} \frac{\partial A}{\partial m} = \frac{a\tau_2^\sigma}{2L} \left\{ \frac{[2w_{t+1}^F - (1+\delta)w_{t+1}]}{(1-\delta)w_{t+1}} (h_{1,2}^A)^2 - (\widetilde{h} + \frac{L}{2})^2 \right\},$$

and its positiveness is granted for:

$$(\widetilde{h} + \frac{L}{2}) < h_{1,2}^A \left[\frac{2w_{t+1}^F - (1+\delta)w_{t+1}}{(1-\delta)w_{t+1}} \right]^{1/2}.$$

On the other hand:

$$\lim_{m \to 1} \frac{\partial A}{\partial m} = \frac{a\tau_2^\sigma}{2L} \left\{ K(h_{1,2}^A)^2 - (\widetilde{h} + \frac{L}{2})^2 \right\},$$

where

$$K = \left[1 - \frac{2(w_{t+1}^F - w_{t+1})\tau_1^\sigma}{(1-\delta)(w_{t+1}^F \tau_2^\sigma - w_{t+1}\tau_1^\sigma)} \right]$$

is negative provided that:

$$(\widetilde{h} + \frac{L}{2}) > \sqrt{K} h_{1,2}^A.$$

Then, we can say that Proposition 3 holds for:

$$h_{1,2}^A \left[1 - \frac{2(w_{t+1}^F - w_{t+1})\tau_1^\sigma}{(1-\delta)(w_{t+1}^F \tau_2^\sigma - w_{t+1}\tau_1^\sigma)} \right]^{1/2} < (\widetilde{h} + \frac{L}{2}) < h_{1,2}^A \left[\frac{2w_{t+1}^F - (1+\delta)w_{t+1}}{(1-\delta)w_{t+1}} \right]^{1/2}.$$

To conclude, let's underline that a sufficient condition for

$$\left[\frac{2w^F_{t+1}-(1+\delta)w_{t+1}}{(1-\delta)w_{t+1}}\right]-\left[1-\frac{2(w^F_{t+1}-w_{t+1})\tau_1^\sigma}{(1-\delta)(w^F_{t+1}\tau_2^\sigma-w_{t+1}\tau_1^\sigma)}\right]=\frac{2(w^F_{t+1}-w_{t+1})(w^F_{t+1}\tau_2^\sigma-2w_{t+1}\tau_1^\sigma)}{(1-\delta)w_{t+1}(w^F_{t+1}\tau_2^\sigma-w_{t+1}\tau_1^\sigma)}$$

to be positive, is simply that:

$$\frac{w^F_{t+1}}{w_{t+1}}>2\left(\frac{\tau_1}{\tau_2}\right)^\sigma.$$