Optimal Migration: A World Perspective*

by

Jess Benhabib and Boyan Jovanovic[1]

December 28, 2010

ABSTRACT

We ask what level of migration would maximize world welfare. Welfare is assumed to be a weighted average of the utilities of the world's various citizens, but the weights are also country specific. Using a calibrated one-sector model we find that unless the weights are heavily biased towards the natives of rich countries, the extent of migration that would be optimal far exceeds the levels observed today. The claim remains true in a two-sector extension of the model. All versions of the model assume that migration is the only redistributive tool.

KEYWORDS: World welfare optimum, inequality, migration.

JEL CLASSIFICATION NUMBER: O15

AUTHOR INFORMATION:

Jess Benhabib	Boyan Jovanovic
jb2@nyu.edu	bj2@nyu.edu

Department of Economics

New York University

19 W. 4th St.

New York, N.Y. 10012

USA

Running head: Optimal Migration

* Original version received September 14, 2009. Final version received December 28, 2010.

[1] We thank Jonathan Morduch and the referees for comments, the NSF and the Kauffman Foundation for support, and Matthias Kredler for doing all the computations and producing the associated plots.

1 Introduction

What is the optimal distribution of personal incomes in the world, and how best can it be attained? Economists have studied the question of domestic redistribution using taxes and subsidies as the instruments.

When dealing with the world as a whole, however, political constraints arise that limit the usefulness of taxes and transfers. The limits and difficulties of foreign aid or international redistribution have been studied by Easterly (2006): foreign aid funds are often wasted or misdirected.

If public foreign aid has been ineffective in reducing international inequality, so have private capital flows failed to equalize returns and factor prices: Lucas (1990) argued that inequality originates in human-capital differences and that physical capital flows can do little to eliminate inequality.

A neglected mechanism for reducing inequality, the flip side of capital flows, is international migration. Using a one-sector model we find that the optimal use of migration alone would seemingly raise welfare to levels far above those achieved at present. The mechanism that achieves this is the spillover of knowledge flowing to immigrants when they work along side highly skilled natives. Starting from the differences in human capital currently in place, optimal migration policy would involve moving far more people from poor to rich countries than the latter admit at present. We then confirm this conclusion, with and without spillovers, in a two-sector extension of the model.

In all versions of the model there is just one consumption good, and therefore only factor mobility (as opposed to trade in goods) can achieve equalization of prices per unit of skill. Moreover, and this is specific to our model, spillovers by assumption occur among people in the same geographical area, so people must locate together in order to share them.

In deriving the optimal migration policy we ignore political constraints, except implicitly those constraints that prevent foreign aid from being the main redistributive tool (for an analysis of political constraints see Benhabib (1996)). Redistribution involves winners and losers, but the losers, typically in rich countries, may block policies based on egalitarian social weights worldwide, and restrict access to their labor markets by foreigners. Our aim here is to derive the egalitarian or near-egalitarian ideal, as the reasons why the world falls short of the ideal are political. It is unlikely that the ideal can be collectively achieved in the near future, even as the poor of the world increasingly press for more open access to the labor markets of the rich countries.

In related work, Klein and Ventura (2007) study the steady state of a dynamic model with two location in which labor can move from one location to another at a cost and in which capital is costlessly mobile. The authors study optimal allocation of labor over the two regions. In spite of this similarity, our model differs in several ways from theirs. On the one hand, our model is static; on the other it endogenizes TFP along the lines suggested by Lucas (1988).

The model and the planner's problem and presented in Section 2. Calibration and Simulation are in Section 3 and several extensions are considered in Section 4. Section 5 concludes, and some proofs are in the Appendix.

2 The Model

We first introduce the production technology and its implications for segregated production. We then introduce the rest of model and the welfare-maximizing planner's problem.

2.1 Production technology

Immigrants in our model affect the well-being of the residents of the host country through an external effect. The effect works through human capital per person, $\bar{h}$, as in Lucas (1990). The output of a country is

$$Y = G\left(\bar{h}\right)H, \tag{1}$$

where H is the total human capital in the country. The private marginal product of human capital is $G\left(\bar{h}\right)$.[2]

Constant returns and decentralizability.—The production function (1) obeys constant returns to scale in the sense that doubling the number of residents while leaving the distribution of individual human capital h unchanged leaves $\bar{h}$ unaffected, but doubles H and, hence, Y. This allows for a competitive situation in which zero-profit firms (of indeterminate size) hire labor and pay a wage of $G\left(\bar{h}\right)$ per efficiency unit.

Efficiency vs. distribution.—The model has a tension between considerations of efficiency and distribution. Efficiency requires that production be segregated geographically. This is the content of Proposition 1. Let $M(h)$ be the world's distribution of human capital, and assume that

$$G(h) = h^{\alpha}, \tag{2}$$

where $\alpha > 0$.

Proposition 1 *World output is maximized when there is complete segregation by h, i.e.,*

$$Y \leq \int h^{1+\alpha} dM(h).$$

Proof. Suppose that there is a location in the world where people are heterogeneous in h. Let the distribution at that location have measure $\mu(h)$, with mean $\bar{h}$. Let the total output at that location be

$$y = G\left(\bar{h}\right)\left(\int h d\mu\right)$$

[2] Evidence that $G' > 0$ is in Clark (1987) who invokes national culture somewhat in line with the 'social capital' interpretation of Coleman (1988), and in Rauch (1993) who attributes it to human-capital spillovers at the regional level.

Then

$$\frac{1}{\int d\mu} y = \bar{h} G\left(\bar{h}\right) = \bar{h}^{1+\alpha} = \left(\frac{\int h d\mu}{\int d\mu}\right)^{1+\alpha} \leq \frac{1}{\int d\mu} \int h^{1+\alpha} d\mu$$

where the inequality follows because $d\mu / \int d\mu$ is a measure adding up to unity, and $h^{1+\alpha}$ is a convex function. Cancelling the multiplicative constant leaves us with

$$y \leq \int h^{1+\alpha} d\mu$$

and the inequality is strict if the support of μ has more than one point. Therefore no location can have heterogeneity of h. ■

The foregoing proposition and many of the other results of the paper extend to a world in which there is a perfectly mobile physical capital K and in which, instead of (1), the production function is $Y = G\left(\bar{h}\right) K^b H^{1-b}$. The reduced-form production function would take the form of (1) and an efficiency-equity tradeoff would remain: Migration from poor to rich countries would reduce both income inequality and world output.[3]

Exogenous TFP differentials.—Proposition 1 holds if the production function is the same in all locations. When this is relaxed, Proposition 1 does not generally hold. Suppose there is a finite number of locations. Instead of (1), let the production function at location i be

$$Y = S_i \bar{h}_i^{\alpha} H_i, \tag{3}$$

where S_i is a TFP specific to location i, and H_i and $\bar{h}_i$ are the total and average skill levels at location i. Let i^* be the location where TFP is the highest, and let

$$\lambda = \frac{S_{i^*}}{\max_{j \neq i^*} S_j} \geq 1$$

[3]The tradeoff arises because spillovers are local in nature because $\bar{h}$ is region specific. Firms in the same location share the same production function function determined by the TFP term $G\left(\bar{h}\right)$. Kremer and Maskin (2006) and Eeckhout and Jovanovic (forthcoming) relax the assumption that workers must be in the same location to participate in the same production process; their models do not have external effects in production.

be the ratio of the highest TFP level to the next highest.

Suppose the total measure of the world's agents is unity. Let $\bar{h}(\varepsilon)$ be the average skill among the measure ε of the most skilled agents in the distribution M. That is,

$$\bar{h}(\varepsilon) = \frac{1}{\varepsilon}\int_{\hat{h}(\varepsilon)}^{\infty} hdM(h)$$

where $\hat{h}(\varepsilon)$ solves the equation $\varepsilon = 1 - M(h)$. Then the world's average skill level is just $\bar{h}(1)$. We shall assume

A.1. $\varepsilon^{1/(1+\alpha)}\bar{h}(\varepsilon)$ is increasing in ε.

Example: To get a feel for **A.1.**, consider the Pareto form $M(h) = 1 - h^{-\rho}$ with $\rho > 1$ and $h \geq 1$. Then $\hat{h}(\varepsilon) = \varepsilon^{-1/\rho}$, and

$$\bar{h}(\varepsilon) = \frac{1}{\varepsilon}\int_{\varepsilon^{-1/\rho}}^{\infty} h\rho h^{-\rho-1}dh = \frac{1}{\varepsilon}\frac{\rho}{1-\rho}h^{1-\rho}\Big|_{h=\varepsilon^{-1/\rho}}^{\infty} = \frac{1}{\varepsilon}\frac{\rho}{\rho-1}\varepsilon^{-(1-\rho)/\rho} = \frac{\rho}{\rho-1}\varepsilon^{-1/\rho}$$

and therefore **A.1.** holds iff $\frac{1}{1+\alpha} > \frac{1}{\rho}$, i.e., iff

$$\rho > 1 + \alpha. \tag{4}$$

That is, the right tail must decay fast enough relative to the gains to segregating the ablest agents.

Proposition 2 *If A.1 holds and if $\lambda \to \infty$, then output is highest if all agents produce together at location i^**

Proof. If all the agents were at location i^*, output would be $Y^* \equiv S_{i^*}\bar{h}^{1+\alpha}$. Let $M_j(h)$ $(j = 1, 2, ...)$ be any other distribution of agents over the locations which adds up to M, i.e., $M(h) = \sum_j M_j(h)$. Let $\varepsilon_j = \int dM_j$ and $\bar{h}_j \equiv \frac{1}{\varepsilon_j}\int hdM_j$.This allocation yields an output of

$$Y = \sum_j S_j \int (\bar{h}_j)^{\alpha} hdM_j(h) = \varepsilon_{i^*}S_{i^*}\bar{h}_{i^*}^{1+\alpha} + \sum_{j\neq i^*}\varepsilon_j S_j \bar{h}_j^{1+\alpha}$$

Then since $\sum_{j \neq i^*} \varepsilon_j S_j \bar{h}_j^{1+\alpha} \leq \max_{j \neq i^*} S_j \sum_{j \neq i^*} \varepsilon_j \bar{h}_j^{1+\alpha}$, as $\lambda \to \infty$,

$$\frac{Y}{Y^*} \leq \frac{\varepsilon_{i^*} S_{i^*} \bar{h}_{i^*}^{1+\alpha} + \max_{j \neq i^*} S_j \sum_{j \neq i^*} \varepsilon_j \bar{h}_j^{1+\alpha}}{S_{i^*} \bar{h}^{1+\alpha}} \to \varepsilon_{i^*} \left(\frac{\bar{h}_{i^*}}{\bar{h}} \right)^{1+\alpha} \leq \left(\frac{\varepsilon^{1/(1+\varepsilon)} \bar{h}(\varepsilon)}{\bar{h}(1)} \right)^{1+\alpha} < 1$$

for any $\varepsilon < 1$. The last weak equality is based on the fact that the highest output producible by a mass ε of agents at i^* is obtained by taking the agents from the right tail of the distribution so that $\bar{h}_i \leq \bar{h}(\varepsilon)$. The last inequality is based on **A.1**, and that completes the proof. ■

Intuitively, because of the synergy in skills, something like A.1. is still needed to prevent it being optimal to segregate the highest-ability agents and place them in the highest-TFP location. The Pareto example expresses this intuition in condition (4).

Until section 4 we shall analyze and simulate the model on the assumption that there are no exogenous TFP differentials. That is, the production function will be assumed to be (1). The straight forward extension to including TFP differentials is in section 4.1.

2.2 The planner's problem

If taxes and subsidies had no disincentive effects and if tax proceeds could be distributed without waste or diversion, the optimal redistribution mechanism would be foreign aid. A world Planner would segregate people by skill, tax the rich and distribute the proceeds to the poor. But if foreign aid is not feasible, the Planner can use migration. This is the problem we shall now analyze.

Analysis.—Let μ_A be the pre-migration mean skills in country A and let the human capital of A's residents be distributed $h \sim F_A(h)$. Let μ_B be the mean skills in country B and let the human capital of B's residents be distributed $h \sim F_B(h)$, with density function $f_B(h)$. Let

$$x = \phi(h)$$

be the probability that a type-h resident of B will be allowed to emigrate to A. That is, $\phi : R \rightarrow [0,1]$.

A *skill-neutral* policy is one in which ϕ is a constant, independent of h. Policies that are not skill neutral are skill biased. Denote the average *post*-migration h levels in A and B by $\bar{h}_A$ and $\bar{h}_B$ respectively. Let n be the population of B relative to that of A. The latter is normalized to unity. Human capital per head in A is

$$\bar{h}_A = \frac{\mu_A + n \int h\phi(h)\, dF_B(h)}{1 + n \int \phi(h)\, dF_B(h)}, \tag{5}$$

and in B it is

$$\bar{h}_B = \frac{\int h\left[1 - \phi(h)\right] dF_B(h)}{\int \left[1 - \phi(h)\right] dF_B(h)}. \tag{6}$$

Migration costs and incentive compatibility.—We assume that each individual migrant loses a fraction t of his or her income in the host country. These are costs of assimilating, finding a job and so on and one would expect them to be proportional to potential income. Migration must be *voluntary* which now means that net of migration costs the migrant must earn more in the country of his or her destination than in the country of origin. This requires that for an immigrant with skill-level h, $G\left(\bar{h}_A\right)(1-t)\, h \geq G\left(\bar{h}_B\right) h$ or simply that[4]

$$G\left(\bar{h}_A\right)(1-t) \geq G\left(\bar{h}_B\right). \tag{7}$$

[4]The implicit assumption is that the immigrant spends t units of time on migrating and adapting to the new environment, leaving him with $1-t$ units of time for work. We nevertheless assume, for simplicity, that immigrants contribute fully to $\bar{h}_A$. One could alternatively assume that their contribution to $\bar{h}_A$ was proportional to $(1-t)$ in which case the RHS of (5) would become $\frac{\mu_A + (1-t)n \int h\phi(h) dF_B(h)}{1 + (1-t)n \int \phi(h) dF_B(h)}$. This would slightly lower the costs of an influx of low-h immigrants, slightly lower the benefits of a high-h influx, but otherwise leave the results largely unchanged. Our formulation gets support from Caponi (forthcoming) who finds that immigrants face a significant loss of capacity to translate their abilities into earnings but no loss of capacity to transfer

Practically, this constraint rules out policies that would send so many unskilled people from B to A (see "skimming from the bottom" defined in (10)) that $\bar{h}_B$ becomes so much higher $\bar{h}_A$ as to imply the negation of (7). Clearly, the higher is t, the larger the range of policies that (7) rules out.

Social welfare function and the Planner's problem.—The Planner is a Stackelberg leader. He announces a policy at the outset, and agents then choose their migration decisions and production takes place. Let θ and $(1-\theta)$ denote the welfare weights that the Planner assigns to utilities of the residents of A and B, respectively. Let $U(c)$ be an agent's utility function of consumption c. Agents simply consume their wages. The Planner chooses a function $\phi(h)$ to maximize

$$\theta \int U\left[G\left(\bar{h}_A\right) h\right] dF_A(h) + (1-\theta) n \int \left\{\phi(h) U\left(G\left[\bar{h}_A\right] h [1-t]\right) + [1-\phi(h)] U\left(G\left[\bar{h}_B\right] h\right)\right\} dF_B(h). \tag{8}$$

subject to (5), (6), and (7).

The planner therefore chooses the migration policy that maximizes the weighted sum of the utilities of agents, with the weights differing between natives of the host country and others. An alternative would be to have the planner assign weights as a function not of countries but of agents' characteristics alone. Such a welfare function would then give the same weight to agents with the same skill level, independently of their country of origin. This is a welfare function that one would regard as fair, and with our parametrization it obtains only when $\theta = 1/2$.

Until Section 4.2., we shall take the distribution of skills F_A and F_B as given, and throughout the paper we shall treat as exogenous the parameter n (the number of residents in the poor world relative to that of the rich). This means that a rise in α will have two opposing effects. It will raise the efficiency losses resulting from mixing the skill levels, but it also will widen the welfare losses resulting from differences in $\bar{h}_A$ and $\bar{h}_B$. In other words as α rises, the tension between distribution and efficiency

their human capital to their children; we loosely associate the latter with how the human capital of immigrants enters $G(\cdot)$.

gets stronger.

2.2.1 The optimal policy

The rest of the paper will assume that $\mu_A > \mu_B$, that h has no upper bound in the supports of F_A and F_B, and that

$$U(c) = \ln c. \tag{9}$$

In this case, we shall show that the optimal policy is skill dependent and of the "bang-bang" type: Among people of type h, either everyone should migrate or no one should do so. Moreover, the set of types is connected in that if type h_0 is allowed to migrate, then either everyone with h below h_0 is also allowed to migrate, or everyone *above* h_0 is allowed to migrate. Let us first describe these policies.

Skimming from the bottom of F_B.—Under this policy there exists a cutoff, $\tilde{h}$, such that everyone with $h < \tilde{h}$

$$\phi(h) = \begin{cases} 1 & \text{for} \quad h < \tilde{h} \\ 0 & \text{for} \quad h > \tilde{h} \end{cases} \tag{10}$$

Under this policy (10), there exists a unique $\tilde{h}$ that will equate the average skill in the two countries. Formally, there exists an unique $\hat{h} < \infty$ such that the RHSs of (5) and (6) are equated:[5]

$$\frac{\mu_A + n\int_0^{\hat{h}} h dF_B(h)}{1 + nF_B\left(\hat{h}\right)} = \frac{\int_{\hat{h}}^{\infty} h dF_B(h)}{1 - F_B\left(\hat{h}\right)}. \tag{11}$$

This policy always helps the non-migrating natives of B, and for $\tilde{h} < \hat{h}$ it helps the migrants too. The planner will use this policy when θ (the social weight on the rich country) is low.

[5]Proof: At $\hat{h} = 0$, the LHS is larger, whereas as $\hat{h} \to \infty$, the LHS$\to \frac{\mu_A + n\mu_B}{1+n}$ while the RHS$\to \infty$. Uniqueness is shown by showing that when evaluated at the solution of (11), the derivative of the RHS exceeds that of the LHS, ruling out multiple crossings

Skimming from the top of F_B.—This we shall refer simply a "brain-drain" policy:

$$\phi(h) = \begin{cases} 0 & \text{for} \quad h < \tilde{h} \\ 1 & \text{for} \quad h > \tilde{h} \end{cases} \tag{12}$$

For $\tilde{h}$ large enough, this policy raises $\bar{h}_A$ and helps the natives of A, while it hurts the non-migrating natives of B. The planner will use this policy when θ is high.[6]

The rest of this section will prove that the policy will indeed be bang-bang, taking the form (10) if θ is low and (12) if θ is high. The full characterization will emerge in section 3 where we shall simulate two versions of the model. To ease notation, let $g_i = \ln G\left(\bar{h}_i\right)$ for $i = A, B$.

Lemma 1 *When $U(C) = \ln C$, the maximand in (8) reduces to the equivalent*

$$W \equiv \theta_A^* g_A + \theta_B^* \left(g_B - \ln[1-t]\right) \tag{13}$$

subject to (5) and (6), where

$$\theta_A^* = \theta + (1-\theta)\omega n, \quad \theta_B^* = (1-\theta)n(1-\omega), \quad \textit{and} \quad \omega \equiv \int \phi(h)\, dF_B.$$

Proof. Substituting for U and leaving out terms that do not depend on ϕ, (8) reads

$$\begin{aligned} & \theta g_A + (1-\theta)\, n \int \{\phi(h)(g_A + \ln h + \ln[1-t]) + (1-\phi[h])(g_B + \ln h)\}\, dF_B \\ = \; & \theta g_A + (1-\theta)\, n \int \{\phi(h)\, g_A + (1-\phi[h])(g_B - \ln[1-t])\}\, dF_B + (1-\theta)\, n \int (h + \ln[1-t])\, dF_B. \end{aligned}$$

But the last terms does not depend on ϕ and we are left with (13). ■

Assume that the density f_B exists for all h, and define

$$z(h) = n f_B(h) \phi(h)$$

[6] For both policies at point $h = \tilde{h}$ we know only that $0 \leq \phi(h) \leq 1$, the Planner being indifferent about whether $\tilde{h}$ should migrate or not.

to be the new control variable that satisfies $z(h) : \; R \rightarrow [0, \; nf_B(h)]$ for all h. In terms of this control variable in (13) we have

$$\bar{h}_A = \frac{\mu_A + \int hz(h)\,dh}{1 + \int z(h)\,dh}, \quad \bar{h}_B = \frac{n\mu_B - \int hz(h)\,dh}{n - \int z(h)\,dh}, \quad \text{and} \quad n\omega \equiv Z = \int z(h)\,dh.$$

The constraint set for z is convex. We attach the multiplier λ_0 to the non-negativity constraint, and the multiplier λ_1 to the upper-bound constraint. The Planner faces the Lagrangian

$$\mathcal{L} = W + \int \lambda_0(h)\,z(h)\,dh - \int \lambda_1(h)\,z(h)\,.$$

We do not include the incentive-compatibility constraint (7) in the Lagrangian. Rather, we shall verify (7) ex post, and in Section 3 characterize the policies that (7) rules out. The FOC is

$$\frac{\partial W}{\partial z(h)} = \lambda_1(h) - \lambda_0(h)\,, \tag{14}$$

where $\frac{\partial W}{\partial z(h)}$ is evaluated at the optimal policy, the latter consisting of an entire function $z(.)$. Note that at most one multiplier can be non-zero and that

$$\frac{\partial W}{\partial z(h)} = \begin{cases} < 0 \Longrightarrow \lambda_0(h) > 0 \text{ and } \phi(h) = 0 \\ > 0 \Longrightarrow \lambda_1(h) > 0 \text{ and } \phi(h) = 1 \end{cases} \tag{15}$$

Now let $\bar{n}_A = 1 + n\omega$ be the post-immigration population of A and $\bar{n}_B = n(1-\omega)$ the post-immigration population of B.

$$\begin{aligned} \tfrac{\partial W}{\partial z(h)} &= \theta_A^* g'(\bar{h}_A)\tfrac{h-\bar{h}_A}{\bar{n}_A} - \theta_B^* g'(\bar{h}_B)\tfrac{h-\bar{h}_B}{\bar{n}_B} + (1-\theta)\left[g(\bar{h}_A) - g(\bar{h}_B) + \ln(1-t)\right] \\ &= \alpha(1-\theta)\left[\ln\tfrac{\bar{h}_A}{\bar{h}_B} + \tfrac{\ln(1-t)}{\alpha} + 1 - m + \left(\tfrac{m}{\bar{h}_A} - \tfrac{1}{\bar{h}_B}\right)h\right]. \end{aligned}, \tag{16}$$

where the second equality follows because $g(h) = \alpha \ln h$ and $g'(h) = \alpha/h$, and where

$$m \equiv \frac{\theta + (1-\theta)Z}{(1-\theta)(1+Z)} \qquad \text{and} \qquad Z \equiv \int z(h)dh. \tag{17}$$

Proposition 3 *The optimal policy has the following properties:*

1. Whenever immigration is positive, it is always skill biased,

2. For θ sufficiently close to unity, the policy is of the form (12), and

3. For $\theta < \frac{1}{2}$, the policy is of the form (10).

Proof. The RHS of the FOC (16) at the status-quo point at which $Z = 0$, i.e. the point at which there is no migration, is illustrated in Figure 3 in which vertical axis measures the marginal benefit of moving a person of type h from B to A. That benefit depends linearly on the migrant's level of h: I.e., in (16), $\frac{\partial W}{\partial z(h)}$ is linear in h, which immediately shows that either (10) or (12) must hold, though possibly with $\tilde{h} = 0$ or $\tilde{h} = +\infty$. The larger is m (which, in turn, is more likely when θ is large) the more likely the slope will be positive and that the more likely that a brain drain is optimal. Conversely when θ is small, the more likely that low-h migrating is optimal, i.e., (10). ■

We also note that for some θ's satisfying $\frac{1}{2} < \theta < 1$, the optimal policy may involve no migration – see Figure 3 and the top panel of Figure 4.

A number of alternative assumptions would overturn the optimality of the bang-bang policy. First, conditional on h, people may differ in their moving costs. Second, there may be h-specific congestion costs such as were imposed by the American Medical Association some decades ago. Third, the production function could entail diminishing returns to each skill if, for instance, $Y = G\left(\bar{h}\right)\left(\int n\left(h\right)^{\phi} dh\right)^{1/\phi}$. Fourth, physical-capital limits in the rich world would, coupled with diminishing returns to H, limit the optimal immigration flow.[7] We do, however, present a two-skill extension

[7]Nevertheless, we do not believe that migration is deterred largely by moving costs or by diminishing returns to labor – the latter would provide additional reasons why immigrants' wages in the host countries would be lower. Rather, it would appear that migrants are eager to move but that they are shut out by the restrictive immigration policies of the rich countries. In fact, there seems to be a vast excess demand for migration. Such excess demand prompted the U.S. to built a fence on the Mexican border and to ration migrants it by lottery. People are willing to risk imprisonment

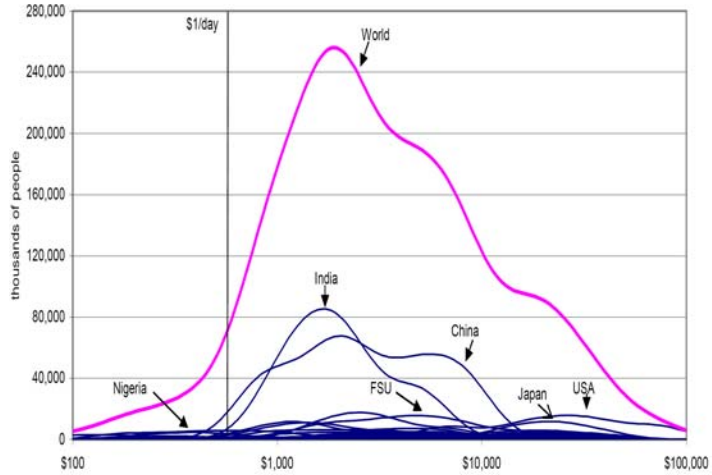


Figure 1: The world income distribution in 2000

that features diminishing returns to skill of each type and show that our conclusions survive this extension.

3 Calibration and simulation

We now wish to illustrate the optimal policy for all θ, and for realistic F_A and F_B. We choose "country A" to be the OECD which we shall think of as the developed world. "Country B" will then be the rest of the world. Sala-i-Martin (2006) reports the world distribution in the year 2000, and how it comprises the distributions of income in individual countries. We reproduce these distributions in Figure 1, which shows them to be roughly log-normal in form.

---

and deportation and estimates are that there are at least 10 million illegal immigrants in the us. Similar conditions exist in Europe; Italian authorities intercept boats from Albania, and the French legal system imposes fines and even jail on those who help illegal immigrants, North Africans risk their lives trying to cross the Mediterranean to the south of Europe.. A further indication that there is excess demand to migrate is the proposal that visas be auctioned to the highest bidders.

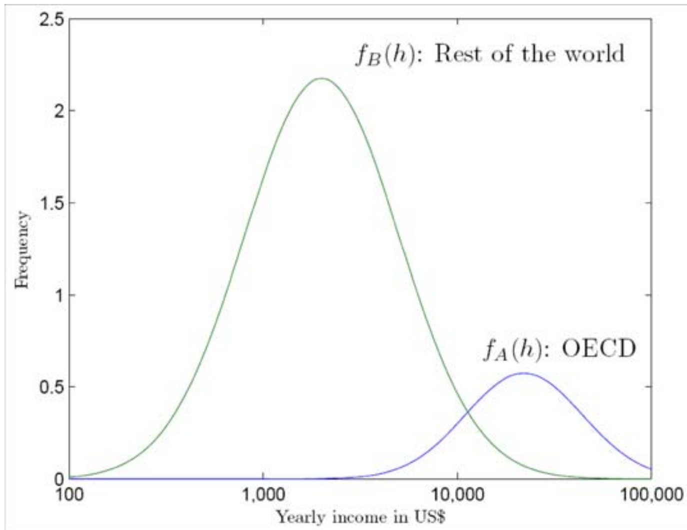


Figure 2: CALIBRATED DISTRIBUTIONS OF A AND B

We observe the distribution of income y for each citizen, which we approximate as follows by a log-normal distribution:

$$\mu_{OECD}(\log y) = \ln 20,000, \qquad \sigma_{OECD}(\log y) = \ln 2 \tag{18}$$
$$\mu_{Rest}(\log y) = \ln 2,000, \qquad \sigma_{Rest}(\log y) = \ln 2.5$$

These are portrayed in Figure 2. We set

$$n = 10, \quad \alpha = 0.35$$

the latter following Lucas (1990). The following equation identifies $\bar{h}$:

$$E(y) = \exp(\mu + \sigma^2/2) = G(\bar{h})E(h) = \bar{h}^{\alpha+1} \Longrightarrow \bar{h}_A = \exp\left(\frac{\mu + \sigma^2/2}{1+\alpha}\right)$$

To infer the human capital, h, of a citizen with income y, we invert the equation $y = G(\bar{h})h = \bar{h}^{\alpha}h$ to get $h = y\bar{h}^{-\alpha}$, i.e., $\ln h = \ln y - \alpha \ln \bar{h}$

Rendon and Cuecuecha (2007) estimate that the out-of pocket moving cost of a Mexican to the U.S. is about US$ 550 in 1992 dollars and Amuedo-Dorantes and Bansak (2007) report the slightly higher estimate of $655-$831. As a fraction of a

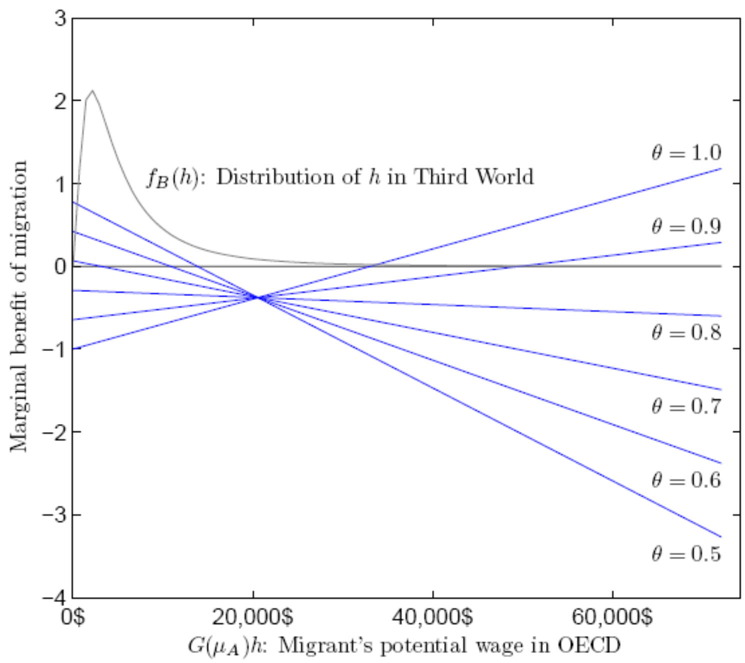


Figure 3: THE FIRST-ORDER CONDITION AT THE STATUS QUO

Mexican's lifetime U.S. earnings this is negligible. In our first calibration, then, we assume that moving costs are zero.

Figure 3 plots the RHS of the FOC (16) at the status-quo point at which $Z = 0$, i.e. the point at which there is no migration. The vertical axis measures the marginal benefit of allowing a migrant in; the benefit depends on the migrant's level of h. The figure shows that for some values of θ – say around $\theta = 0.8$, the marginal benefit of migration is negative at all levels of migration. Because the first-order condition is linear in h, the gain to migrating a worker of type h is either decreasing or increasing in h depending on the sign of $\frac{\theta}{\bar{h}_A} - \frac{1-\theta}{\bar{h}_B}$.

The brain-drain region $\theta \in [\theta_{\mathrm{BD}}, 1]$ — The slope of the FOC changes sign at

$$\theta_{\mathrm{BD}} = \frac{1}{1 + \bar{h}_B/\bar{h}_A} \approx 0.88 \text{ in the calibrated example,}$$

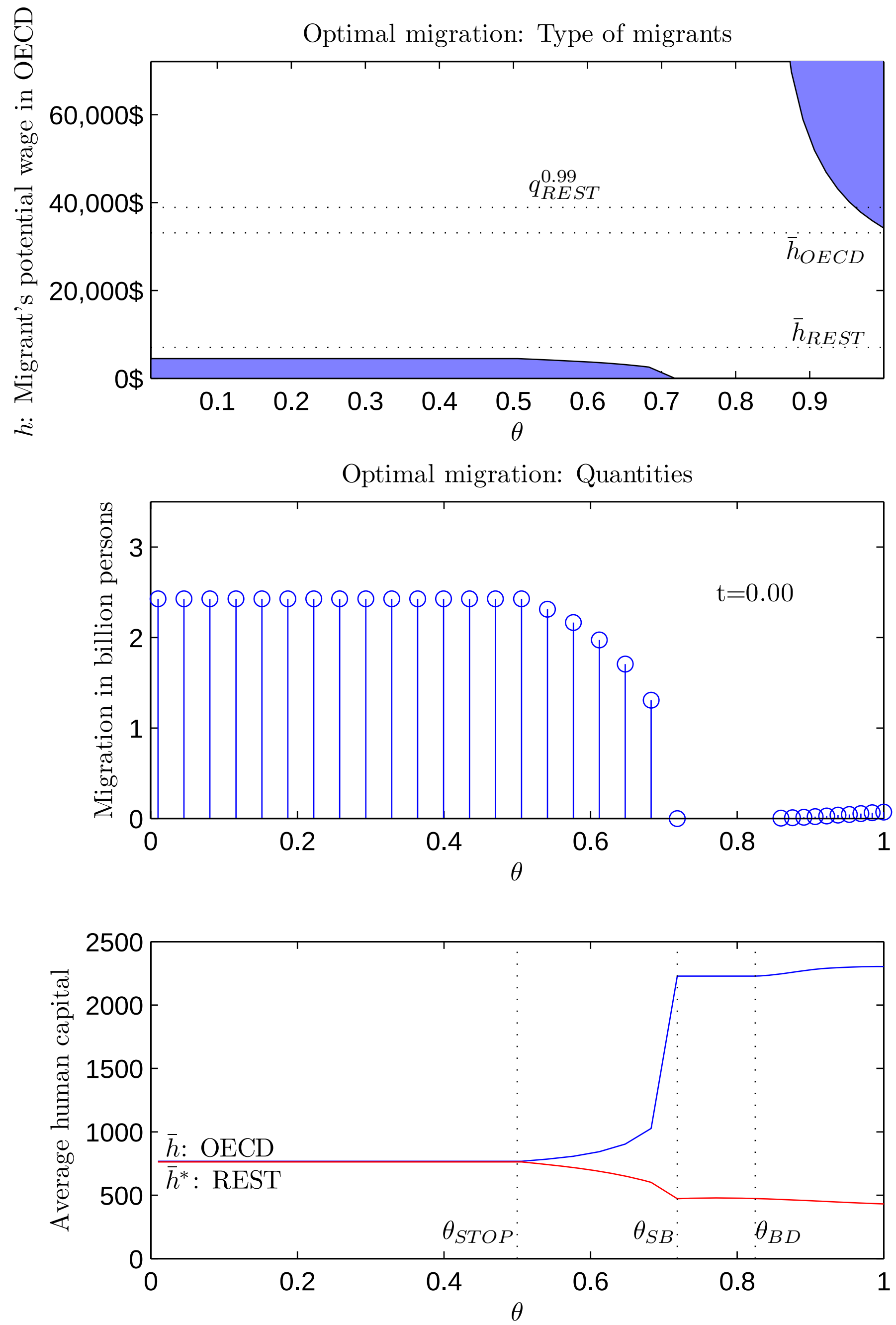


Figure 4: THE OPTIMAL POLICY

where BD is for brain-drain: If (as is the case in this calibration) F_B has unbounded support, then for any $\theta > \theta_{BD}$, some very smart B-people should go to A, and there will be a brain drain. This will remain so even when we add migration costs.

The skim-the-bottom region of $\theta \in [0, \theta_{\text{SB}}]$.—There is another threshold, call it θ_{SB}, below which country A will only receive low-h types. Suppose that the lowest level of h in the support of F_B is zero (as, again, is the case in the calibration). Then as shown in Figure 4,

$$\theta_{\text{SB}} = \frac{1+C}{2+C} \approx 0.71$$

where $C = \ln(\bar{h}_A/\bar{h}_B)$. This is also apparent in Figure 3 in which the FOC for $\theta = 0.7$ barely crosses the zero axis in the neighborhood of zero.

The inaction region $\theta \in (\theta_{\text{SB}}, \theta_{\text{BD}})$.—In this region, efficiency losses stemming from the mixing overwhelm the redistributive gains. It is not worth moving the high-skilled B-natives to A because, while this would raise $G\left(\bar{h}_A\right)$, it would reduce $G\left(\bar{h}_B\right)$ by too much. At these intermediate θs, it is not that the Planner does not value the A-natives; he simply values the B-natives too much to allow a brain drain from B to occur.

The optimal policy.—The optimal policy is described in Figure 4. The horizontal axis in each panel measures θ, the weight that the planner assigns to the rich country's natives. The top panel describes the skill of the movers. The shaded area is the set of people who can move under the optimal policy. Rather than measure the level of a migrant's h, however, the vertical axis in the top panel measures $G(\mu_A)\,h$, the wage that a migrant of type h would earn in country A assuming that no one else was allowed to move so that average skills in A were at their pre-migration level of μ_A. For $\theta \leq 0.72$, the unskilled B-natives migrate to A, and for $\theta \geq 0.89$, the skilled B-natives migrate. In between, migration is zero.

The middle panel of Figure 4 plots the numbers optimally moving at various levels of θ. At the egalitarian weight $\theta = 1/2$ the optimal number migrating is 2.3 billion, and then as θ rises the number declines at an increasing rate, and reaches zero when $\theta = 0.72$. When θ reaches 0.88, skimming from the top starts, but the numbers are seen to be small, simply because the developing countries have few highly-skilled people. The bottom panel Figure 4 shows that when evaluated at $t = 0$, the incentive-compatibility constraint (7) is binding for $\theta \leq 1/2$. That is because, for $\theta < 1/2$, the planner would like the residents of B to enjoy higher utility, and this requires $\bar{h}_B > \bar{h}_A$, but in that case migrants would not wish to move from B to A.

3.1 Broader migration costs

One can argue that after crossing the border, an immigrant needs to look for a job and a place to live, and so on, and that these additional costs should be included in t. We now pursue this possibility. One can estimate such costs from wage differentials net of cost-of-living differentials within, say, the U.S., within which mobility is unrestricted. Gemici (2007) estimates an individual's cost of moving from one U.S. census region to another[8] to be about seven percent of lifetime income. If we add the Rendon-Cuecuecha or the Amuedo-Dorantes and Bansak estimates to Gemici's estimate, we end up with at most 7.5 percent of U.S. lifetime income.

The counterpart of Figure 4 when $t = 0.075$ is Figure 5. We find that at all levels of θ the effect of $t > 0$ is to reduce the amount of optimal migration. At the egalitarian weight of of $\theta = 0.5$, the number of people optimally moving drops from 2.5 billion to 1.8 billion, which is still two orders of magnitude higher than the estimates of current migration levels of 100-200 million.[9] In the bottom panel we see that (7) starts to

[8] The U.S. is divided into 9 such regions so that the average region contains six states. Gemici controls for unobserved attributes of regional locations so that the effect of cost-of-living differentials on mobility costs would be reflected in the influence of these attributes. She estimates the residual pecuniary costs of moving between a pair of census divisions to be about $19,000 (in 1982 dollars), which is seven percent of the lifetime income ($240,000) of an average white male with 9 years of regional tenure.

[9] Freedberg and Hunt (1995) report that all but 100 million of the world's 6 billion people, i.e., all but 1.7 percent, live in the country of their birth.

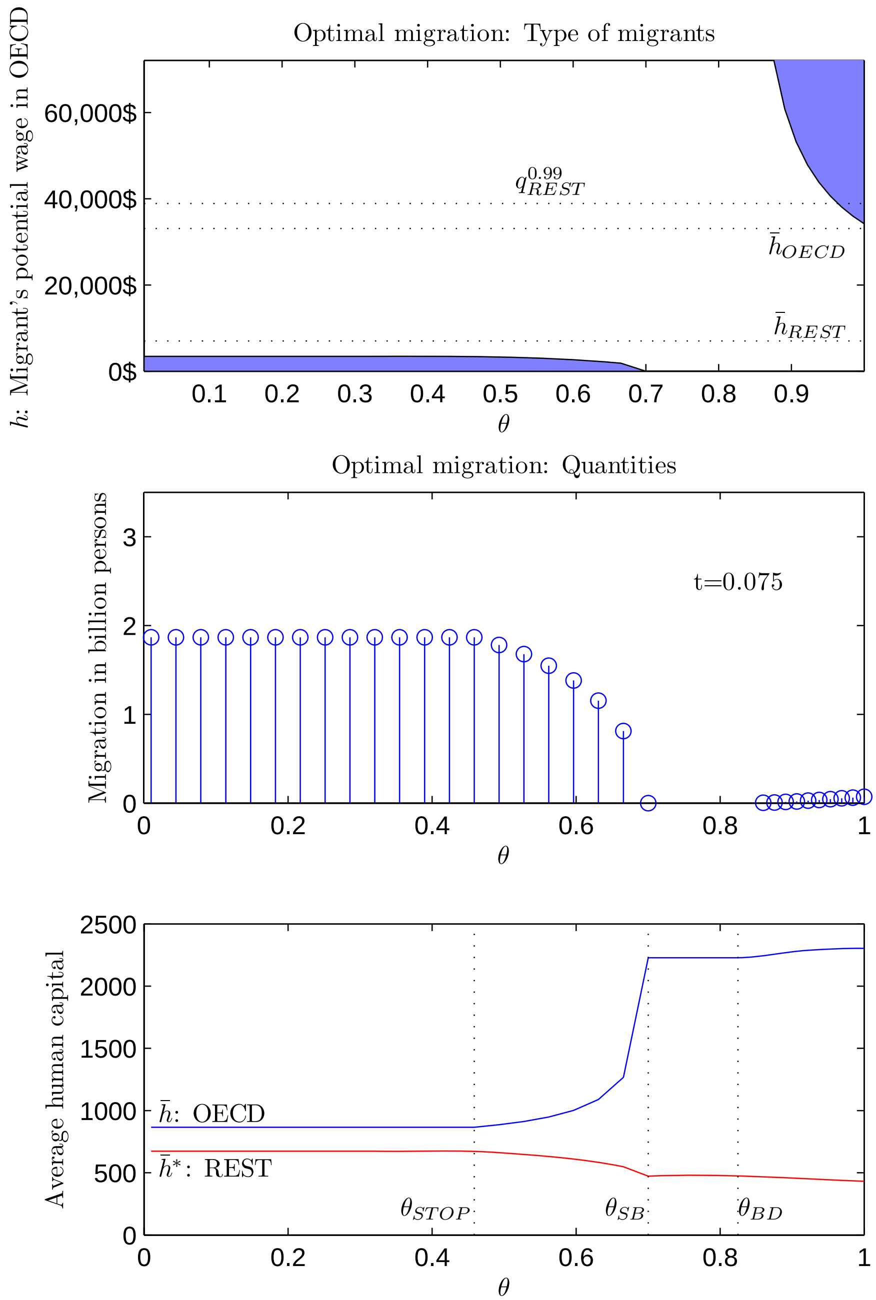


Figure 5: Optimal migration when migration costs are 7.5% of lifetime income in the host country

bind when $\bar{h}_A/\bar{h}_B = (1-t)^{-1/\alpha} = 1.25$, which binds for $\theta \leq 0.46$. In neither case, therefore does the incentive constraint impede the Planner from implementing the egalitarian allocation.

Now the region of θ's for which zero migration is optimal grows only slightly, from $[0.72, 0.88]$ to $[0.70, 0.88]$, and it would now take a value of θ of almost 0.7 to justify the currently observed migration levels in terms of numbers alone. But such an inference would not be correct because of the composition of the migrants which still come mainly from the right tail of the h distribution in the source countries – what the advanced countries have in place at the moment may not be a brain drain policy, but it certainly is not a 'skim from the bottom policy' that is optimal when $\theta < 0.70$. A significant number of rich countries are evidently behaving as if the Planner weighed their citizens with a value of θ in excess of 0.88.[10] This is not at all surprising: in

The International Organization for Migration estimates that there are 191 million transnational migrants worldwide comprising 3% of the global population. See http://www.iom.int/jahia/page254.html.

[10]The U.S. today follows a mixture of skill-biased policies and skill-neutral policies based on four principles: The reunification of families, the admission of immigrants with needed skills, the protection of refugees, and the diversity of admissions by country of origin. While special legislation now allows for special consideration for medical professionals for example, the majority of legal immigrants enter the US through the family-reunification program. While Canadian policy also allows immigration based on family reunification, preferences stress skills and youth: During 1990–2002, 65 per cent of permanent immigrants to the United States were admitted under family preferences. In Canada, the equivalent proportion was 34 per cent (International Migration and Development: Regional Factsheet, The Americas, http://www.un.org/migration/presskit/factsheet_america.pdf). Similarly, Australia heavily emphasizes skills and youth in its preference system for immigrants. See for example http://www.workpermit.com/Australia/australia.htm.

as far as a rich country's migration policy is set unilaterally to maximize the utility of its native citizens, one would expect migration policy to follow the predictions of the model when $\theta = 1$, and this is indeed what seems to be happening, with world migration flows reaching only about 200 million people per year.

3.1.1 Lower values of α and TFP differentials

The value of $\alpha = 0.35$ underlies both the simulations, the same value that Lucas (1988) uses. Recent evidence suggests that external effects may be smaller (Acemoglu and Angrist 2000), or about zero (Ciccone and Peri 2006). What effects would a lower value of α have? The answer to this question depends on whether or not the model includes exogenous differentials in TFP

With no exogenous TFP differentials as in (1) which underlies the simulations, an absence of external effects removes all incentives to migrate. Thus at $\alpha = 0$, migration is zero. On the other hand, when $\alpha = 0.35$ optimal migration is positive, and it is plotted in the middle panels of Figures 4 and 5 for two different values of t. Denote by $N(\theta; 0.35)$ the levels of migration corresponding to $\alpha = 0.35$ and to a welfare weight θ. This quantity is plotted in the middle panels of Figures 4 and 5 for two different values of t. A linear approximation states that the number of migrants for other values α denoted by $N(\theta; \alpha)$ would be

$$N(\theta; \alpha) = \frac{\alpha}{0.35} N(\theta; 0.35) \tag{19}$$

so that, indeed, with no external effects optimal migration is zero, and there should be no need to restrict immigration motivated by higher wages in developed countries. Since this is not the case, the conclusion needs qualifying in three ways.

Recently France has also moved towards a skill biased immigration policy: see http://www.migrationpolicy.org/pubs/Backgrounder2_ Caponi (forthcoming [a]), however, finds that an empirical analog of $\phi(h)$, relation between skill and the probability of migration of Mexicans to the U.S. is U-shaped: The highest and lowest educated tend to migrate more than the middle educated.

First, a second type of external effect of $\bar{h}$ may operate through the utility function of the natives of the host country: A large influx of uneducated and unskilled immigrants who are not familiar with the host country's culture and mores could reduce the natives' utility directly. This externality would algebraically show up as a positive α. Therefore, even if production externalities are small, α need not be small.

The second qualification to (19), elaborated on Section 4.1, is the likely presence of exogenous TFP differentials. If we add such TFP differentials as in (3), then Proposition 2 suggests we should have more migration. Suppose we set $\alpha = 0$ so that output at location i is just $Y = S_i H_i$. Then since moving costs, t are proportional to h, it is optimal to move agent h from B to A if $S_A (1 - t) h. > S_B h$, i.e., iff

$$\frac{S_A}{S_B} > \frac{1}{1-t}. \tag{20}$$

Since this condition does not involve h, it is optimal to move either everyone or no one. Now the simulation depicted in Figure 4 assumes that $t = 0$, in this case if α were zero, everyone in country B should have moved to A. The second simulation depicted in Figure 5 assumes that $t = 0.075$, and in this case if α were zero, everyone in country B should move to A as long as $S_A/S_B > 1.081$. Therefore even modest TFP differentials would produce migration levels far in excess of those currently observed.

Third, as we shall show in Section 4.3 and Proposition 4 in particular, (20) says that in the one-input model there is no incentive to migrate if alpha is zero. In the two input case, even without external effects or TFP differences, there would be incentives for migration towards equating factor ratios and prices.

4 Three extensions

We now ask if our conclusions remain valid when we extend the model to incorporate three realistic features that have so far been left out. Each feature is likely to change the quantitative predictions, but they leave intact the conclusion that a welfare function that considers individuals alone (and not their country of origin) would imply that much larger flows of migrants are optimal.

4.1 Cross-country differences in TFP

Recent studies that differences in output per worker across countries can to a large extent be explained by TFP differences. Hall and Jones (1999) show that the TFP differences are correlated with output per worker, and that TFP ratios between the rich and poor countries can be of the order of 6 to 8. See also Klenow, Peter and Andres Rodriguez-Clare (1997). More recently, Klein and Ventura (2007) also focus on TFP differences to explain the gap in output per worker between OECD and non-OECD countries, as we do in the simulations that we report in Figures 3 and 4. Klein and Ventura come up with TFP ratios ranging from 1.5 to 4.

As we stressed in discussing condition (20), exogenous cross-country TFP differentials raise the optimal level of migration for any welfare function except at $\theta = 1$. To consider this more generally, we revert to the production function (3) but re-written in the slightly more general form

$$Y_i = S_i G\left(\bar{h}_i\right) H_i \tag{21}$$

where S_i is TFP in country $i = A, B$. Then (13) becomes

$$W \equiv \theta_A^* \left(g_A + s_A\right) + \theta_B^* \left(g_B + s_B - \ln\left[1 - t\right]\right) \tag{22}$$

where $s_i = \ln S_i$. Instead (16), the first order condition is now

$$\begin{aligned} \frac{\partial W}{\partial z(h)} &= \theta_A^* g'(\bar{h}_A)\frac{h-\bar{h}_A}{\bar{n}_A} - \theta_B^* g'(\bar{h}_B)\frac{h-\bar{h}_B}{\bar{n}_B} + (1-\theta)\left[g(\bar{h}_A) + s_A + \ln(1-t) - g(\bar{h}_B) - s_B\right] \\ &= \alpha(1-\theta)\left[\ln\frac{\bar{h}_A}{\bar{h}_B} + \frac{\ln(1-t)}{\alpha} + \frac{s_A - s_B}{\alpha} + 1 - m + \left(\frac{m}{\bar{h}_A} - \frac{1}{\bar{h}_B}\right)h\right]. \end{aligned} \tag{23}$$

The first observation on (23) is that as $\alpha \to 0$, it becomes equivalent to (20). If we assume that productivity is higher in the host country A, then $s_A > s_B$ and the derivative $\frac{\partial W}{\partial z(h)}$ is now larger than it was in (16) where the TFP differences were absent. Therefore larger is the TFP difference $s_A - s_B$, the more likely it is that the slope will be positive at all levels of h.

Second, when $\alpha > 0$, the displacements are the same at all levels of h, but they are larger when θ is small. In terms of Figure 3 which plots the RHS of the FOC (16), the RHS of (23) entails an upward displacement of the lines by the amount $(1-\theta)(s_A - s_B)$. Therefore the addition of TFP differences raises the optimal migration for all welfare functions except the one in which $\theta = 1$. Since TFP is exogenous, the cost to migration that the planner faces are the same as they were in the absence of TFP differences – the foregone earnings costs is the efficiency loss entailed by the reduction in $G\left(\bar{h}_A\right)$. But the benefits are now larger, and so more migration would be optimal.

To see more precisely the effect that exogenous TFP variation across countries makes in the context of $\alpha = 0.35$, we shall use the mid-point of Klein and Ventura's estimates of S_A/S_B which ranging from 1.5 to 4, so that the midpoint is 2.75, and its logarithm is 1.01. Figure 6 shows how Figure 3 changes as a result of this adjustment. There is absolutely no effect on the optimal policy when $\theta = 1$. But when $\theta = 0.7$, the effect is to raise the identity of the marginal immigrant from a potential wage of \$3,000 to one of \$18,000, and when $\theta = 0.5$, the identity of the marginal immigrant rises from a potential wage of \$16,000 to one of \$22,000. These effects on the numbers of immigrants are especially large in the first case, as can be seen from the rest-of-the-world income distribution in Figure 2.

4.2 Dynamics and investment in h

To keep things simple we assume that within countries A and B agents are homogeneous, with skills μ_A and μ_B respectively, once again we assume that $\mu_A > \mu_B$.

Accumulation technology.—The technology for accumulating human capital is the same in both countries. The fraction of the period-0 time spent working is u_A and u_B, respectively. The remaining time is spent training, with the resulting human capital

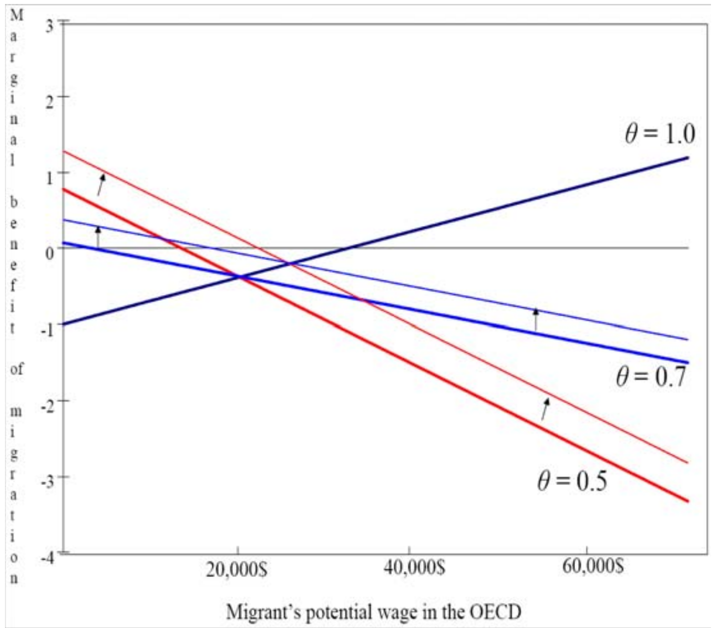


Figure 6: THE FOC WITH THE KLEIN-VENTURA TFP ADJUSTMENT

levels being, as in Lucas (1988),

$$h_A = \delta(1 - u_A)\mu_A \qquad \text{and} \qquad h_B = \delta(1 - u_B)\mu_B$$

in period one.

Equal migration probabilities.—Let x be the probability that *any* B-native will be allowed to move to A. Within-country homogeneity implies that (5) and (6) become

$$\bar{h}_A = \frac{h_A + xnh_B}{1 + xn}, \qquad \text{and} \qquad \bar{h}_B = h_B \tag{24}$$

respectively.

Preferences.—Lifetime utility is $U(c_0) + \rho U(c_1)$. There is no borrowing and lending, so that each worker simply consumes his wages. We also assume that immigrants cannot send money back to their own country.

The investment decision in A.—Each resident of A starts life with μ_A. In the first period there are no immigrants and therefore the first-period efficiency wage is $G(\mu_A)$. The only variable that the immigration policy x affects is the second-period

$\bar{h}_A$, and through it, each A-native's second period wage, $G\left(\bar{h}_A\right)$. It is convenient to index the decision problems by $\bar{h}_A$. Therefore a native of A maximizes his lifetime utility

$$v\left(\bar{h}_A\right) \equiv \max_{c_0, c_1, u} \left\{U\left(c_0\right) + \rho U\left(c_1\right)\right\}$$

subject to

$$c_0 = uG\left(\mu_A\right)\mu_A, \qquad \text{and} \qquad c_1 = \delta\left(1-u\right)G\left(\bar{h}_A\right)\mu_A. \tag{25}$$

He takes the wages in both periods as given, because these depend only on the skill composition of the population at large. Then the A-native's problem becomes

$$v\left(\bar{h}_A\right) = \max_u \left\{U\left(uG\left(\mu_A\right)\mu_A\right) + \rho U\left(\delta\left(1-u\right)G\left(\bar{h}_A\right)\mu_A\right)\right\}.$$

We once again assume (9). Then the problem boils down to maximizing $\ln u + \rho \ln\left(1-u\right)$. The first-order condition is $\frac{1}{u} - \frac{\rho}{1-u} = 0$, and it yields optimal investment

$$u_A = \frac{1}{1+\rho}. \tag{26}$$

which does not depend on the immigration policy x.

The investment decision in B.—The policy is assumed to be skill neutral, which means x is not influenced by the B-native's h. Thus, a B-native must choose his first-period investment before he knows if he will immigrate or not. Should he have to remain in B, his efficiency wage will be $G\left(\bar{h}_B\right)$, and if he is allowed to emigrate, his wage will be $G\left(\bar{h}_A\right)$. His lifetime utility is

$$v\left(\bar{h}_A, \bar{h}_B, x\right) = \max_u \left\{U\left(c_0\right) + \rho\left[xU\left(c_{1A}\right) + \left(1-x\right)U\left(c_{1B}\right)\right]\right\}$$

subject to:

$$\begin{aligned} c_0 &= uG\left(\mu_B\right)\mu_B \\ c_{1A} &= \delta\left(1-u\right)G\left(\bar{h}_A\right)\mu_B \\ c_{1B} &= \delta\left(1-u\right)G\left(\bar{h}_B\right)\mu_B \end{aligned} \tag{27}$$

where c_{1A} is B's consumption if he wins the lottery and moves to A and c_{1B} is his consumption if he loses and stays in B. The worker takes $\bar{h}^*$ and $\bar{h}$ as given. Once again, under (9),the investment rate is the same as in A, i.e.,

$$u_B = \frac{1}{1+\rho} = u_A. \tag{28}$$

Planner's choice of x.—The planner is a Stackelberg leader. He announces a migration policy at the outset, and carries it out at the end of the first period. Given the policy, agents invest in h at $t = 0$. The planner then chooses x to solve the problem

$$\max_x \left\{ \theta v\left(\bar{h}_A\right) + (1-\theta)\, nv\left(\bar{h}_A, \bar{h}_B, x\right) \right\}$$

subject to (24) and (28). The first-period utilities do not depend on x, and the discount factor drops out. The problem reduces to

$$\max_x \left\{ \theta U(c_1) + (1-\theta)\, n \left[xU(c_{1A}) + (1-x)\, U(c_{1B}) \right] \right\}.$$

Substituting from (9), (25), and (27), and letting $g = \ln G$,the problem boils down to $\max_x J$, where (writing $g_i \equiv g\left(\bar{h}_i\right)$ for $i \in \{A, B\}$)

$$J \equiv \theta g_A + (1-\theta)\, n \left(x g_A + (1-x)\, g_B \right) \tag{29}$$

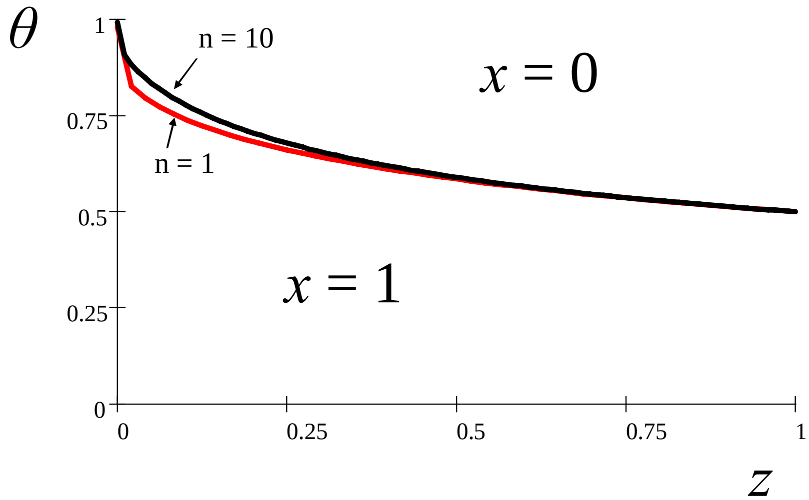
subject to (24). The following Lemma (proved in Appendix 2) shows that under some reasonable conditions the solution for x is bang-bang, i.e., either $x = 0$ or $x = 1$:

Lemma 2 *For $G(h) = h^\alpha$, and for any $\alpha > 0$, $n > 0$ and $\theta \in (0,1)$, x is either zero or one*

This result implies that the planner's problem cannot have an interior maximum. Rather, the planner's maximum is at a corner: Either $x = 0$ or $x = 1$.

Characterizing the solution for x.—Define the initial, date-zero productivity of a B-native relative to that of an A-native by

$$\text{Relative backwardness} \equiv z = \frac{\mu_B}{\mu_A} < 1. \tag{30}$$



The bang-bang policy

Figure 7: Optimal policy when h is engogenous

The Appendix also shows that the optimal policy is

$$x = \begin{cases} 0 & \text{if } \theta > \frac{n\ln\left(\frac{z^{-1}+n}{1+n}\right)}{\ln z^{-1}+(n-1)\ln\left(\frac{z^{-1}+n}{1+n}\right)}, \\ 1 & \text{otherwise} \end{cases} \tag{31}$$

We plot the indifference locus for $n = 1$ and $n = 10$ in Figure 7.

The planner is more inclined to a policy of immigration if B is poor, and if B is large (large n), though the latter is not a quantitatively important consideration. In the plot, the action $x = 0$ is preferred in the north east quadrant (NE) and $x = 1$ is preferred in the south east quadrant (SW). So, the planner chooses maximal immigration if he cares enough for B (low value of θ), and if B is poor enough (low value of z) and if B is large (high values of n). Empirically, $z = 0.1$ is a good approximation to the average non-U.S. income, which means, for $n = 1$, (since in fact x is close to zero) we may infer that $\theta \geq 3/4$. For example, at $z = 0.1$, the critical θ needed to switch to a zero-migration regime is 0.78, and at $z = 0.5$ the critical θ

then drops to 0.59

Freedberg and Hunt's (1995) evidence tells us that we are, effectively, in the $x = 0$ region. But, since z must be rather small – say 1/10 – the action $x = 0$ is optimal only if θ is at least 0.8. Thus, from the homogeneous-residents model the following conclusions emerge:

1. Investment in human capital leaves intact the main conclusion that the policy outcome we now have is incompatible with even approximately equal weights in the social welfare function.

2. The model of this subsection assumes that within each country agents were homogeneous before and after training. If the planner could and did redistribute incomes within economies, he would still want to use migration to redistribute *across* economies.

4.3 Two skills, and no external effects

In this section we show that our main conclusions are robust to the introduction of a second skill in the production function. They also are robust to the removal of external effects and their replacement with exogenous productivity differences that do not depend on the within-country composition of skills. In particular, the following two conclusions (see Figure 4),

1. When $\theta = 1/2$, the Planner equalizes $\bar{h}$ in the two countries, and

2. When $\theta = 1$, the Planner pursues a 'brain-drain' policy,

remain valid when restated in their two-skill version as in Propositions 2 and 3.

Let production now depend on two homogeneous groups of workers: the 'skilled,' the number of which is s, and the 'unskilled,' the number of which is u. Output is

$$Y = xG\left(\frac{s}{u}\right) s^{\beta} u^{1-\beta}, \tag{32}$$

where $G\left(\frac{s}{u}\right)$ is an external effect operating through the ratio of skilled to unskilled workers, and where G is an increasing function. The parameter x is introduced to proxy for all other country-specific variables that may affect the productivity of the two factors. When external effects are absent $G \equiv 1$.

Let $r = s/u$. The wage of a skilled worker in that country then is

$$w_s = \beta r^{\beta-1} x G(r), \tag{33}$$

and the wage of an unskilled worker is

$$w_u = (1-\beta) r^{\beta} x G(r). \tag{34}$$

In general, x and s depend on the location A or B. We assume that $x_A > x_B$ so that other things the same the Planner wants to move people from B to A.

Let s' and u' denote the number of skilled and unskilled that move from B to A. Post-migration wages are for each country then given by (33) and (34), but with the appropriate r substituted in, namely

$$r_A = \frac{s_A + s'}{u_A + u'} \quad \text{and} \quad r_B = \frac{s_B - s'}{u_B - u'}.$$

The Welfare criterion.—As before, let θ be the weight on country A natives. Assume moving costs are zero. The Planner's criterion then is

$$W = \theta\left(s_A U\left(w_s^A\right) + u_A U\left(w_u^A\right)\right) + (1-\theta)\left\{\begin{array}{c} s' U\left(w_s^A\right) + u' U\left(w_u^A\right) + \\ \left(s_B - s'\right) U\left(w_s^B\right) + \left(u_B - u'\right) U\left(w_u^B\right) \end{array}\right\}. \tag{35}$$

The incentive-compatibility constraint.—Each factor can flow from B to A only if its wages in A exceed those in B. Instead of (7), then, we now have two IC constraints:

$$\left(w_s^A - w_s^B\right) s' \geq 0, \text{ and } \left(w_u^A - w_u^B\right) u' \geq 0. \tag{36}$$

The Planner's tradeoff.—As before, the Planner faces a tradeoff between efficiency and utility. But because of the complementarity between s and u, (in contrast to Proposition 1) output is now lo longer largest under complete segregation. If x_A were to equal x_B, because (32) is homogeneous of degree 1 in (s, u), output would be at a maximum as long as $r_A = r_B$, and the distribution of activity between A and B would not matter. But if $x_A > x_B$, the world would produce the most output if everyone were to be moved to location A. The Planner's question then is how far the Planner can shift domestic factor ratios without reducing too much the consumption of one or the other group of A-natives.

As before, we shall assume that

$$G(r) = r^{\alpha} \quad \text{and } \alpha + \beta < 1, \tag{37}$$

so that $w_s = \beta r^{\alpha+\beta-1} x$, and $w_u = (1-\beta) r^{\alpha+\beta} x$. Then, skilled immigration from B to A is incentive compatible as long as

$$r_B > \left(\frac{x_B}{x_A}\right)^{1/[1-(\alpha+\beta)]} r_A \tag{38}$$

and unskilled immigration from B to A is feasible as long as, i.e.,

$$r_B < \left(\frac{x_A}{x_B}\right)^{1/(\alpha+\beta)} r_A. \tag{39}$$

Proposition 4 *If $\theta = 1/2$, and if $x_A = x_B$, the Planner welfare criterion is maximized by immigration flows that equate factor ratios*

$$\frac{s_B - s'}{u_B - u'} = \frac{s_A + s'}{u_A + u'}.$$

This result, proved in the Appendix, parallels the one-skill result (see Figure 4) that when $t = 0$ and $\theta = 1/2$, the Planner equalizes $\bar{h}$ in the two countries. It also assumes that (9) holds. Thus if the Planner's preferences were egalitarian, immigration flows would be much larger than they are in practice.

In the Appendix we also prove that when $\theta = 1$, a 'brain-drain' policy is optimal, just as it was in the one-skill case.

Proposition 5 *If $\theta = 1$ and if*

$$r_A < \frac{\alpha + \beta}{1 - (\alpha + \beta)}, \tag{40}$$

the optimal policy is to allow only the skilled emigrate from B to A until $w_s^A = w_s^B$.

The converse is also true: If the inequality in (40) is reversed, then it is optimal for only the unskilled to emigrate from B to A until $w_u^A = w_u^B$.

Since the nature of the policy depends on whether (40) holds or not, we need evidence on the parameters. Condition (40) pertains to the pre-migration relation between factor endowments. Taking the U.S. evidence as an indication, (40) is likely to hold for in fact. If we consider the skilled as college educated labor, then in the US roughly half the working population now holds a college degree, so that $r_A \approx 1$. On the other hand, the wage bill of the skilled is roughly twice that of the unskilled. In our context this means that $\beta \approx 2/3$. Therefore even if $\alpha = 0$ the RHS of (40) is at least 2. Therefore (40) is the relevant case empirically.

Proposition 4 is illustrated in Figure 8. The planner's FOCs both hold along that dashed line with its intercept at $r_A = \frac{\alpha+\beta}{1-(\alpha+\beta)}$.(See (41) and (42) of the Appendix). If he did not have to worry about (36), the Planner would end up on the dashed line. The location of that line does not depend on x_A or x_B since the Planner cares only about the A-natives.

Incentive compatibility fails in the shaded areas in Figure 8. Now (38) is violated in the bottom-right shaded area, and (39) is violated in the top-right shaded area. If $x_A = x_B$, both lines would coincide with the 45^0 line. The figure assumes that $x_A > x_B$.

Equation (40) holds to the left of the dashed line, and we have shown that this is the empirically relevant starting point. But there is a second consideration: Because

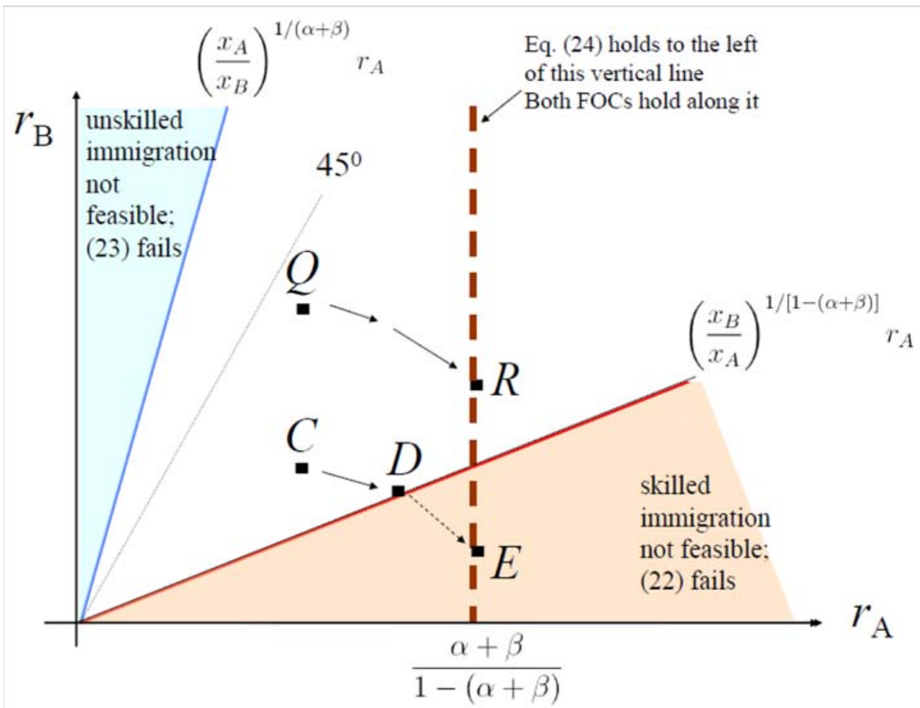


Figure 8: Illustration of Proposition 3

country A is skill abundant, we have $r_A > r_B$ which means that we start somewhere to the right of the 45^0 line, and as the skilled move from A to B, we move southeast. Now there are two subcases, depending on how high the initial r_B is.

1. If country B is fairly well endowed with skilled labor and if r_B is fairly high so that the starting point is Q, say, then, the Planner's optimum is to move to the unconstrained interior maximum point R.

2. If r_B is low, however, far below the 45^0 line, then the starting point is a point such as C. At this point w_S^B is closer to w_S^A than in the first case, and that means that less immigration is feasible before (36) is violated (which it would be if the planner were to continue to an interior point such as E). Therefore less immigration of the skilled can occur in this case.

Whether the move is a large one from Q to R, or a smaller one from C to D, the empirically relevant optimal policy involves an influx of the skilled factor from the poor world, i.e., a brain drain. Thus when we use empirically relevant parameter values, this result parallels our previous result (see Figure 4) that in the one-skill case, when $\theta = 1$, the Planner should pursue a brain-drain policy.

Therefore the main implications of our one skill analysis survive the model's extension to two skills. In particular, current immigration policies still resemble far more the $\theta = 1$ case than they do the egalitarian case of $\theta = 1/2$.

5 Conclusion

We have studied the role that migration can play in redistributing income to the world's poor. We argued that there is an equity-efficiency tradeoff. In spite of that tradeoff, we found that for reasonable social weights on the rich and the poor, the extent of migration that would be optimal would appear to be much larger than the levels observed today.

Any policy conclusion drawn from this exercise must, however, be quite tentative, for several reasons. First, except in section 4.2, the model is static, yet the fraction of the world's population that is migrating appears to be rising – the stock of migrants has increased from 155 million in 1990 to 213 million today. The model is silent on why this shift may have taken place. Second, all versions of the model assume that migration is the only redistributive tool, excluding foreign aid, FDI. Third, the model allows no redistributive role for trade which can be a substitute for migration; we expect that free trade, under certain conditions, could reduce differences in factor prices and mitigate the need for migration. And, fourth, aside from foregone-earnings costs, the diminishing returns to migration operate entirely by reducing the human-capital externality. Overall, several reasonable modifications of the model may reduce the optimal migration flows, but the thrust of our conclusion appears to survive: World-welfare would be higher if migration flows from poor to rich countries were considerably larger than they are today.

6 Appendix

We prove Propositions 2 and 3 after some preliminary derivations. Since $\ln w_s^A - \ln w_s^B = \ln w_s^A / w_s^B$, and since $x_A = x_B$,

$$\ln w_s^A - \ln w_s^B = \ln \frac{G_A}{G_B} + (1-\beta) \ln \frac{(u_A + u')(s_B - s')}{(u_B - u')(s_A + s')},$$

and

$$\ln w_u^A - \ln w_u^B = \ln \frac{G_A}{G_B} - \beta \ln \frac{(u_A + u')(s_B - s')}{(u_B - u')(s_A + s')}.$$

The terms involving the $\ln x_i$ appear additively and drop out of the calculations. We can write the remaining part of the RHS of (35) as

$$\begin{aligned} W &= (\theta s_A + (1-\theta) s') (\ln \beta + (1-\beta) \ln (u_A + u') + (\beta - 1) \ln (s_A + s') + \ln G_A) \\ &+ (\theta u_A + (1-\theta) u') (\ln (1-\beta) - \beta \ln (u_A + u') + (\beta) \ln (s_A + s') + \ln G_A) \\ &+ (1-\theta) (s_B - s') (\ln \beta + (1-\beta) \ln (u_B - u') + (\beta - 1) \ln (s_B - s') \ + \ln G_B) \\ &+ (1-\theta) (u_B - u') (\ln (1-\beta) - \beta \ln (u_B - u') + (\beta) \ln (s_B - s') + \ln G_B). \end{aligned}$$

We maximize it with respect to s' and u' subject to (36). We do so in turn for $\theta = 1$ and $\theta = 1/2$.

Proof of Proposition 3 (the case $\theta = 1/2$).—In this case θ cancels from the FOCs which, after rearrangement, become

$$\begin{aligned} \frac{\partial W}{\partial s'} &= (1-\beta) \left(\ln \frac{(u_A + u')(s_B - s')}{(u_B - u')(s_A + s')} \right) + \ln \frac{G_A}{G_B} + \frac{G'_A}{G_A} \frac{s_A + s'}{u_A + u'} - \frac{G'_B}{G_B} \frac{s_B - s'}{u_B - u'} \\ &+ \beta \left(\frac{u_A + u'}{s_A + s'} - \frac{(u_B - u')}{s_B - s'} \right) + \frac{G'_A}{G_A} - \frac{G'_B}{G_B}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial W}{\partial u'} &= -\beta \left(\ln \frac{(u_A + u')(s_B - s')}{(u_B - u')(s_A + s')} \right) + \ln \frac{G_A}{G_B} + (1-\beta) \left(\frac{s_A + s'}{u_A + u'} - \frac{s_B - s'}{u_B - u'} \right) \\ &+ \frac{G'_B}{G_B} \frac{(s_B - s')^2}{(u_B - u')^2} - \frac{G'_A}{G_A} \frac{(s_A + s')^2}{(u_A + u')^2} + \frac{G'_B}{G_B} \frac{s_B - s'}{(u_B - u')} - \frac{G'_A}{G_A} \frac{s_A + s'}{u_A + u'}. \end{aligned}$$

Now if

$$\frac{s_B - s'}{u_B - u'} = \frac{s_A + s'}{u_A + u'} = \frac{s_A + s_B}{s_B + u_B},$$

$\frac{G'_A}{G_A} = \frac{G'_B}{G_B}$, and $\frac{\partial W}{\partial s'} = \frac{\partial W}{\partial u'} = 0$. So the first order conditions hold. Using (37) we can show that the Hessian of W is negative semi-definite along a ray. Thus the optimal immigration flows equalize factor ratios.

Proof of Proposition 4 (case $\theta = 1$).—In this case the welfare criterion reduces to $W = s_A \ln w_s^A + u_A \ln w_u^A$, or (dropping the A subscript from G),

$$\begin{aligned} W &= s_A \left(\ln \beta + (1-\beta) \ln (u_A + u') + (\beta - 1) \ln (s_A + s') + \ln G\right) \\ &\quad + u_A \left(\ln (1-\beta) - \beta \ln (u_A + u') + \beta \ln (s_A + s') + \ln G\right). \end{aligned}$$

The derivative with respect to s' is

$$\begin{aligned} \frac{\partial W}{\partial s'} &= \frac{-s_A}{s_A + s'} \left(1 - \beta - \frac{G'}{G} \frac{s_A + s'}{u_A + u'}\right) + \frac{u_A}{s_A + s'} \left(\beta + \frac{G'}{G} \frac{s_A + s'}{u_A + u'}\right) \\ &= \left(\frac{s_A}{s_A + s'}\right) \left[\beta \left(\frac{u_A}{s_A} - \frac{1-\beta}{\beta}\right) + \alpha \left(\frac{s_A + u_A}{s_A}\right)\right], \end{aligned} \tag{41}$$

and with respect to u' it is

$$\begin{aligned} \frac{\partial W}{\partial u'} &= \frac{1}{u_A + u'} \left[\left((1-\beta) s_A - s_A \frac{G'}{G} \left(\frac{s_A + s'}{u_A + u'}\right)\right) + \left(-\beta u_A - u_A \frac{G'}{G} \left(\frac{s_A + s'}{u_A + u'}\right)\right)\right] \\ &= \frac{-s_A}{u_A + u'} \left[\beta \left(\frac{u_A}{s_A} - \frac{1-\beta}{\beta}\right) + \alpha \left(\frac{s_A + u_A}{s_A}\right)\right]. \end{aligned} \tag{42}$$

Then the claim holds if $\beta \left(\frac{u_A}{s_A} - \frac{1-\beta}{\beta}\right) + \alpha \left(\frac{s_A + u_A}{s_A}\right) > 0$, which reads $\frac{\beta}{r_A} - 1 + \beta + \alpha + \frac{\alpha}{r_A} > 0$, i.e., (40).

APPENDIX 2: Proof of Lemma 2.

We start by noting that since neither $\bar{h}_B$ nor $u_A = u_B = 1/(1+\rho)$ depend on x,

$$\frac{dJ}{dx} = (\theta + [1-\theta] nx) \frac{dg_A}{dx} + (1-\theta) n (g_A - g_B), \tag{43}$$

and

$$\frac{d^2J}{dx^2} = (\theta + [1-\theta]\, nx)\, \frac{d^2 g_A}{dx^2} + 2\,(1-\theta)\, n \frac{dg_A}{dx}. \qquad (44)$$

If (43) does not hold for any $x \in (0,1)$, the solution must be at a corner and the claim is proved. Conversely, if (43) does hold at an interior point, we now show that (44) also does. Note that

$$\frac{d^2J}{dx^2} > 0 \Longleftrightarrow \frac{d^2 g_A/dx^2}{dg_A/dx} < -\frac{2\,(1-\theta)\, n}{\theta + (1-\theta)\, nx}. \qquad (45)$$

We shall now show that if (43) holds, then (45) also does. Since $G(h) = h^\alpha$,

$$\begin{aligned} g_A &= \alpha \ln \frac{\delta\,(1-u_A)\,\mu_A + nx\delta\,(1-u_B)\,\mu_B}{1+nx} = \alpha \ln \delta\,(1-u_A) + \alpha \ln \frac{h + nx\mu_B}{1+nx} \\ &= \text{constant} + \alpha\,(\ln(\mu_A + nx\mu_B) - \ln(1+nx)). \end{aligned}$$

Therefore in light of (30), $\frac{1}{\alpha}\frac{dg_A}{dx} = \frac{n}{z^{-1}+nx} - \frac{n}{1+nx} < 0$, and

$$\frac{1}{\alpha}\frac{d^2 g_A}{dx^2} = \frac{d}{dx}\left(\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}\right) = \left(-\left(\frac{n}{z^{-1}+nx}\right)^2 + \left(\frac{n}{1+nx}\right)^2\right) > 0.$$

Therefore

$$\begin{aligned} \frac{d^2 g_A/dx^2}{dg_A/dx} &= \frac{\left(-\left(\frac{n}{z^{-1}+nx}\right)^2 + \left(\frac{n}{1+nx}\right)^2\right)}{\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}} \\ &= \frac{\left[\left(\frac{n}{1+nx}\right) - \left(\frac{n}{z^{-1}+nx}\right)\right]\left[\left(\frac{n}{1+nx}\right) + \left(\frac{n}{z^{-1}+nx}\right)\right]}{\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}} \\ &= -\left[\left(\frac{n}{1+nx}\right) + \left(\frac{n}{z^{-1}+nx}\right)\right]. \end{aligned}$$

Note that $h_B = \delta\,(1-u_B)\,\mu_B$, and that $h_A = \frac{\delta(1-u_A)\mu_A + nx\delta(1-u_B)\mu_B}{1+nx}$. Then,

$$\begin{aligned} g - g^* &= \alpha \ln \delta\,(1-u_A) + \alpha\left(\ln\left(\frac{\mu_A + nx\mu_B}{(1+nx)}\right) - \alpha \ln(\delta\,(1-u_B)\,\mu_B)\right) \qquad (46) \\ &= \alpha \ln \delta\,(1-u_A) + \alpha \ln\left(\frac{\mu_A + nx\mu_B}{(1+nx)\,\delta\,(1-u_B)\,\mu_B}\right) \quad \text{so that} \end{aligned}$$

$$\begin{aligned} \frac{g_A - g_B}{dg_A/dx} &= \frac{\alpha \ln \delta\,(1-u_A) + \alpha \ln\left(\frac{\left(\frac{\mu_A}{\mu_B}+nx\right)}{(1+nx)\delta(1-u_B)}\right)}{\alpha\left(\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}\right)} = \frac{\alpha \ln \delta\,(1-u_A) + \alpha \ln\left(\frac{\left(z^{-1}+nx\right)}{(1+nx)\delta(1-u_B)}\right)}{\alpha\left(\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}\right)} \\ &= \frac{\ln\left(\frac{\left(z^{-1}+nx\right)}{(1+nx)}\right)}{\left(\frac{n}{z^{-1}+nx} - \frac{n}{1+nx}\right)}. \end{aligned}$$

Now from the definitions of h_A, h_B, and g,

$$\begin{aligned}
\frac{d^2 g_A/dx^2}{dg_A/dx} &= \frac{\left(-\left(\frac{n}{z^{-1}+nx}\right)^2+\left(\frac{n}{1+nx}\right)^2\right)}{\frac{n}{z^{-1}+nx}-\frac{n}{1+nx}} \\
&= \frac{\left[\left(\frac{n}{1+nx}\right)-\left(\frac{n}{z^{-1}+nx}\right)\right]\left[\left(\frac{n}{1+nx}\right)+\left(\frac{n}{z^{-1}+nx}\right)\right]}{\frac{n}{z^{-1}+nx}-\frac{n}{1+nx}} \\
&= -\left[\left(\frac{n}{1+nx}\right)+\left(\frac{n}{z^{-1}+nx}\right)\right].
\end{aligned}$$

So, for (45) to hold, one needs that

$$-\left[\left(\frac{n}{1+nx}\right)+\left(\frac{n}{z^{-1}+nx}\right)\right] < -\frac{2(1-\theta)n}{\theta+(1-\theta)nx},$$

or that

$$\frac{1}{2}\left[\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right] > \frac{(1-\theta)}{\theta+(1-\theta)nx} = \frac{1}{\frac{\theta}{1-\theta}+nx}.$$

Now (43) implies $\frac{\theta}{(1-\theta)} = -\left(n\left(\frac{g-g^*}{g'}\right)+nx\right)$, which is equivalent to

$$\frac{1}{2}\left[\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right] > \frac{1}{-\left(n\left(\frac{g-g^*}{g'}\right)+nx\right)+nx}$$

or, from (46), to

$$\frac{1}{2}\left[\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right] > \frac{1}{-n\left(\frac{g-g^*}{g'}\right)} = \frac{-1}{\frac{\ln\left(\frac{(z^{-1}+nx)}{(1+nx)}\right)}{\left(\frac{1}{z^{-1}+nx}-\frac{1}{1+nx}\right)}} = \frac{-\left(\frac{1}{z^{-1}+nx}-\frac{1}{1+nx}\right)}{\ln\left(\frac{(z^{-1}+nx)}{(1+nx)}\right)}.$$

Now, this condition can be re-written as

$$\begin{aligned}
\ln\left(\frac{(z^{-1}+nx)}{(1+nx)}\right) &> \frac{2\left(\frac{1}{1+nx}-\frac{1}{z^{-1}+nx}\right)}{\left[\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right]} = 2\frac{\left(\frac{1}{1+nx}\right)^2-\left(\frac{1}{z^{-1}+nx}\right)^2}{\left(\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right)^2} \\
\ln\left(\frac{1}{(1+nx)}\right)-\ln\left(\frac{1}{(z^{-1}+nx)}\right) &> \frac{2\left(\frac{1}{1+nx}-\frac{1}{z^{-1}+nx}\right)}{\left[\left(\frac{1}{1+nx}\right)+\left(\frac{1}{z^{-1}+nx}\right)\right]},
\end{aligned}$$

or, if we write $A = \frac{1}{(1+nx)}$ and $B = \frac{1}{(z^{-1}+nx)}$, to

$$\ln(A)-\ln(B) = \ln\left(\frac{A}{B}\right) > 2\left(\frac{A-B}{A+B}\right) = 2\frac{\frac{A}{B}-1}{\frac{A}{B}+1}.$$

That is, we need, for all $y \equiv A/B > 1$, that

$$\ln y > 2\frac{y-1}{y+1}. \tag{47}$$

The LHS and RHS of (47) are both zero at $y = 1$. Therefore it's enough to show that the first derivative of the LHS is more positive than the first-derivative of the RHS for all y. That is, it suffices to show that

$$\begin{aligned} \frac{1}{y} &> 2\left(\frac{1}{y+1} - \frac{y-1}{(y+1)^2}\right) = \frac{2}{y+1}\left(1 - \frac{y-1}{y+1}\right) \\ &= \frac{2}{y+1}\left(\frac{y+1-y+1}{y+1}\right) = \frac{4}{(y+1)^2}. \end{aligned}$$

So, we need to show that $(y+1)^2 > 4y$, or that $y^2 + 2y + 1 > 4y$. But $y^2 + 1 - 2y = (y-1)^2 > 0$. Therefore (45) holds. ■

Derivation of (31).—Consider the set of (θ, z) pairs at which $J(1) = J(0)$. We shall show that the boundary of the two regions in (31) is the indifference curve $\theta = \frac{n \ln\left(\frac{z^{-1}+n}{1+n}\right)}{\ln z^{-1} + (n-1)\ln\left(\frac{z^{-1}+n}{1+n}\right)}$. First note from (24), that $\bar{h}_A$ is proportional to $1 - u_A (= 1 - u_B)$, and so is h_B. Neither u_A nor u_B depends on x, and after being substituted into g_A and g_B they become additive constants. Therefore $J(1) = J(0)$ is equivalent to

$$g\left(\frac{\mu_A + n\mu_B}{1+n}\right)(\theta + (1-\theta)n) = \theta g(\mu_A) + (1-\theta)ng(\mu_B)$$

Since $g(s) = \alpha \ln s$, the parameter α cancels and upon exponentiating both sides we have

$$\left(\frac{\mu_A + n\mu_B}{1+n}\right)^{\theta + n(1-\theta)} = \mu_A^{\theta}\mu_B^{n(1-\theta)}$$

Now $\mu_A = z^{-1}\mu_B$, so that $\left(\frac{z^{-1}+n}{1+n}\right)^{\theta+(1-\theta)n} = z^{-\theta}$ and this then reduces to

$$\begin{aligned} \left(\frac{z^{-1}+n}{1+n}\right)^{\theta+(1-\theta)n} &= z^{-\theta}, \quad \text{i.e.,} \\ [\theta + (1-\theta)n]\ln\left(\frac{z^{-1}+n}{1+n}\right) &= -\theta \ln z, \quad \text{i.e.,} \\ \theta\left[(1-n)\ln\left(\frac{z^{-1}+n}{1+n}\right) + \ln z\right] + n\ln\left(\frac{z^{-1}+n}{1+n}\right) &= 0, \end{aligned}$$

i.e.,

$$\theta = \frac{n \ln\left(\frac{z^{-1}+n}{1+n}\right)}{\ln z^{-1} + (n-1)\ln\left(\frac{z^{-1}+n}{1+n}\right)}.$$

References

[1] Acemoglu, D. and J. Angrist. "How Large Are The External Returns to Education: Evidence from Compulsory Schooling Laws" *NBER Macro Annual* 15 (2000), 9-74.

[2] Amuedo-Dorantes, C. and C. Bansak. "Illegal Border Crossings Attempts Following Apprehension." San Diego State University, January 2007.

[3] Benhabib, J. "On the Political Economy of Immigration" *European Economic Review* 40, (December 1996), 1737-1743.

[4] Caponi, V. "Intergenerational Transmission of Abilities and Self Selection of Mexican Immigrants." *International Economic Review,* forthcoming,

[5] Caponi, V. "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the US?" *Annales d'Economie et de Statistique*, Forthcoming (a).

[6] Ciccone, A. and G. Peri. "Identifying Human-Capital Externalities: Theory with Applications." *Review of Economic Studies* 73 (2006), 381-412.

[7] Clark, G. "Why Isn't the Whole World Developed? Lessons from the Cotton Mills." *Journal of Economic History* 47 (March 1987), 141-73.

[8] Coleman, J. "Social Capital in the Creation of Human Capital" *American Journal of Sociology* 94, Supplement: *Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure* (1988), S95-S120.

[9] Easterly, W. *The White Man's Burden: Why The West's Efforts To Aid The Rest Have Done So Much Ill And Little Good.* Penguin Press, New York, 2006.

[10] Eeckhout, J. and B. Jovanovic. "Occupational Choice and Development." *Journal of Economic Theory*, forthcoming.

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

[12] Gemici, A. "Family Migration and Labor Market Outcomes." NYU, June 2008.

[13] Hall, R. and C. Jones "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* (February 1999), 83-116.

[14] Klein, P. and G. J. Ventura. "TFP Differences and the Aggregate Effects of Labor Mobility in the Long Run." *The B.E. Journal of Macroeconomics*, Berkeley Electronic Press 7 (2007), Article 10.

[15] Klenow, P. and A. Rodriguez-Clare. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" *NBER Macroeconomics Annual* 12 (1997), 73-104.

[16] Kremer, M., and E. Maskin. "Globalization and Inequality." Harvard University, 2006.

[17] Lucas, R. E., Jr. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July 1988), 3-42.

[18] Lucas, R. E., Jr. "Why Doesn't Capital Flow from Rich to Poor Countries?" *A.E.A. Papers and Proceedings* 80 (May 1990), 82-86.

[19] Rauch J. E. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities" *Journal of Urban Economics* 34 (November 1993), 380-400.

[20] Rendon, S. and A. Cuecuecha. "International Job Search: Mexicans In and Out of the US." IZA DP #3219, December 2007.

[21] Sala-i-Martin, X. "The World Distribution of Income." *Quarterly Journal of Economics* 121 (May 2006), 351-97.