Optimal Migration: A World Perspective

by

Jess Benhabib and Boyan Jovanovic

June 13, 2008

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. Using a calibrated one-sector model we find that unless the weights are heavily biased towards the rich, 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.

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

Acknowledgement: We thank Jonathan Morduch 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 deriving the result we ignore the political constraints. 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. It is unlikely that this 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.

2 The Model

Immigrants in our model affect the well-being of the residents of the host country through a group effect. This group effect operates through the effect that immigrants have on "social capital," originally discussed by Coleman (1988). We define "social capital" as nationwide human capital per person, $\bar{h}$, which raises the marginal productivity of human capital H. We abstract from physical capital and assume that 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. We assume that private returns to H are constant in human capital alone, so that the private marginal product of human capital is $G\left(\bar{h}\right)$, where $G' > 0$ and $G'' < 0$.[1]

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

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

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

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

Then

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

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

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

[1]Evidence that G should increase as a function of $\bar{h}$ is given by Clark (1987) who invokes national culture, and by Rauch (1993) who attributes it to human-capital spillovers at the level of the Standard Metropolitan Statistical Area. Thus if immigrants decrease the average level of human capital, they depress productivity.

A second, non-market interpretation of G, is one of a cultural externality operating not through production, but through preferences of natives. That is, since agents' utility is a monotone transform of their output or consumption, we can interpret $G\left(\bar{h}\right)$ as an externality acting directly on utility, reflecting a cultural distaste for unskilled immigrants, so that the enjoyment of consumption is diluted in a society where the arrival of less skilled immigrants lowers $\bar{h}$. In this sense, culture is a public good that is diminished by the arrival of unskilled immigrants, but enriched by highly skilled immigrants, so that it is not simply xenophobia.

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

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 ˜ $F_A(h)$. Let μ_B be the mean skills in country B and let the human capital of B's residents be distributed h ˜ $F_B(h)$, 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, $\phi : R \to [0,1]$.

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. Human capital per head in A is

$$\bar{h}_A = \frac{\mu_A + n \int h\phi(h)\, dF_B(h)}{1 + n \int \phi(h)\, dF_B(h)}, \tag{3}$$

and in B it is

$$\bar{h}_B = \frac{\int h [1 - \phi(h)]\, dF_B(h)}{\int [1 - \phi(h)]\, dF_B(h)}. \tag{4}$$

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 he 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, $G(\bar{h}_A)(1-t)h \geq G(\bar{h}_B)h$ or simply that

$$G(\bar{h}_A)(1-t) \geq G(\bar{h}_B). \tag{5}$$

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

$$\theta \int U\left[G\left(\bar{h}_A\right)h\right]dF_A(h)+(1-\theta)\,n\int\left\{\phi(h)\,U\left(G\left[\bar{h}_A\right]h\,[1-t]\right)+[1-\phi(h)]\,U\left(G\left[\bar{h}_B\right]h\right)\right\}dF_B(h). \tag{6}$$

subject to (3), (4), and (5).

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

$$U(c) = \ln c. \tag{7}$$

In this case, the optimal policy is generally 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. The first policy we call "skimming from the bottom" of the F_B distribution; under that policy there exists a cutoff, $\tilde{h}$, such that

$$\phi(h) = \begin{cases} 1 & \text{for } h < \tilde{h} \\ 0 & \text{for } h > \tilde{h} \end{cases} \tag{8}$$

The second "skimming from the top," or simply a "brain-drain" policy:

$$\phi(h) = \begin{cases} 0 & \text{for } h < \tilde{h} \\ 1 & \text{for } h > \tilde{h} \end{cases} \tag{9}$$

At point $\tilde{h}$ we know only that $0 \leq \phi\left(\tilde{h}\right) \leq 1$, the Planner being indifferent about whether $\tilde{h}$ should migrate or not. The rest of this section will prove the first three claims below, and the last is shown by means of simulations.

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 (6) reduces to the equivalent*

$$W \equiv \theta_A^* g_A + \theta_B^*\left(g_B - \ln[1-t]\right) \tag{10}$$

subject to (3) and (4), where

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

Proof. Substituting for U and leaving out terms that do not depend on ϕ, (6) 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 (10). ■

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

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

to be the new control variable that satisfies $z(h):\; R \rightarrow [0,\; n f_B(h)]$ for all h. In terms of this control variable in (10) we have

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

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

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

The FOC is

$$\frac{\partial W}{\partial z(h)} = \lambda_1(h) - \lambda_0(h), \tag{11}$$

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

$$\frac{\partial W}{\partial z(h)} = \begin{cases} < 0 \Longrightarrow \lambda_0(h) > 0 \text{ and } \phi(h) = 0 \\ > 0 \Longrightarrow \lambda_1(h) > 0 \text{ and } \phi(h) = 1 \end{cases} \tag{12}$$

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.

$$\begin{aligned}\frac{\partial W}{\partial z(h)} &= \theta_A^* g'(\bar{h}_A)\frac{h-\bar{h}_A}{\bar{n}_A} - \theta_B^* g'(\bar{h}_B)\frac{h-\bar{h}_B}{\bar{n}_B} + (1-\theta)\left[g(\bar{h}_A) - g(\bar{h}_B) + \ln(1-t)\right] \\ &= \alpha(1-\theta)\left[\ln\frac{\bar{h}_A}{\bar{h}_B} + \frac{\ln(1-t)}{\alpha} + 1 - m + \left(\frac{m}{\bar{h}_A} - \frac{1}{\bar{h}_B}\right) h\right].\end{aligned}, \tag{13}$$

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

$$m \equiv \frac{\theta + (1-\theta) Z}{(1-\theta)(1+Z)} \qquad \text{and} \qquad Z \equiv \int z(h) dh. \tag{14}$$

By (13), $\frac{\partial W}{\partial z(h)}$ is linear in h, which immediately shows that either (8) or (9) must hold, though possibly with $\tilde{h} = 0$ or $\tilde{h} = +\infty$. It is straightforward to show the following features of the optimal policy when $t = 0$.

The case $\theta = 1$.—In this case the Planner cares only about country A. The optimal policy is then characterized by (9), with $\tilde{h} = \bar{h}_A$.

Assumption: Suppose that there are enough unskilled people in B so that migration can equalize the average skills, that is $\bar{h}_A = \bar{h}_B$. That is, there exists an $\hat{h} < \infty$ such that

$$\left(\bar{h}_A =\right) \frac{\mu_A + n \int_0^{\hat{h}} h dF_B(h)}{1 + nF_B\left(\hat{h}\right)} = \frac{\int_{\hat{h}}^{\infty} h dF_B(h)}{1 - F_B\left(\hat{h}\right)} \left(= \bar{h}_B\right). \tag{15}$$

Since $\bar{h}_B = E\left(h_B \mid h_B \geq \hat{h}\right)$, $\hat{h} < \bar{h}_B$. But then, since $\bar{h}_B\left(\hat{h}\right) = \bar{h}_A\left(\hat{h}\right)$, $\hat{h} < \bar{h}_A$. The derivative of the LHS of (15) is $\frac{nf_B}{1+nF_B}\left(\hat{h} - h_A\right) < 0$, whereas the derivative of the RHS is $-\frac{f_B}{1-F_B}\left(\hat{h} - h_A\right) > 0$. The two sides of (15) are continuous, and at any solution the LHS must cut the RHS from above. Therefore if $\hat{h}$ exists, it is unique.

The case $\theta \leq 0.5$.—The policy (8) is optimal. Only the less-skilled are allowed emigrate.

The case $\theta < 0.5$.—Now $m < 1$ for all policies z, and $\tilde{h} \in \left[0, \hat{h}\right]$, there exists a unique optimal policy of the form (8) that $\bar{h}_A\left(\tilde{h}^*\right) = \bar{h}_B\left(\tilde{h}^*\right)$ so that (5) binds.

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.

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

$$\begin{aligned} \mu_{OECD}(\log y) &= \ln 20{,}000, \qquad \sigma_{OECD}(\log y) = \ln 2 \\ \mu_{Rest}(\log y) &= \ln 2{,}000, \qquad \sigma_{Rest}(\log y) = \ln 2.5 \end{aligned}$$

These are portrayed in Figure 2. The following equation identifies $\bar{h}$:

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

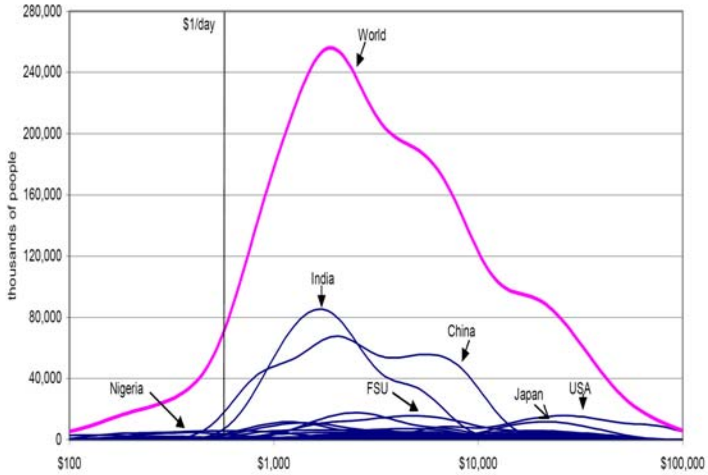

Figure 1: The world income distribution in 2000

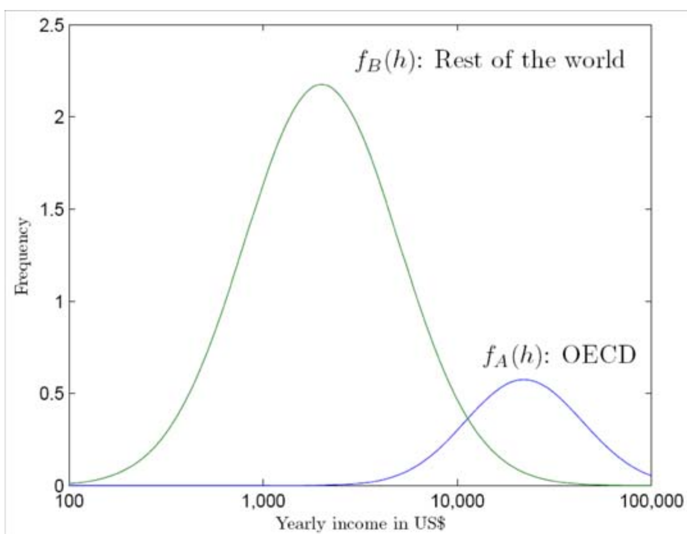

Figure 2: Calibrated distributions of A and B

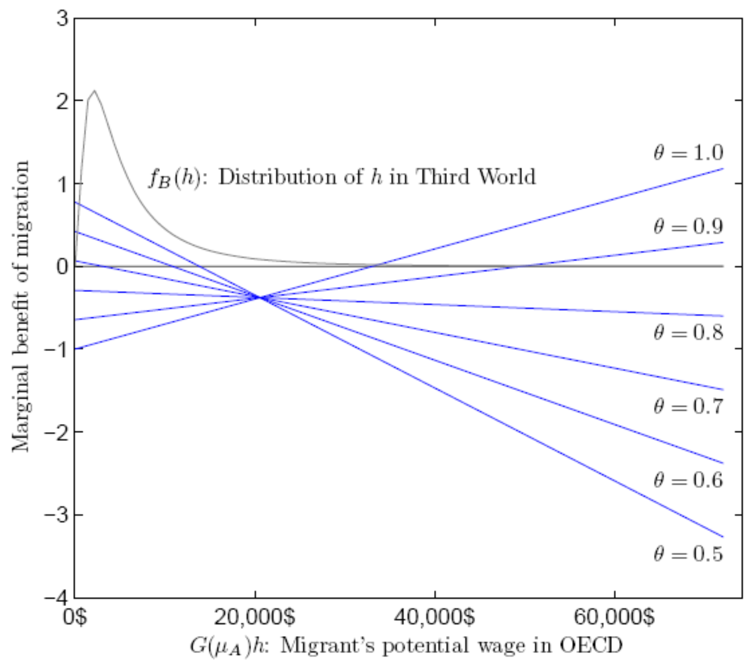

Figure 3: THE FIRST-ORDER CONDITION AT THE STATUS QUO

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 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 (13) 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 $\theta \in [\theta_{\text{BD}}, 1]$ — The slope of the FOC changes sign at

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

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.

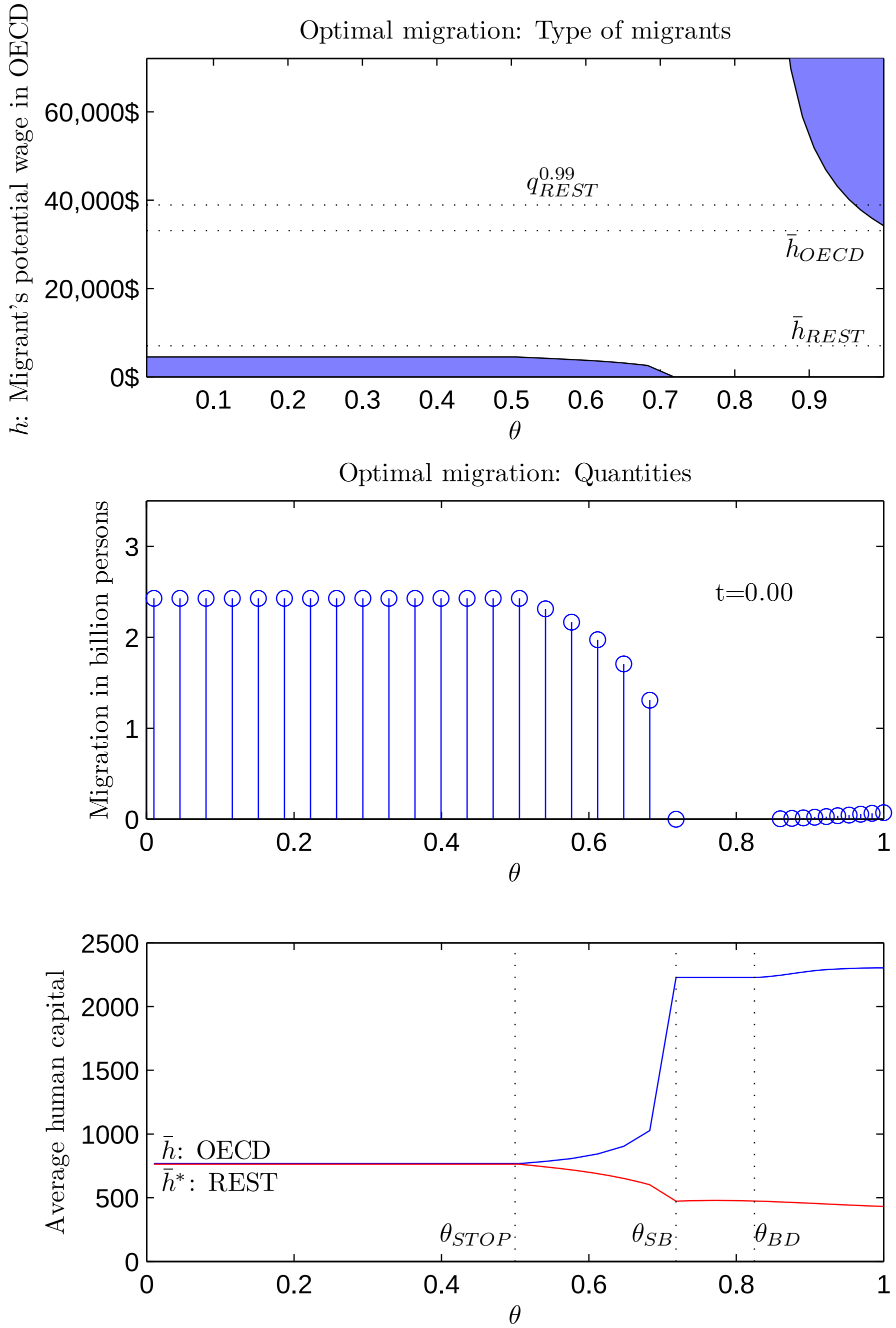

Figure 4: The optimal policy

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 (see Proposition 1) 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 θ's, 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 purple 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 $G\left(\mu_A\right)h$, 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 (5).is binding for $\theta \leq 1/2$.

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[2] to be about seven percent of lifetime income. If we add the Rendon-

[2] 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

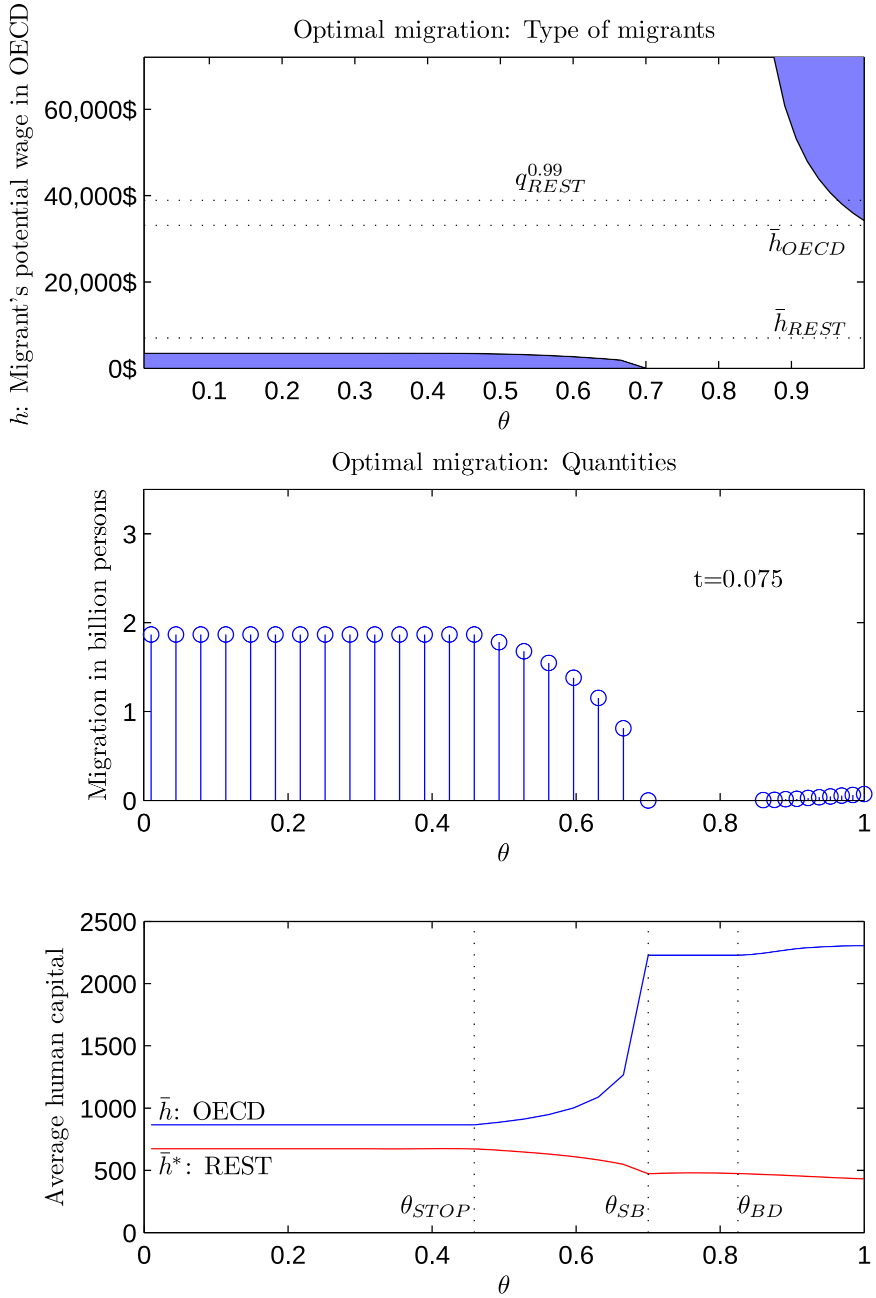

Figure 5: Optimal migration when migration costs are 7.5% of lifetime income in the host country

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.[3] In the bottom panel we see that (5) starts to 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.[4]

4 Robustness: 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

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.

[3]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. 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.

[4]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. Recently France has also moved towards a skill biased immigration policy: see http://www.migrationpolicy.org/pubs/Backgrounder2_

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

$$Y = xG\left(\frac{s}{u}\right) s^{\beta} u^{1-\beta}, \tag{16}$$

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

$$w_s = \beta r^{\beta-1} xG(r), \tag{17}$$

and the wage of an unskilled worker is

$$w_u = (1-\beta) r^{\beta} xG(r). \tag{18}$$

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 (17) and (18), 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

$$W = \theta\left(s_A U\left(w_s^A\right) + u_A U\left(w_u^A\right)\right) + (1-\theta)\left\{\begin{array}{c} s'U\left(w_s^A\right) + u'U\left(w_u^A\right) + \\ \left(s_B - s'\right)U\left(w_s^B\right) + \left(u_B - u'\right)U\left(w_u^B\right) \end{array}\right\}. \tag{19}$$

The incentive-compatibility constraint.—Each factor can flow from B to A only if its wages in A exceed those in B. Instead of (5), then, we now have two IC constraints:

$$\left(w_s^A - w_s^B\right) s' \geq 0, \text{ and } \left(w_u^A - w_u^B\right) u' \geq 0. \tag{20}$$

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 (16) 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

$$G(r) = r^{\alpha} \quad \text{and } \alpha + \beta < 1, \tag{21}$$

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

$$r_B > \left(\frac{x_B}{x_A}\right)^{1/[1-(\alpha+\beta)]} r_A \tag{22}$$

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

$$r_B < \left(\frac{x_A}{x_B}\right)^{1/(\alpha+\beta)} r_A. \tag{23}$$

Proposition 2 *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 $\theta = 1/2$, the Planner equalizes $\bar{h}$ in the two countries. It also assumes that (7) 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.

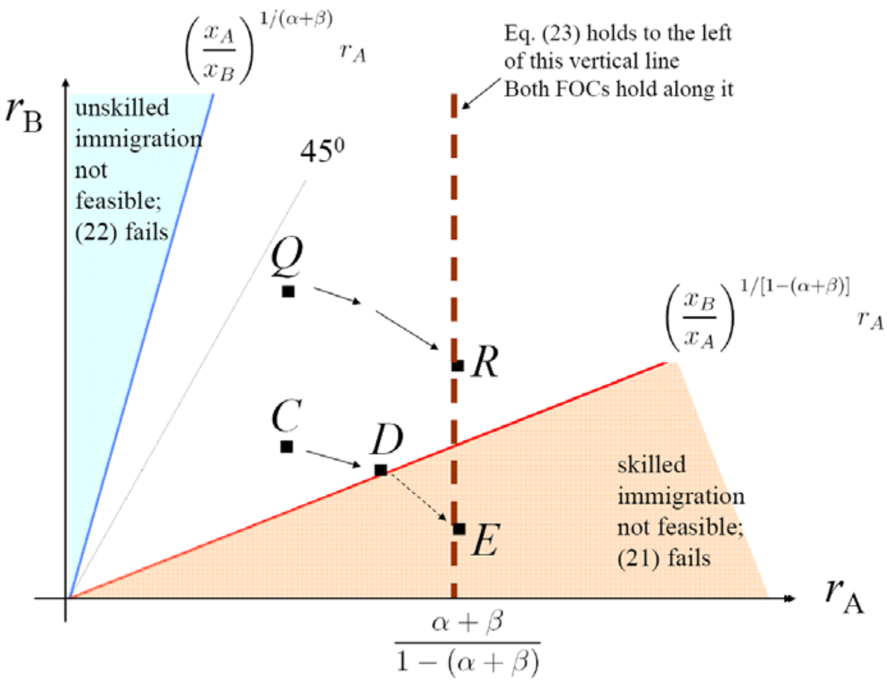

Figure 6: ILLUSTRATION OF PROPOSITION 3

Proposition 3 *If $\theta = 1$ and if*

$$r_A < \frac{\alpha+\beta}{1-(\alpha+\beta)} \tag{24}$$

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 (24) 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 (24) holds or not, we need evidence on the parameters. Condition (24) pertains to the pre-migration relation between factor endowments. Taking the U.S. evidence as an indication, (24) 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 (24) is at least 2. Therefore (24) is the relevant case empirically.

Proposition 7 is illustrated in Figure 6. The planner's FOCs both hold along that dashed line with its intercept at $r_A = \frac{\alpha+\beta}{1-(\alpha+\beta)}$.(See (25) and (26) of the Appendix). If he did not have to worry about (20), 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 6. Now (22) is violated in the bottom-right shaded area, and (23) 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 (24) 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 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 (20) 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

Egalitarian optimal immigration policy from the world perspective, taking into account the economic costs of immigration to the host and source countries, may still require a significant and abrupt relaxation of the restrictive immigration policies currently imposed by the rich countries. With increasing globalization, the third world countries are likely to acquire a greater voice and request greater access to world labor markets. It will may become harder for richer countries to justify their non-discriminatory and redistributive welfare policies at home, while denying the citizens of poorer countries access to their labor markets, and basing the exclusion simply on ethnicity and nationality. While the deep contradictions between the democratic values of the West and the limitations on free access to world labor markets based on nationality have only recently began to surface, they are likely to become increasingly apparent in the future, and enter political discourse through international organizations like the UN or the World Bank. Political negotiations and compromises, however, may at best yield a gradual relaxation of restrictions on labor mobility, as in the case of a slowly expanding EU or the phased legalization of illegal immigrants in the US, rather than an abrupt switch to free immigration that an egalitarian parametrization of our model suggests.

References

[1] Amuedo-Dorantes, Catalina and Cynthia Bansak. "Illegal Border Crossings Attempts Following Apprehension." SDSU, January 2007.

[2] Butcher, Kristin and Anne Morrison Piehl. "Recent Immigrants: Unexpected Implications for Crime and Incarceration," *Industrial and Labor Relations Review* 51, no. 4 (1998): 654-79.

[3] Clark, Gregory. "Why Isn't the Whole World Developed? Lessons from the Cotton Mills." *Journal of Economic History* 47, no. 1 (March 1987): 141-73.

[4] Coleman, James. "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.

[5] Easterly, William. *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.

[6] Friedberg, Rachel, and Jennifer Hunt. "The Impact of Immigrants on Host Country Wages, Employment and Growth." *Journal of Economic Perspectives* 9, no. 2 (Spring, 1995): 23 - 44.

[7] Gemici, Ahu. "Family Migration and Labor Market Outcomes." University of Pennsylvania, August 2007.

[8] Lucas, Robert E., Jr. "Why Doesn't Capital Flow from Rich to Poor Countries?" *A.E.A. Papers and Proceedings* 80, no. 2 (May 1990): 82-86.

[9] Rauch James E. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities" *Journal of Urban Economics* 34, Issue 3 (November 1993): 380-400.

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

[11] Sala-i-Martin, Xavier. "The World Distribution of Income." *Quarterly Journal of Economics* 121, no 2 (May 2006): 351-97.

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 (19) 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 (20). We do so in turn for $\theta = 1$ and $\theta = 1/2$.

Proof of Proposition 6 (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 (21) 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 7 (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

$$\begin{aligned} \frac{\partial W}{\partial s'} &= \frac{-s_A}{s_A + s'} \left(1 - \beta - \frac{G'}{G} \frac{s_A + s'}{u_A + u'}\right) + \frac{u_A}{s_A + s'} \left(\beta + \frac{G'}{G} \frac{s_A + s'}{u_A + u'}\right) \qquad (25) \\ &= \left(\frac{s_A}{s_A + s'}\right) \left[\beta \left(\frac{u_A}{s_A} - \frac{1-\beta}{\beta}\right) + \alpha \left(\frac{s_A + u_A}{s_A}\right)\right], \end{aligned}$$

and with respect to u' it is

$$\begin{aligned} \frac{\partial W}{\partial u'} &= \frac{1}{u_A + u'} \left[\left((1-\beta) s_A - s_A \frac{G'}{G} \left(\frac{s_A + s'}{u_A + u'}\right)\right) + \left(-\beta u_A - u_A \frac{G'}{G} \left(\frac{s_A + s'}{u_A + u'}\right)\right)\right] \qquad (26) \\ &= \frac{-s_A}{u_A + u'} \left[\beta \left(\frac{u_A}{s_A} - \frac{1-\beta}{\beta}\right) + \alpha \left(\frac{s_A + u_A}{s_A}\right)\right]. \end{aligned}$$

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., (24).