Human capital acquisition and international migration in a model of educational market

Roman Zakharenko[*]

November 15, 2011

Abstract

This paper analyzes international high-skilled migration caused by financial frictions in educational market. I develop a model of learning in which acquisition of skill is only possible through personal interaction with a skilled individual; the income of the skilled is sensitive to financial constraints for the unskilled. Cross-country differences in such constraints have a multiplicative effect on the skill premium, causing outmigration of skilled individuals from a less developed country. I study welfare implications of such brain drain for the sending and receiving countries. Although it makes more difficult skill acquisition in the sending country, the unskilled may still be better off: increased cost of skill acquisition is offset by higher income once the skill has been acquired. For the receiving country, I identify a phenomenon of *immiserizing immigration*: a depletion of the stock of skill in the sending country due to brain drain hinders further production of skill, which may hurt the receiving country. Additionally, I find that increased openness of the sending country to migration and the resultant accelerated brain drain increase the incentives of the country government to reduce financial frictions.

Keywords: skill acquisition; brain drain; immiserizing immigration; financial constraints.
JEL codes: F22, J61, O15

1 Introduction

Despite the unprecedented development of long-distance communication technologies, knowledge continues to diffuse from one person to another mainly by means of personal interaction. One can become a scientist only through a continuous interaction with other scientists. In stable political environments, virtually all successful politicians have an experience of interaction with politicians from previous cohorts. In most jobs, young workers learn from experienced workers. Teaching services continue to be local in nature, and university professors in the United States do not expect that their jobs will ever be exported to India. Even the acquisition of skills that are labeled by economists as "low," such as taxi driving, require frequent personal interaction with people who have been in the business for some time. Numerous studies find that the first destination of immigrant workers is usually a location where many immigrants from the same country live,[1] despite the fact that the new immigrants, whose skills are usually similar to that of incumbent immigrants, would face less competition on the job market in other locations.

Given the local nature of transmission of productive knowledge, part of the welfare gain from enhanced skills is likely to be shifted from the learners (young, unskilled workers) to the teachers (experienced, skilled workers) through a bargaining process, pushing current income flow of the learners below their current marginal product of labor, and vice versa, raising the income of the teachers. Further, a compensation for education may have a multiplicative effect on the willingness to acquire skill and on the return to skill. When skilled individuals increase their earnings by receiving a compensation for education, unskilled individuals have an increased willingness to acquire skill. Since the best way to acquire skill is to interact with and learn from existing skilled individuals, the latter get a further increase in earnings, further increasing the willingness of the unskilled to acquire skill. This multiplicative positive effect on demand for education offsets the traditional law of demand and makes the demand for education highly inelastic.

With inelastic demand, even small exogenous cross-country differences in the educational technology, such as differences in student access to educational credit, may lead to large differences in the return to skill and create a basis for brain drain from a country with a poor access to educational credit. The objective of this paper is to examine, theoretically, the effects of international migration of skilled individuals caused by such institutional differences, in a model of learning that captures the above mentioned stylized facts. The proposed

---

[1]E.g. Winters, Janvry, and Sadoulet (2001), Bauer, Epstein, and Gang (2007), Munshi (2003); Vergalli (2006) provides a theoretical analysis of the location choice of migrants.

theory differs from the existing literature that tends to explain international migration by the differences in how much *current* output migrants can produce in different countries. The theoretical model developed in this paper examines migration between countries with identical fundamental parameters (productivity of skilled and unskilled, the learning technology), with the only exogenous heterogeneity being the heterogeneity of institutions that facilitate the transfer of wealth from the unskilled to the skilled in compensation for education.

I find the welfare effects of international migration to be ambiguous for both sending and receiving countries. For the sending country, the departure of a fraction of skilled workers reduces the number of potential teachers and makes it more difficult to acquire skill (the negative effect of openness). On the other hand, increased country openness increases the return to skill, which makes unskilled individuals, who expect to acquire skill in the future, better off (the positive effect of openness). I show that with very high transaction costs in the home educational market, the negative effect of brain drain dominates, while with medium values of transaction costs, the positive effect may overwhelm once the magnitude of migration becomes sufficiently high.

For the receiving country, the welfare effects of skilled immigration are generally positive. I show however that once the skilled emigration *rate* becomes sufficiently high, the *number* of migrants begins to diminish, which hurts not only the sending but also the receiving country. The explanation is simple: if skilled individuals are an essential input of new skill production in the sending country, an excessive brain drain may reduce the number of skilled individuals in the sending country by so much that it can no longer supply skilled individuals to the receiving country. I label this phenomenon as *immiserizing immigration* by analogy with Bhagwati's immiserizing growth in the international trade literature. The immiserizing immigration effect is also reminiscent of the overfishing effect in the Economics of natural resources.

Additionally, I study the welfare gains of the sending country from a marginal increase in the quality of educational credit institutions. I find that such gain is always higher in a more open country: the welfare of the unskilled is more sensitive to the quality of educational institutions when the skilled have a better opportunity to leave. Thus, the increased openness and the resultant increased brain drain help "discipline"a home country government to improve institutions facilitating education.

Although there exists an extensive literature on the diffusion of knowledge[2] and on posi-

---

[2]Jovanovic and Rob (1989) is an example of theoretical analysis; Keller (2002) is an empirical account of geographic localization; Keller (2004) contains a review of literature on knowledge spillovers.

tive externalities of human capital,[3] virtually all of this literature assumes that all the welfare gains of skill acquisition accrue only to those who acquire the skill. Park (1997) is, to my knowledge, the only discussion of the fact that part of the welfare gain may be shifted to those skilled from whom new skill has been acquired. The main focus of Park (1997), however, is different from that of this paper, and does not include the examination of migration patterns.

Much of the modern literature that relates education and migration begins with Stark, Helmenstein, and Prskawetz (1997) and Mountford (1997) who point out the potential incentive effect of a brain drain prospect: unskilled individuals who face a prospect of emigration have an incentive to study more if emigration is conditioned on skill acquisition. According to the theories, brain drain may thus increase the total number of skilled in the world, and, assuming that emigration is a probabilistic outcome, may even increase the number of skilled that remain in the sending country. Beine, Docquier, and Rapoport (2001) and Beine, Docquier, and Rapoport (2008) take the theory to cross-country data; Chand and Clemens (2008) tests the theory in the context of migration from Fiji to New Zealand. All of these studies assume that the technology of skill acquisition does not include the presence of individuals already possessing the skill, the assumption challenged in this paper.

The relationship between financial constraints and human capital accumulation has received a considerable attention in the economic development literature. At the theoretical level, Galor and Zeira (1993) is a seminal contribution to the area that shows that, in the presence of financial constraints, a higher initial income inequality may lead the economy to a poverty trap. Mejia and St-Pierre (2008) is a more recent theoretical investigation of a related topic. On the empirical side, the issue has been studied, among others, by by Flug, Spilimbergo, and Wachtenheim (1998) who discover a significant and robust negative effect of the financial constraints on the human capital accumulation; Schady (2004) examines the effect of macroeconomic crises on education.

The role of this paper is combine, theoretically, the two stands of literature – the one that relates educational attainment with migration prospects, and the one that relates educational attainment with financial frictions – to analyze their interaction. In the remainder of the paper, I first analyze a one-country ("closed economy") dynamics and a steady-state, and then proceed to a two-country setting to model migration between the two countries.

---

[3]This literature starts with Lucas (1988); applications of this concept to migration include Stark, Helmenstein, and Prskawetz (1997), Mountford (1997), Stark and Wang (2002), and Stark and Zakharenko (2012).

2 Closed Economy

2.1 Overview

2.1.1 Demographics

This is a general equilibrium dynamic model. Time is discrete; at each moment of time t, there is a continuum of mass L of individuals that are endogenously divided into two types – skilled and unskilled. I denote the fraction of skilled individuals in the economy in period t by m_t. Between any two time periods, a randomly selected fraction $1 - \delta$ of all individuals dies. The same mass of new individuals is born; therefore, the total population remains constant. Every newly-born individual is unskilled.

2.1.2 Consumption and savings

There is one consumption good, which is produced using the only input – skill – in a manner specified below. The price of the good is normalized to unity. A representative individual maximizes his discounted stream of consumption by:

$$U_{i,t_0} = \begin{cases} \sum_{t=t_0...\infty} \beta^{t-t_0} c_{i,t} & \text{if } c_{i,t} \geq 0 \, \forall t \\ -\infty & \text{otherwise} \end{cases} \tag{1}$$

where t_0 is the birth date of the individual, t is the index of time, $c_t \geq 0$ is consumption at time t, and $\beta < 1$ is the discount factor. Given that death is a random occurrence, individuals do not know the moment of their death and calculate their utility on an infinite time horizon. I assume that the death probability is already built into the discount factor (thus $\beta < \delta < 1$), and therefore the parameter δ does not explicitly enter the decision-making process.

Individuals earn a stream of (stochastic) income which depends on skill and which is detailed below; for now, denote it $y_{i,t}, t \geq t_0$. Individuals borrow and save at rate r_t. To model financial frictions, we make a somewhat simplistic assumption that those who need to borrow one dollar actually have to borrow $K \geq 1$ dollars; the difference of $K-1$ dollars is the sunk cost due to financial frictions. Throughout the paper, we refer to K as the "transaction cost"; it reflects the quality of financial institutions in a country.

Denote by

$$a_{i,t} \equiv y_{i,t} - c_{i,t} \tag{2}$$

the flow of savings and by $A_{i,t}$ its stock at the beginning of period t; given the above assumptions, the evolution of $A_{i,t}$ is determined by

$$A_{i,t+1} = a_{i,t}K^{I(a_{i,t}<0)} + (1 + r_t)A_{i,t} \tag{3}$$

where $I(\cdot)$ is the indicator function.

In every period, the savings rate r_t is such that it clears the savings market: $\int_i a_{i,t}K^{I(a_{i,t}<0)}di = 0$, with the initial aggregate stock of savings also equal to zero. The dividends are payable at the beginning of each period.

Given that the death occurrence is random, the average savings of those who die are equal to the average savings of the entire population (i.e. zero). We assume that savings of those who die are cleared, and all newly-born start with zero savings.

2.1.3 Production and learning

In each period of time, each skilled individual produces two types of output: one unit of consumption good and one unit of teaching services. Both outputs are supplied inelastically.

The productivity of unskilled individuals is normalized to zero.[4] The only way for unskilled individuals (students) to become skilled is through personal interaction with existing skilled individuals (teachers). The process of learning is stochastic and depends on the *learning intensity* of students $x_{i,t} \geq 0$, which has two possible interpretations. First, it could be viewed as a fraction of time a student has spent learning within each time period. Alternatively, we can assume that all students learn full time, but in classes of variable sizes: low intensity $x_{i,t}$ implies learning in a large class, while high intensity implies learning in a small class, or individually, or even individually with several teachers. In this stylized model, $x_{i,t}$ of a representative unskilled individual must be equal to the ratio of teachers to students in equilibrium.

If an unskilled individual learns with intensity $x_{i,t}$, his probability of becoming skilled at that point in time is $P(x_{i,t})$ where $P : R_+ \to [0,1]$ is a smooth, strictly increasing, and strictly concave "learning" function that satisfies the following property:

$$P'(0) > \frac{1 - \delta}{\delta} \tag{4}$$

Teaching services are supplied in an educational market at a prevalent "wage" (i.e. com-

---

[4]Generalizing the unskilled productivity to a positive value would reduce the incentive to acquire skill, which would entail a lower return to skill, but it would not qualitatively change the results that follow.

pensation for teaching), denoted w_t. Thus the income of a skilled individual is $y_{i,t} = 1 + w_t$, while the income of the unskilled, net of educational expenses, is $y_{i,t} = -w_t x_{i,t}$.

2.1.4 Objectives

The objective of an individual i is to maximize (1) by making savings decisions $a_{i,t}$ given expectations about the flow of income, and subject to (2), (3), and to the transversality condition

$$\lim_{t \to \infty} \beta^{t-t_0} c_{i,t} = 0 \tag{5}$$

2.2 Equilibrium analysis

Skilled individuals have reached their terminal state of knowledge, and supply both their outputs, consumption good and teaching services, inelastically. As a result, the only decision they have to make is the savings decision. The unskilled individuals make learning intensity decisions $x_{i,t}$ and savings decisions $a_{i,t}$. The appendix shows that, given that the utility is linear in consumption, the financial part of the model, i.e. savings decisions, may be suppressed to reduce the dimensionality of the problem. In the rest of the paper, we analyze the simplified "savings-free" version of the model.

2.2.1 Skilled individuals

In the "savings-free" modification of the model, skilled individuals do not make any decisions (as they supply both outputs, consumption good and teaching services, inelastically);[5] the value of being skilled is the discounted stream of future earnings:

$$v_1^t \equiv v_1(w_t, w_{t+1}, ...) = \sum_{\tau=t}^{\infty} \beta^{\tau-t}(1 + w_\tau) \tag{6}$$

To analyze the aggregate supply of teaching services, it is sufficient to analyze the dynamics of the fraction of skilled individuals in the population, m_t. Observe that, in the closed economy, such fraction does not depend on transaction costs K and wage levels w_τ: the equilibrium learning intensity at time t, x_t, is determined by the current proportion of

---

[5]One can develop a model in which skilled individuals have a tradeoff between production and teaching, which would result in a traditional upward-sloping supply of teaching services. This would, however, complicate the analysis that follows, without adding new insights to the migration part of the model.

skilled to unskilled individuals,

$$x_t = \frac{m_t}{1 - m_t} \tag{7}$$

while the dynamics of m_t is fully determined by the learning intensity:

$$m_{t+1} = \delta \left(m_t + P(x_t)(1 - m_t) \right) \tag{8}$$

Taking into account (4), we can find that the dynamic system given by (7) and (8) has two steady states: one, unstable, at $m_t = 0$, and a stable one at some $m_t = m_A \in (0, 1)$. Thus, a closed economy that starts at time t_0 from any $m_{t_0} \in (0, 1]$ converges to $\lim_{t \to \infty} m_t = m_A$. The corresponding value of the learning intensity is denoted by $x_A \equiv \frac{m_A}{1 - m_A}$ and is determined by

$$\frac{P(x_A)}{x_A} = \frac{1 - \delta}{\delta} \tag{9}$$

2.2.2 Unskilled individuals, demand for education and its properties

The appendix demonstrates that the "savings-free" unskilled optimization problem may be reduced to the following:

$$v_0^t = \max_z \left(-K w_t z + \beta \left[P(z) v_1^{t+1} + (1 - P(z)) v_0^{t+1} \right] \right) \tag{10}$$

Denote by $x^D(w_t, w_{t+1}, ...; K)$ the argmaximum of (10), and by $V(z; w_t, w_{t+1}, ...; K)$ its maximand:

$$V(z; w_t, w_{t+1}, ...; K) = -K w_t z + \beta \left[P(z) v_1^{t+1} + (1 - P(z)) v_0^{t+1} \right] \tag{11}$$

We next outline some properties of the demand for education.

Lemma 1 *The difference between the values of being skilled and unskilled, $v_1^t - v_0^t = v_1(w_t, w_{t+1}, ...) - v_0(w_t, w_{t+1}, ...; K)$, is increasing in w_τ for every t and for every $\tau \geq t$, such that*

$$\frac{\partial (v_1^t - v_0^t)}{\partial w_t} = 1 + K x_t^D > 0 \tag{12}$$

$$\left. \frac{\partial (v_1^t - v_0^t)}{\partial w_{t+k}} \right|_{k>0} = \beta^k \prod_{i=0}^{k-1} (1 - P(x_{t+i}^D))(1 + K x_{t+k}^D) > 0 \tag{13}$$

8

Proof. For every t, we have that $\frac{\partial v_1^t}{\partial w_t} = 1$, while $\frac{\partial v_0^t}{\partial w_t} = -Kx_t^D$, hence (12). To prove (13), observe that $\frac{\partial v_1^t}{\partial w_{t+k}} = \beta^k = \beta\frac{\partial v_1^{t+1}}{\partial w_{t+k}}$, while

$$\frac{\partial v_0^t}{\partial w_{t+k}} = \beta P(x_t^D)\left[\frac{\partial v_1^t}{\partial w_{t+k}} - \frac{\partial v_0^t}{\partial w_{t+k}}\right]$$

The difference between the two is then

$$\frac{\partial v_1^t}{\partial w_{t+k}} - \frac{\partial v_0^t}{\partial w_{t+k}} = \beta(1 - P(x_t^D))\left[\frac{\partial v_1^{t+1}}{\partial w_{t+k}} - \frac{\partial v_0^{t+1}}{\partial w_{t+k}}\right]$$

By iterating the last result, we obtain (13). ∎

Lemma 1 is used in the proof of the following

Proposition 1 *The optimal demand for education* $x^D(w_t, w_{t+1}, ...; K)$ *is decreasing in current wage* w_t *and is increasing in every subsequent wage* $w_\tau, \tau > t$.

Proof. From the implicit function theorem it follows that

$$\frac{\partial x_t^D}{\partial w_\tau} = \frac{\frac{\partial^2 V_t}{\partial z \partial w\tau}}{-\frac{\partial^2 V_t}{\partial z^2}} \tag{14}$$

Given concavity of V_t with respect to z, the denominator of (14) is positive and therefore the sign of $\frac{\partial x_t}{\partial w_\tau}$ is equal to the sign of $\frac{\partial^2 V_t}{\partial z \partial w\tau}$. The latter is equal to

$$\frac{\partial^2 V_t}{\partial z \partial w_t} = -K < 0$$
$$\left.\frac{\partial^2 V_t}{\partial z \partial w_\tau}\right|_{\tau > t} = \beta P'(z)\left[\frac{\partial v_1^{t+1}}{\partial w_\tau} - \frac{\partial v_0^{t+1}}{\partial w_\tau}\right] > 0$$

∎

The optimal demand for education increases in future wages for two reasons. First, should the individual become skilled in the future, his wealth will be improved by higher future wage, hence the willingness to increase the likelihood to become skilled by demanding more education today, which we call the *income effect*. Second, should the individual remain unskilled in the future, his cost of education in the future will be increased by higher future teacher's wage, hence the willingness to reduce the likelihood of remaining unskilled and demand more education today, which we call the *substitution effect*. Figure 1 displays the demand for education x_t, as a function of current wage w_t, assuming that future wages $w_\tau, \tau > t$ are held time-invariant at various levels.
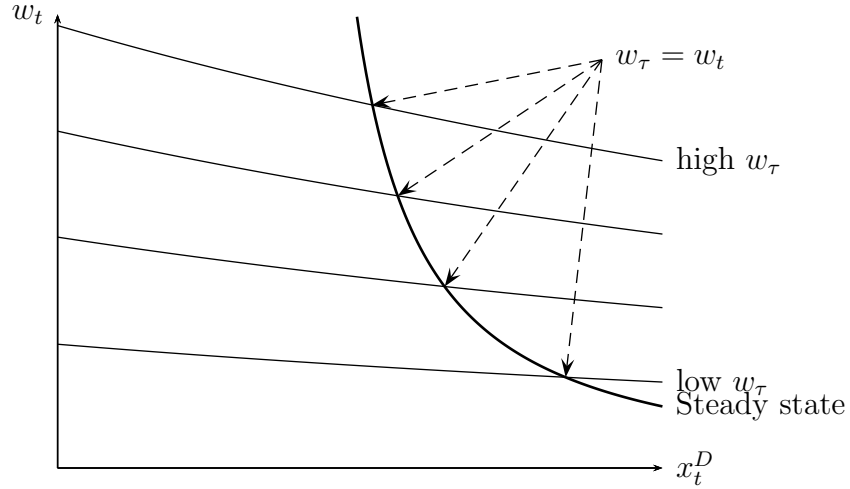
Figure 1: Demand for education x_t^D as function of current wage w_t at various levels of future wage w_τ

2.2.3 Equilibrium wage path

The equilibrium wage path is determined by setting the demanded learning intensity $x^D(w_t, w_{t+1}, ...; K)$ equal to the supplied learning intensity x_t, defined by (7), at every time period $t \in [t_0, \infty)$.

2.3 Steady state analysis

In a steady state where the wage is constant over time, $w_t = w$, the value of being skilled (6) simplifies to

$$v_1(w) = \frac{1 + w}{1 - \beta} \tag{15}$$

while the steady-state supply of teaching services is given by (9).

The steady-state version of the value of being unskilled, (10), is:

$$v_0(w, K) = \max_{z \geq 0} \left(-Kwz + \beta \left[P(z)v_1(w) + (1 - P(z))v_0(w, K) \right] \right) \tag{16}$$

The following Proposition characterizes the steady-state demand for education by unskilled individuals.

Proposition 2 *The demand for skill by the unskilled individuals, $x(w, K)$, is characterized by*

$$w \geq \frac{G(x(w, K))}{K - G(x(w, K))} \tag{17}$$

10

with strict equality if $x(w, K) > 0$, where

$$G(x) \equiv \frac{P'(x)}{\frac{1-\beta}{\beta} + s(x)} \tag{18}$$

$$s(x) \equiv P(x) - P'(x)x \tag{19}$$

Refer to the Appendix for proof.

From the properties of $P(x)$, it follows that $s(x)$ is increasing from zero to one, while $G(x)$ is decreasing from $G(0) = \frac{\beta}{1-\beta}P'(0)$ to zero.

Define by $\underline{x}(K)$ the lower bound of demand:

$$\underline{x}(K) \equiv \min_w x(w, K) = x(\infty, K)$$

When transaction costs are sufficiently low, $K < G(0) = P'(0)\frac{\beta}{1-\beta}$, demand for education is bounded away from zero, $\underline{x}(K) > 0$ and is determined by

$$G(\underline{x}(K)) = K \tag{20}$$

which can be interpreted as follows. A marginal increase of w has two effects. First, w can be seen as a tuition, and the law of demand prescribes unskilled individuals to demand less of teaching services when tuition increases. Second, an increased w means an increased return to skill (since education is provided by skilled individuals), which means an increased willingness to acquire skill, implying an increased demand for education. With high transaction costs K, the former effect overwhelms the latter, and demand for education is a standard textbook demand function with negative and bounded away from zero own-price elasticity. When transaction costs K are sufficiently low, the two effects offset each other, and the learning intensity approaches its lower bound $\underline{x} > 0$, while demand elasticity approaches zero, as the tuition w approaches infinity.

Next, we investigate the steady-state expected value ("welfare") of being unskilled $v_0(w, K)$.

Lemma 2

$$v_0(w, K) = \frac{w + 1}{1 - \beta}S(x(w, K)) \tag{21}$$

$$S(x(w, K)) \equiv \frac{s(x(w, K))}{\frac{1-\beta}{\beta} + s(x(w, K))} \tag{22}$$

The proof is contained in Appendix.

How does $v_0(w, K)$ depend on its arguments? The effect of increased transaction costs K is rather obvious: at a constant wage w, a rise in the transaction costs must be offset by a proportional rise in $G(x)$ (cf.(17)), which implies a decline in learning intensity x. A decline in x causes a decline in welfare (21).

The effect of an increased wage is less obvious: besides the direct positive effect, it has a negative effect on demanded learning intensity $x(w, K)$ thus the total effect is uncertain. The Proposition below clarifies the ambiguity.

Proposition 3 *At a given K, the welfare $v_0(w, K)$ reaches its minimum at $w^*(K)$ and $x^*(K) = x(w^*(K), K)$ such that*

$$\frac{P(x^*(K))}{x^*(K)} \leq \frac{1 - \beta}{\beta} K \tag{23}$$

with strict equality if $x^*(K) > 0$. Moreover, welfare increases with w when $w > w^*(K)$ $(x < x^*(K))$ and decreases when $w < w^*(K)$ $(x > x^*(K))$. In other words, the point $(x^*(K), w^*(K))$ is the unique point of minimum welfare on the demand curve (17).

Proof. The statement is proved by direct computation and analysis of the first derivative of (21) with respect to w.

$$
\begin{aligned}
(1 - \beta)\frac{\partial v_0(w, K)}{\partial w} &= S(x) + (w + 1)\frac{\partial S(x)}{\partial x}\left(\frac{\partial w}{\partial x}\right)^{-1} \\
&\underbrace{=}_{\text{cf.(17)}} S(x) + \frac{K}{K - G(x)}\frac{\partial S(x)}{\partial x}\frac{(K - G(x))^2}{K}\left(\frac{\partial G(x)}{\partial x}\right)^{-1} \\
&= S(x) + (K - G(x))\frac{\partial S(x)}{\partial x}\left(\frac{\partial G(x)}{dx}\right)^{-1} \tag{24}
\end{aligned}
$$

The Appendix proves that (24) is equal to

$$(1 - \beta)\frac{\partial v_0}{\partial w} = \frac{P(x) - \frac{1 - \beta}{\beta}Kx}{\left(\frac{1 - \beta}{\beta} + P(x)\right)}$$

which is positive if $x < x^*(K)$ $(w > w^*(K))$, negative if $x > x^*(K)$ $(w < w^*(K))$, and is zero otherwise. Thus, a country with transaction costs K reaches the minimum of its steady-state welfare when the steady-state wage is equal to $w^*(K)$ and learning intensity is equal to $x^*(K)$. $\blacksquare$

We next investigate the relationship between points of demanded learning intensity minimum $\underline{x}(K)$, unskilled welfare argminimum $x^*(K)$, and autarky learning intensity x_A.

Proposition 4 (a) *If* $K \geq P'(0)\frac{\beta}{1-\beta}$, *then* $0 = \underline{x}(K) = x^*(K) < x_A$.

(b) *If* $K < P'(0)\frac{\beta}{1-\beta}$, *then* $0 < \underline{x}(K) < x^*(K) < x_A$ *and both* $\underline{x}(K)$ *and* $x^*(K)$ *are strictly diminishing with* K.

Proof.

(a) If $K \geq P'(0)\frac{\beta}{1-\beta}$, the denominator of (17) is positive at every positive value of learning intensity x, and therefore the wage achieves its upper bound at $\underline{x}(K) = 0$. To prove $x^*(K) = 0$, suppose the opposite, $x^*(K) > 0$. Then we have that

$$P'(0) \leq \underbrace{K\frac{1-\beta}{\beta}}_{\text{cf.}(23)} = \frac{P(x^*(K))}{x^*(K)} \underbrace{<}_{P''(\cdot)<0} \underbrace{\lim_{z\to 0}\frac{P(z)}{z}}_{\text{L'Hopital's rule}} = P'(0)$$

which is a contradiction.

(b) When $K < P'(0)\frac{\beta}{1-\beta}$, the value of $\underline{x}(K)$ is determined by (20), the left-hand side of which is decreasing in $\underline{x}$. At $\underline{x} = 0$, the left-hand side is greater than the right-hand side, hence the equality is achieved at $\underline{x}(K) > 0$. Additionally, an increase in K increases the right-hand side of (20), hence the equality is achieved at a lower value of $\underline{x}$, meaning $\frac{\partial \underline{x}(K)}{\partial K} < 0$.

To prove $x^*(K) > \underline{x}(K)$, observe that

$$\frac{\beta}{1-\beta}P'(x^*(K)) < \frac{\beta}{1-\beta}\frac{P(x^*(K))}{x^*(K)} \equiv K \equiv G(\underline{x}(K)) \equiv \frac{P'(\underline{x}(K))}{\frac{1-\beta}{\beta} + s(\underline{x}(K))} < \frac{\beta}{1-\beta}P'(\underline{x}(K))$$

By assumption, $P'(x)$ is strictly decreasing in x, which ensures $x^*(K) > \underline{x}(K)$.

To prove $\frac{\partial x^*(K)}{\partial K} < 0$, observe that the left-hand side of (23) is decreasing in x, hence a higher value of K yields a lower x.

To prove $x^*(K) < x_A$, observe that

$$\frac{\beta}{1-\beta}\frac{P(x^*(K))}{x^*(K)} \equiv K \geq 1 \equiv \frac{\delta}{1-\delta}\frac{P(x_A)}{x_A} > \frac{\beta}{1-\beta}\frac{P(x_A)}{x_A}$$

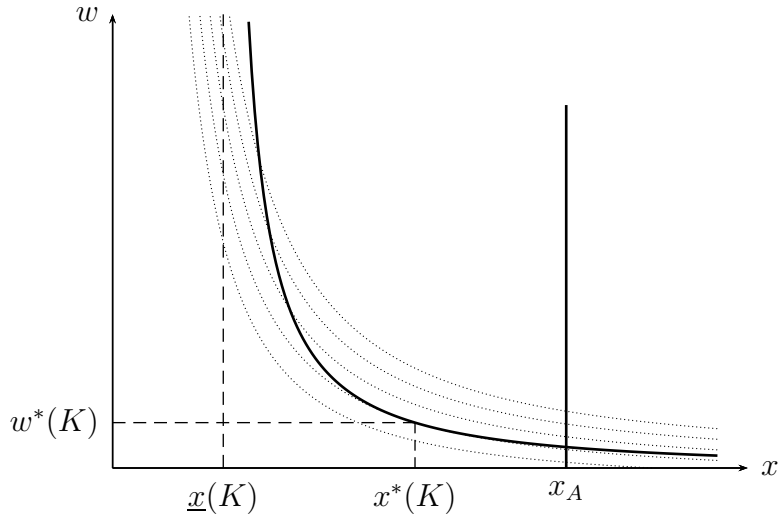The ratio $\frac{P(x)}{x}$ is strictly decreasing in x, which ensures $x^*(K) < x_A$.

13

Figure 2: Isowelfare lines versus the demand curve

■

Figure 2 illustrates the point of welfare minimum on a typical demand curve.

3   International migration

3.1   Overview

In this paper, I suggest a new potential cause of the brain drain from less developed to more developed countries. Even if the origin and destination countries have identical "real-sector" parameters such as marginal product of labor of both skilled and unskilled workers, fertility, life expectancy, and learning technology, brain drain may still exist due to differences in the technology of transferring wealth from those who are willing to acquire skill to those who provide teaching services. Such differences, manifested in the form of costs of transactions between teachers and students, lead to differences in return to skill across countries. Moreover, lower return to skill in a less developed country decreases the willingness to acquire such skill, which leads to further decrease in return to skill. This multiplicative effect of institutional differences creates an incentive for skilled individuals to migrate from less developed to more developed countries. In this section, I conduct a formal analysis of this effect and study its welfare implications.

To isolate the proposed effect, I assume that all other characteristics of the origin and the destination countries of migrants, such as labor productivity or "psychic costs" of living,

14

are equal across countries; the only difference among countries is the ease of borrowing for students. If students have limited access to credit, they offer lower rewards to their teachers which makes the latter emigrate; students' access to education thus becomes limited.

Suppose there are two countries, North, with low transaction costs $K_N = 1$, and less developed South with higher transaction costs $K_S > K_N$. I make a simplistic assumption that the number of births each period in each country is fixed and does not depend on migration patterns; this way, the number of individuals of a given nationality is exogenous. In country i, the number of births is $(1-\delta)L_i$, so that the number of nationals of the country is equal to L_i and is not affected by migration patterns.

Only skilled individuals can migrate across countries; this assumption is consistent with the selective immigration policy exercised by most developed recipient countries. An individual who chooses to migrate pays a one-time migration cost M.[6] Thus, migration is possible when the cost of doing so is low enough: $M \leq \frac{w_N^A - w_S^A}{1-\beta}$, where w_i^A is the autarky wage.

3.2 Wage curve

Under the assumption of sufficiently low cost of migration, the Northern and Southern wages are related as follows:

$$w_S = w_N - (1 - \beta)M$$

or, equivalently, (cf.(17))

$$\frac{G(x_S)}{K_S - G(x_S)} = \frac{G(x_N)}{K_N - G(x_N)} - (1 - \beta)M \tag{25}$$

Equation (25) defines a hypersurface in the (x_N, x_S) space, which we refer to as the *wage curve*. The wage curve originates from the $(\underline{x}(K_N), \underline{x}(K_S))$ point and slopes upward; a decrease of migration cost M results in a lower x_S (i.e. higher w_S) at a given x_N, hence a reduction of migration costs rotates the wage curve clockwise in the (x_N, x_S) space.

---

[6]One important phenomenon not captured in this paper is that limited access to credit negatively affects not only the acquisition of skill, but also the ability of individuals to migrate. In this paper, the cost of migration M is assumed to be independent of institutions K: skilled individuals that consider migration are assumed to have enough funds for such migration.

3.3 Demographic balance curve

Denote by r the endogenous fraction of skilled Southerners who emigrate to the North; then, the steady-state learning intensity of Northern students is

$$x_N = \frac{m_N + r\frac{L_S}{L_N}m_S}{1 - m_N} \tag{26}$$

while the learning intensity of Southern students is

$$x_S = \frac{(1 - r)m_S}{1 - m_S} \tag{27}$$

The steady-state share of skilled of nationality $i \in \{N, S\}$ is determined by (cf.(8))

$$m_i = \delta \left(m_i + P(x_i)(1 - m_i) \right) \tag{28}$$

By manipulating with (26), (27), and (28), this system of four equations with five unknowns may be reduced to the following equation with two unknowns:

$$L_N D(x_N) = -L_S D(x_S) \tag{29}$$

$$D(x) \equiv \frac{(1 - \delta)x - \delta P(x)}{(1 - \delta) + \delta P(x)} \tag{30}$$

Equation (29) defines another hypersurface in the (x_N, x_S) space, referred to as the *demographic balance curve*. The function $D(x)$ is equal to the per-capita demand for educational services in a country, $\frac{(1-\delta)x}{(1-\delta)+\delta P(x)} = (1 - m)x$, minus the per-capita supply of educational services by the skilled nationals of that country, regardless of their country of residence, $\frac{\delta P(x)}{(1-\delta)+\delta P(x)} = m$. Naturally, both sides of (29) must be equal to the supply of Southern educational services in the North, which is also equal to the number of Southern migrants. The following Lemma outlines the properties of $D(\cdot)$.

Lemma 3 *The function $D(\cdot)$ has the following properties:*

(a) $D(0) = D(x_A) = 0$;

(b) *It has a unique extremum $x_D \in (0, x_A)$ such that*

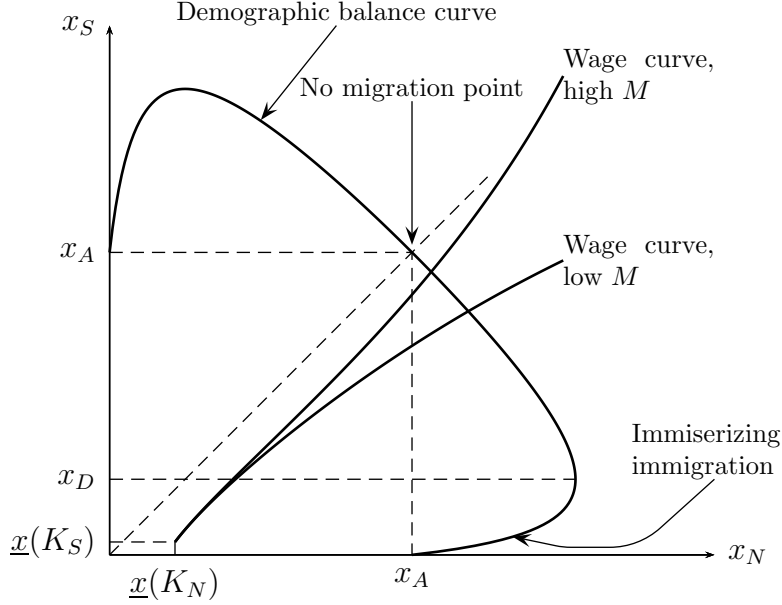$$(P'(x_D) - s(x_D)) \frac{\delta}{1 - \delta} \equiv 1 \tag{31}$$

Figure 3: Steady states with migration

$D(\cdot)$ *is decreasing on* $[0, x_D]$ *and increasing thereafter;*

(c) $x_D > \underline{x}(K)$ *for every* $K \geq 1$.

Proof. Item (a) is proved by direct computation of $D(0)$ and $D(x_A)$. For (b), we calculate the first derivative of D:

$$D'(x) = (1 - \delta)\frac{(1 - \delta) + \delta(s(x) - P'(x))}{[(1 - \delta) + \delta P(x)]^2} \qquad (32)$$

The numerator is increasing in its argument from $-\infty$ to 1, and is equal to zero at x_D, which proves (b). For (c), proof is trivial if $\underline{x}(K) = 0$. Otherwise observe that

$$\frac{\delta}{1 - \delta}\left(P'(x_D) - s(x_D)\right) \equiv 1 \leq K \underbrace{\equiv}_{\text{cf.}(20)} \frac{\beta}{1 - \beta}\left(P'(\underline{x}(K)) - Ks(\underline{x}(K))\right) < \frac{\delta}{1 - \delta}\left(P'(\underline{x}(K)) - s(\underline{x}(K))\right)$$

The fact that $P'(x) - s(x)$ is strictly decreasing in its argument proves the claim. ∎

The steady-state levels of learning intensity in both countries, x_N and x_S, is then determined by the intersection of the wage curve (25) and the demographic balance curve (29), which are illustrated by Figure 3. Since South is assumed to be less developed, migration occurs from North to South, decreasing x_S, the proportion of remaining in the South skilled to Southern unskilled.

17

3.4   Immiserizing immigration

As seen on Figure 3, when brain drain from the South becomes sufficiently large, and the Southern learning intensity x_S falls below x_D, further increases in the volumes of brain drain not increase but *decrease* the learning intensity in the North – a phenomenon I label as the *immiserizing immigration*, following Bhagwati's immiserizing growth (Bhagwati 1958). The immiserizing immigration is due to the fact that increasing openness of the South has two effects: (i) it increases the share of Southern skilled that migrate, r, and (ii) decreases the total number of Southern skilled, $m_S L_S$. The stock of Southern migrants in the North is the product of the two. With low migration rates, the first effect dominates, but as the cost of migration continues to decline, the second effect may overwhelm, reducing the number of skilled individuals not only in the South (the standard brain drain effect) but also in the North. In the most extreme case, when every Southern skilled emigrates, Southern unskilled have no one to learn from and thus there is no production of skill in the South, meaning that North has only as many skilled individuals as it can produce on its own.

3.5   Welfare effects of brain drain

To analyze the steady-state effects of a change of various model parameters, it is necessary to introduce the notion "welfare" in the proposed model. To fix ideas, by "welfare" I assume the value of being newly born in a steady state. In other words, does a change of a model parameter make, on average and in the long run, happier those individuals who have just entered this world? Since all newly-born are unskilled, it is sufficient to analyze the steady-state value of being unskilled, $v_0(w, K)$.

   With the results of Proposition 3, we can state the welfare effects of brain drain from South to North. The effects depend on relationship between Southern welfare-minimizing point $x^*(K_S)$ and immiserizing immigration point x_D.

Proposition 5 *A lowered migration cost M and an associated increase in migration rate r have the following welfare effects:*

In the South, welfare declines if initial migration rate is low such that $x_S \in (x^*(K_S), x_A]$, and increases if initial migration rate is high such that $x_S \in (\underline{x}(K_S), x^*(K_S)]$.

In the North, welfare increases if initial migration rate is low such that $x_S \in (x_D, x_A]$, and declines if initial migration rate is high such that $x_S \in (\underline{x}(K_S), x_D]$.

In other words, if initial migration rate is low, its increase benefits North and hurts South; the opposite happens if migration is high. The relationship between Southern welfare minimum point $x^*(K_S)$ and Northern welfare maximum point x_D depends on the value of K_S. Thus, if K_S is low, there may be a region when both North and South benefit from increasing brain drain ($x_S \in (x_D, x^*(K_S))$); a high K_S implies the existence of a region when both North and South lose from increased brain drain ($x_S \in (x^*(K_S), x_D)$).

Proof. The results for the South follow directly from the definition of $x^*(K_S)$, the welfare minimum point. The results for the North follow from the fact that $x_N > x_A > x^*(K_N)$, thus welfare increases with x_N, and from the fact that x_N is increased by Southern brain drain if $x_S \in (x_D, x_A)$, and is decreased otherwise. ∎

3.6 Reduction of transaction costs

It is intuitively obvious that lowered transaction costs K lead to increased welfare. Improvement of financial institutions, however, usually comes at a social cost, and it is therefore useful to know whether greater openness of a country to migration increases or decreases the social gain from lower transaction costs. For mathematical tractability of subsequent analysis, we assume henceforth that North is a large country and is not affected by migration patterns; then, (25) implies that Southern wage does not depend on Southern transaction costs K_S because it is pinned down by (and is inversely related to) migration costs M. This property enables us to interpret any change is Southern wage as the effect of an inverse change of migration costs, and to study the interaction between transaction and migration costs in the $K - w$ space.

Figure 4 plots the isowelfare lines of the newly born in the space of transaction costs K and wage w. The position of isowelfare lines can be predicted by Proposition 4 which states, among other things, the following: if transaction costs are high $K \geq P'(0)\frac{\beta}{1-\beta}$, the welfare minimum point is achieved at $x^*(K) = 0$ and thus at $w^*(K) = \infty$, i.e. welfare declines with increasing wage and the isowelfare curves must be downward sloping. With low transaction costs $K < P'(0)\frac{\beta}{1-\beta}$, minimum welfare is achieved at some intermediary $(x^*(K), w^*(K))$ such that $x^*(K)$ decreases with K from a positive finite value to zero, and thus $w^*(K)$ increases with K from a positive finite value to infinity. Isowelfare lines are downward sloping below $w^*(K)$ and are upward sloping above.

The following proposition helps to analyze the (horizontal) *distance* between isowelfare lines.
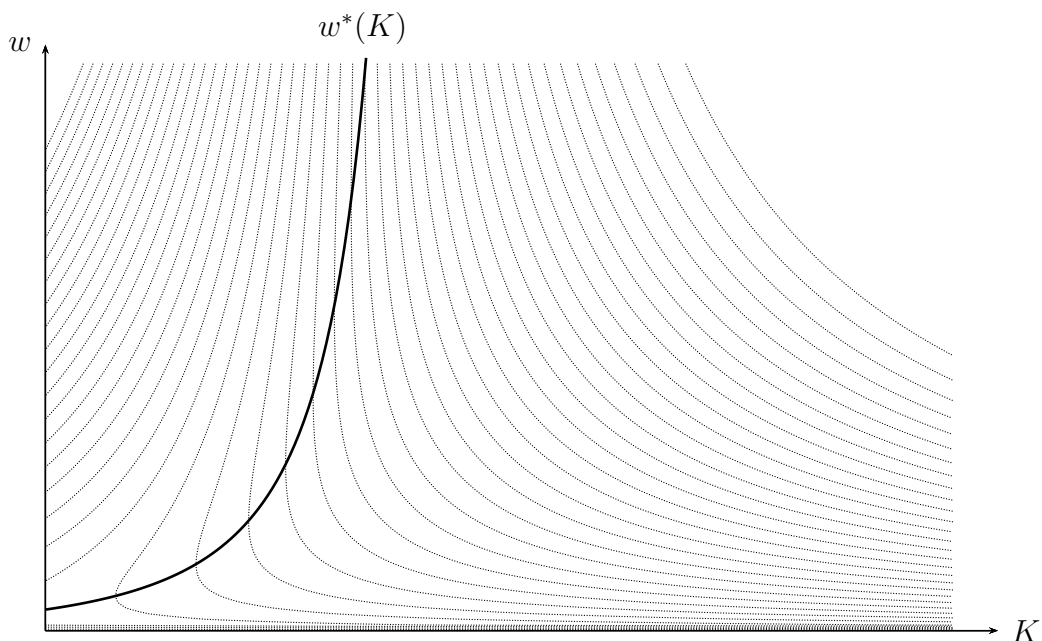
Figure 4: Isowelfare lines of newly born in the $K - w$ space

Proposition 6 *A greater openness of South to migration and an associated increased wage w, lead to a higher marginal gain from lowered transaction costs:*

$$\frac{d}{dw}\left|\frac{dv_0(w, K)}{dK}\right| = -\frac{d^2v_0}{dwdK} > 0$$

In other words, the horizontal distance between isowelfare curves diminishes as the wage rises: with higher wage due to greater openness, an equivalent reduction of transaction costs leads to a greater welfare gain (more isowelfare lines are crossed). Thus, as less developed South becomes more open, its government has a higher incentive to improve financial institutions.

Refer to the Appendix for the proof of the Proposition.

4 Conclusion

Modern civilization is only possible because people acquire most of their human capital from those who already possess it rather than accumulate it in isolation. While this phenomenon is well-studied at both theoretical and empirical levels, its logical extension that knowledge transfer from skilled to unskilled may be compensated by the latter, and that this compensation may positively affect the willingness to acquire skill, received little attention in the literature. In the present paper, I elaborate on this idea by developing a simple general

equilibrium model of skill acquisition, in which the unskilled individuals acquire skill only by interacting with the skilled, and by compensating the latter for knowledge transfer they generate. An exogenous increase in the demand for education, for example, due to lower costs of transaction between learners and teachers, has a multiplicative effect: it leads to an increase in the return to skill, which encourages the unskilled to acquire even more skill and boosts the demand for education even further.

I apply this idea to a model of international migration between countries with exogenously different ability of the unskilled to pay for their education, which leads to large differences in the return to skill and creates a basis for skilled migration from a less developed country to a more developed country. I find that an increased openness of the less developed country to such emigration may lead to an increase of welfare of the unskilled, despite the fact that they lose potential teachers, when transaction costs in the less developed country are sufficiently low. I also find that such increased openness increases economic payoffs to improvement of institutions.

Appendix

Optimal savings and "savings-free" optimization problem By assumptions of the model, the life of an individual is divided into two time periods: first, they are unskilled, have no income, and have to borrow to fund education; second, they become skilled and earn positive income. To balance the savings market, skilled individuals should (on average) thus make positive savings.

The optimization problem of a skilled individual is

$$\tilde{v}_1^t(A_{i,t}) = \max_{a_{i,t}} y_{i,t} - a_{i,t} + \beta \tilde{v}_1^{t+1}(a_{i,t} K^{I(a_{i,t}<0)} + (1+r_t)A_{i,t}) \tag{33}$$

subject to $a_{i,t} \leq y_{i,t}$ and to the transversality condition (5), where $\tilde{v}_1^t(A_{i,t})$ is the value of being skilled with assets $A_{i,t}$ at the beginning of period t, and $y_{i,t} = 1 + w_t$ is the income of a skilled individual. The first-order condition of the maximum is then

$$\beta \frac{d\tilde{v}_1^{t+1}}{dA_{i,t+1}} \begin{cases} < \frac{1}{K}, & a_{i,t} = a_t^{\min} \\ = \frac{1}{K}, & a_{i,t} \in \left(a_t^{\min}, 0\right) \\ \in \left[\frac{1}{K}, 1\right], & a_{i,t} = 0 \\ = 1, & a_{i,t} \in (0, y_{i,t}) \\ > 1, & a_{i,t} = y_{i,t} \end{cases} \tag{34}$$

where $a_t^{\min}$ is the minimum possible level of savings flow imposed by the transversality condition. The functional equation (33) can be shown to have the following solution:

$$\tilde{v}_1^t(A_{i,t}) = (1 + r_t)A_{i,t} + v_1^t \tag{35}$$

where v_1^t is defined by (6); thus, $\frac{d\tilde{v}_1^{t+1}}{dA_{i,t+1}}$ is equal to $1 + r_t$. To ensure that any skilled individual makes positive savings, it must be that $(1 + r_t)\beta \geq 1$, which also ensures that *every* skilled individual makes nonnegative savings. Further, skilled individuals ever enjoy positive consumption and the savings market clears only if

$$(1 + r_t)\beta = 1 \tag{36}$$

which implies that the equilibrium savings rate equals the discount rate.

The optimization problem of the unskilled individual is

$$\tilde{v}_0^t(A_{i,t}) = \max_{a_{i,t}, x_{i,t}} -a_{i,t} - w_t x_{i,t} + \beta \left[P(x_{i,t})\tilde{v}_1^{t+1}(A_{i,t+1}) + (1 - P(x_{i,t}))\tilde{v}_0^{t+1}(A_{i,t+1}) \right] \tag{37}$$

where $A_{i,t+1} = a_{i,t}K^{I(a_{i,t}<0)} + (1 + r_t)A_{i,t}$, and subject to

$$a_{i,t} + w_t x_{i,t} \leq 0 \tag{38}$$

and to the transversality condition (5). Note that the constraint (38) implies that the savings are non-positive (i.e. unskilled individuals borrow). The first-order conditions of the maximum are then

$$-1 \quad + \quad \beta K \left(P(x_{i,t})\frac{d\tilde{v}_1^{t+1}}{dA_{i,t+1}} + (1 - P(x_{i,t}))\frac{d\tilde{v}_0^{t+1}}{dA_{i,t+1}} \right) + \lambda_{i,t} = 0 \tag{39}$$

$$-w_t \quad + \quad \beta P'(x_{i,t}) \left(\tilde{v}_1^{t+1}(A_{i,t+1}) - \tilde{v}_0^{t+1}(A_{i,t+1}) \right) + \lambda_{i,t} w_t = 0 \tag{40}$$

where $\lambda_{i,t}$ is the Lagrange multiplier for the constraint (38). The functional equation (37) can be shown to have the following solution:

$$\tilde{v}_0^t(A_{i,t}) = (1 + r_t)A_{i,t} + v_0^t \tag{41}$$

where v_0^t is defined by (10). Therefore, we have that $\frac{d\tilde{v}_0^{t+1}}{dA_{i,t+1}} = \frac{d\tilde{v}_1^{t+1}}{dA_{i,t+1}} = 1 + r_t$, thus the first-order condition for savings, (39), simplifies to $\beta K(1 + r_t) + \lambda_{i,t} = 1$. Using (36), we conclude that $\lambda_{i,t} = 1 - K$. Using this finding and the equations (35) and (41), we can

22

rewrite the first-order condition for x, (40), as follows:

$$-Kw_t + \beta P'(x_{i,t}) \left(v_1^{t+1} - v_0^{t+1}\right) = 0 \tag{42}$$

which coincides with the first-order condition of the "savings-free" optimization problem described by (10). In other words, for the analysis of optimal educational decisions, the savings part of the model may be suppressed.

Proof of Proposition 2 Maximization of (16) with respect to learning intensity z results in the following first-order condition:

$$-Kw + \beta P'(x(w,K))(v_1(w) - v_0(w,K)) \begin{cases} = 0 & x > 0 \\ \leq 0 & x = 0 \end{cases} \tag{43}$$

where $x(w,K)$ is the $\arg\max$ of (16). Solving for $v_1 - v_0$, we get (all subsequent inequalities hold with equality if $x > 0$)

$$(v_1(w) - v_0(w,K)) \leq \frac{Kw}{\beta P'(x(w,K))} \tag{44}$$

Substituting the inequality for $v_1 - v_0$ back into (16) yields, after some rearrangement,

$$(1 - \beta)v_0(w,K) \leq -Kwx(w,K) + Kw\frac{P(x(w,K))}{P'(x(w,K))}$$

Given that (cf.(15)) $v_0(w,K) = v_1(w) - (v_1(w) - v_0(w,K)) \geq \frac{1+w}{1-\beta} - \frac{Kw}{\beta P'(x(w,K))}$, we have that

$$(1 + w) - Kw\frac{1 - \beta}{\beta P'(x(w,K))} \leq (1 - \beta)v_0(w,K) \leq Kw\left(\frac{P(x(w,K))}{P'(x(w,K))} - x(w,K)\right)$$

Rearranging, we get

$$\begin{aligned} \frac{1+w}{Kw} &\leq \frac{P(x(w,K))}{P'(x(w,K))} - z + \frac{1 - \beta}{\beta}\frac{1}{P'(x(w,K))} \\ &= \frac{\frac{1-\beta}{\beta} + P(x(w,K)) - P'(x(w,K))x(w,K)}{P'(x(w,K))} = \frac{1}{G(x(w,K))} \end{aligned}$$

where G is defined in (18). Solving for w, we get (17). ∎

Proof of Lemma 2 If w and K are such that $x(w, K) = 0$, proof is trivial since $v_0 = 0$. Otherwise, from (43), we have that

$$v_0(w, K) = v_1(w) - \frac{Kw}{\beta P'(x(w, K))} \tag{45}$$

From (15), we know that $v_1(w) = \frac{1+w}{1-\beta}$. From (17), we have that (the arguments of $x(w, K)$ are dropped for brevity) $Kw = (w + 1)\frac{P'(x)}{\frac{1-\beta}{\beta} + P(x) - P'(x)x}$, which allows us to rewrite (45) as follows:

$$
\begin{aligned}
v_0(w, K) &= (w + 1)\left[\frac{1}{1 - \beta} - \frac{1}{(1 - \beta) + \beta(P(x) - P'(x)x)}\right] \\
&= (w + 1)\frac{1}{1 - \beta}\frac{\beta(P(x) - P'(x)x)}{(1 - \beta) + \beta(P(x) - P'(x)x)} \\
&= v_1(w)\frac{P(x) - P'(x)x}{\frac{1-\beta}{\beta} + P(x) - P'(x)x} \\
&= v_1(w)\frac{s(x)}{\frac{1-\beta}{\beta} + s(x)}
\end{aligned}
$$

∎

Proof of Proposition 3 continued

$$
S(x) + (K - G(x))\frac{\partial S(x)}{\partial x}\left(\frac{\partial G(x)}{\partial x}\right)^{-1}
$$

$$
\underbrace{=}_{\text{cf.(18),(19),(22)}} \frac{s(x)}{\frac{1-\beta}{\beta} + s(x)} + \left(K - \frac{P'(x)}{\frac{1-\beta}{\beta} + s(x)}\right)\frac{\frac{1-\beta}{\beta}x}{\frac{1-\beta}{\beta} + P(x)}
$$

$$
= \frac{\left(\frac{1-\beta}{\beta} + P(x)\right)(P(x) - P'(x)x) - \left(K\frac{1-\beta}{\beta} + Ks(x) - P'(x)\right)\frac{1-\beta}{\beta}x}{\left(\frac{1-\beta}{\beta} + s(x)\right)\left(\frac{1-\beta}{\beta} + P(x)\right)}
$$

$$
= \frac{P\left(\frac{1-\beta}{\beta} + P(x) - P'(x)x\right) - \frac{1-\beta}{\beta}Kx\left(\frac{1-\beta}{\beta} + P(x) - P'(x)x\right)}{\left(\frac{1-\beta}{\beta} + s(x)\right)\left(\frac{1-\beta}{\beta} + P(x)\right)}
$$

$$
= \frac{P(x) - \frac{1-\beta}{\beta}Kx}{\left(\frac{1-\beta}{\beta} + P(x)\right)}
$$

Proof of Proposition 6 $\frac{\partial^2 v_0}{\partial w \partial K} = \frac{\partial}{\partial K} \frac{\partial v_0}{\partial w}$ can be obtained by differentiating (24) with respect to K. We get:

$$\frac{\partial}{\partial K}\left(S(x(w,K)) + (K - G(x(w,K)))\frac{\partial S(x(w,K))}{\partial x}\left(\frac{\partial G(x(w,K))}{\partial x}\right)^{-1}\right)$$

$$= (K - G(x))\frac{\partial}{\partial x}\left(\frac{\partial S(x)}{\partial x}\left(\frac{\partial G(x)}{\partial x}\right)^{-1}\right)\frac{\partial x}{\partial K} + \frac{\partial S(x)}{\partial x}\left(\frac{\partial G(x)}{\partial x}\right)^{-1} \qquad (46)$$

To prove that (46) is negative, it is sufficient to show that each of its two components is negative. Indeed, in the first component that consists of three multipliers, two multipliers are positive while the third is negative:

$$K - G(x) > 0$$

$$\frac{\partial}{\partial x}\left(\frac{\partial S(x)}{\partial x}\left(\frac{\partial G(x)}{\partial x}\right)^{-1}\right) = \frac{\partial}{\partial x}\left(\frac{\frac{1-\beta}{\beta}x}{\frac{1-\beta}{\beta} + P(x)}\right) = \left(\frac{1-\beta}{\beta}\right)\frac{\frac{1-\beta}{\beta} + P(x) - P'(x)x}{\left(\frac{1-\beta}{\beta} + P(x)\right)^2} > 0$$

We can compute $\frac{\partial x}{\partial K}$ from (17), using the implicit function theorem and holding w fixed: $\frac{\partial x}{\partial K} = \frac{G(x)}{K}\left(\frac{\partial G(x)}{\partial x}\right)^{-1} < 0$ due to $\frac{\partial G(x)}{\partial x} < 0$.

The second component of (46) is also negative because $\frac{\partial S(x)}{\partial x} > 0$ while $\frac{\partial G(x)}{\partial x} < 0$.

References

BAUER, T., G. EPSTEIN, AND I. GANG (2007): "The Influence of Stocks and Flows on Migrants' Location Choices," *Research in Labor Economics*, 26, 199–229.

BEINE, M., F. DOCQUIER, AND H. RAPOPORT (2001): "Brain drain and economic growth: theory and evidence," *Journal of Development Economics*, 64(1), 275–289.

——— (2008): "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers.," *Economic Journal*, 118(528), 631 – 652.

BHAGWATI, J. (1958): "Immiserizing Growth: A Geometrical Note," *The Review of Economic Studies*, 25(3), pp. 201–205.

CHAND, S., AND M. CLEMENS (2008): "Skilled emigration and skill creation: A quasi-experiment," *Crawford School of Economics and Government working papers*.

25

COMMANDER, S. J., M. KANGASNIEMI, AND L. A. WINTERS (2003): "The Brain Drain: Curse or Boon?," *SSRN eLibrary*.

FLUG, K., A. SPILIMBERGO, AND E. WACHTENHEIM (1998): "Investment in education: do economic volatility and credit constraints matter?," *Journal of Development Economics*, 55(2), 465 – 481.

GALOR, O., AND J. ZEIRA (1993): "Income Distribution and Macroeconomics," *The Review of Economic Studies*, 60(1), 35–52.

HANSON, G. H. (2009): "The Economic Consequences of the International Migration of Labor," *Annual Review of Economics*, 1(1), 179–208.

HAQUE, N. U., AND S.-J. KIM (1995): ""Human Capital Flight": Impact of Migration on Income and Growth," *Staff Papers - International Monetary Fund*, 42(3), 577–607.

JOVANOVIC, B., AND R. ROB (1989): "The Growth and Diffusion of Knowledge," *The Review of Economic Studies*, 56(4), 569–582.

KELLER, W. (2002): "Geographic Localization of International Technology Diffusion," *The American Economic Review*, 92(1), 120–142.

——— (2004): "International Technology Diffusion," *Journal of Economic Literature*, 42(3), 752–782.

LUCAS, R. E. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22(1), 3 – 42.

MEJIA, D., AND M. ST-PIERRE (2008): "Unequal opportunities and human capital formation," *Journal of Development Economics*, 86(2), 395 – 413.

MOUNTFORD, A. (1997): "Can a brain drain be good for growth in the source economy?," *Journal of Development Economics*, 53, 287–303(17).

MUNSHI, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 118(2), 549–599.

PARK, K. S. (1997): "A Theory of on-the-Job Learning," *International Economic Review*, 38(1), 61–81.

SCHADY, N. R. (2004): "Do Macroeconomic Crises Always Slow Human Capital Accumulation?," *The World Bank Economic Review*, 18(2), 131–154.

STARK, O., C. HELMENSTEIN, AND A. PRSKAWETZ (1997): "A brain gain with a brain drain," *Economic Letters*, 55, 227–234.

STARK, O., AND Y. WANG (2002): "Inducing human capital formation: migration as a substitute for subsidies," *Journal of Public Economics*, 86(1), 29 – 46.

STARK, O., AND R. ZAKHARENKO (2012): "Differential Migration Prospects, Skill Formation, and Welfare," *Review of International Economics*, forthcoming.

VERGALLI, S. (2006): "The Role of Community in Migration Dynamics," *SSRN eLibrary*.

WINTERS, P., A. D. JANVRY, AND E. SADOULET (2001): "Family and Community Networks in Mexico-U.S. Migration," *The Journal of Human Resources*, 36(1), 159–184.