Migration, Learning, and Development

Zakharenko, Roman

Pennsylvania State University

November 2007

Migration, Learning, and Development

Roman Zakharenko*

November 9, 2007

Abstract

US-educated Indian engineers played a major role in the establishment of the "Silicon Valley of Asia" in Bangalore. The experience of India and other countries shows that returning well-educated emigrants, despite their small numbers, can make a difference. This paper builds a model of "local" knowledge spillovers, in which migration of a small number of highly skilled individuals greatly affects country-level human capital accumulation. All economic activity occurs in pairs of individuals randomly matched to each other. Each pair produces the consumption good; the skills of the two partners are complementary. At the same time, the less skilled partner increases human capital by learning from the more skilled colleague. With poor institutions at home, highly skilled individuals leave the country seeking better opportunities abroad. On the contrary, improved institutions foster return migration of emigrants who have acquired more knowledge while abroad. These return migrants greatly amplify the positive effect of better institutions.

JEL classifications: O15, F22, J61, C68, C78

*Department of Economics, Pennsylvania State University.
Email: r.zakharenko@gmail.com. I thank Edward Green, Vijay Krishna, James Tybout, and Neil Wallace for their helpful suggestions. I am especially indebted to Barry Ickes for his guidance of my work

1 Motivation

High-skilled emigrants returning home can make a difference. Saxenian (2006) describes how the rapid growth of the information technology industry in Israel, India, Taiwan, and later mainland China was tightly related to return migration of Israeli, Indian, and Chinese high-skilled engineers living in the US, mainly in the Silicon Valley. These engineers used their US experience to start new businesses at home, train local employees, and enter the global market with their new products.

Recent economic models of growth and development can do little to explain such rapid productivity growth on such a large scale. There exist models of brain drain (for example, Haque and Kim 1995) and return migration (such as Dos Santos and Postel-Vinay 2003) which are based on the assumption that average human capital of the previous generation has a positive effect on human capital acquisition by the new generation. Emigration of the highly-skilled reduces average human capital in the country, thus reducing human capital of future generations. Likewise, return of the highly-skilled increases average human capital, which has a positive externality on the young people.

Empirically, however, the number of returning high-skilled emigrants is usually too small to have any significant effect on average human capital. Few hundreds of Indian talented engineers cannot change the average human capital of the Indian labor force with its half-billion people; they can only change the 99th percentile of the human capital distribution. There has to be another mechanism, in which people at the top of human capital distribution play a much greater role than those at the bottom.

Another problem is to explain the incentives of return migrants. It is relatively easy to create a model of one-way migration, but it proved more difficult to explain why a person who migrated from one country to another decides to reverse his decision after a while. The existing literature tends to explain return by "homesickness": although they are more productive abroad, some emigrants choose to return home after accumulating enough knowledge or wealth.[1]

One more problem is to explain why emigrants do return to some home countries, and don't return to others. Borjas and Bratsberg (1996) in their empirical cross-country study conclude that the probability of a US immi-

[1] Alternatively, some papers assume that individual productivity is exogenously lower if he works abroad, which creates return incentives

grant returning home positively depends on GDP per capita at home, which is easy to explain: wealthier countries typically have lower crime and provide better public goods, and therefore more attractive for living. Yet, the above mentioned countries (India, Taiwan, mainland China, Israel) did not have high income levels and good infrastructure from the start, but still experienced significant high-skilled return migration. An explanation of this fact may come from another finding of Borjas and Bratsberg (1996): they show that the "communist country" dummy is highly negatively significant,[2] which suggests that home country institutions may be an important factor affecting migration decisions.

In this paper, I propose a model of "local" knowledge spillovers. Instead of assuming that a high-skilled individual has a *small* positive externality on *all* young individuals by increasing average human capital, I assume that such an individual has a *large* positive externality on *someone*, and no immediate effect on the rest of population. For example, Amartya Sen returning to India would have a far greater influence on his immediate colleagues and students than on illiterate people living in remote villages.

After a while, those who learn from high-skilled returnees become high-skilled themselves, which enables them to train more unskilled individuals. Thus, the number of individuals with high human capital increases exponentially. With this "bootstrap" training technology, a small number of new high-skilled individuals may lead to a major shift of the human capital distribution over time.

Both production and learning occur in partnerships which consist of two individuals, randomly matched to each other. The amount of output they produce is a complementary function of their human capitals; they divide their joint output according to a bargaining rule. At the same time, the less skilled individual (the "apprentice") learns from the more skilled partner (the "master").[3] Due to skill complementarities of the two partners in production, the opportunity cost of such education is wasted talent of the master, in terms of current production. Obviously, in order to learn, the apprentice should compensate this wasted talent by accepting a lower, or even negative, share

[2]Their data was collected in the 1970's, at the peak of the cold war

[3]The terms "master" and "apprentice" are used here because learning occurs simultaneously with production. However, the "apprenticeship" here is somewhat different from its original meaning – medieval apprentices were bonded to their masters until they pay off their debt, while in my model they typically consume out of their own savings and have no financial obligations

of their joint output. The two partners may choose to split anytime; then, they are randomly matched to new partners after a waiting period.

The time needed to find a new partner may exogenously differ across countries, and it serves as a proxy for institutional quality in the model. In countries with high corruption and bureaucracy, starting a new business typically requires much more time and money; it is widely believed that these entry costs have a significant impact on country development. For example, the startup cost is one of the components of business environment indicators constructed by the World Bank and by the World Economic Forum. I this paper, I show that such entry cost differences alone may lead to major differences in income levels. The higher entry barriers affect bargaining over output: highly skilled people get a lower reward for training their low-skilled partners. Also, individuals of different skills match to each other less optimally, which results in a lower distribution of human capital. When emigration is allowed, some people with high skill emigrate from countries with poor institutions, which further lowers the human capital distribution at home; nobody wants to return.

When the home country improves its institutions, it drastically changes migration patterns: people with average human capital emigrate to acquire knowledge abroad, and return once their human capital becomes high enough. As a result, the home country grows three times faster than it would grow without return migration, despite the fact that the number of returnees per year is only about 0.1% of home country population.

In my model, return migration is a perfectly rational choice even without “homesickness”or exogenous productivity differences. When enough human capital is acquired, it may become optimal to return, because high skill is endogenously rewarded better in countries with scarce skill but good institutions.

2 The model of closed economy

2.1 Individuals

This is a one-sector dynamic model set up in discrete time. The economy is populated by a continuum of individuals of a finite mass. Each individual i at each point in time t is endowed with human capital $h_{i,t} \in [0, \infty)$ which evolves endogenously.

For each individual, there exists a small probability δ of death at each moment of time; for simplicity, death rate does not depend on individual's age. The same number of new people is born; their (initial) human capital is zero. As a result, the country population remains constant.

All individuals have identical risk-neutral preferences over the only consumption good:

$$U_i = \sum_{t=t_{0,i}}^{\infty} \beta^{t-t_{0,i}} c_{i,t}$$

where β is the discount factor, $c_{i,t}$ is consumption, and $t_{0,i}$ is the birth date of individual i. The date of death is uncertain; the probability of death is built into the discount factor β.

Due to complete credit markets, people can borrow and save. Assuming the interest rate equals the discount rate, people are indifferent between having more consumption today and more consumption tomorrow due to their linear preferences; they simply maximize their discounted stream of earnings. As a result, there is no need to model savings explicitly.

2.2 Production and learning

The production of the good occurs in partnerships; each partnership consists of two individuals. The only inputs in production are the human capitals of the two partners. When individuals i and j work together, they produce

$$y(h_i, h_j) = \min\{h_i, h_j\} \tag{1}$$

Note there exists a complementarity between human capitals of the two partners.[4]

The evolution of an individual's human capital depends on her partner's human capital. Suppose an apprentice with human capital h_1 works with a master with human capital h_2 (which implies $h_1 \leq h_2$). Then, the next period human capitals are

$$\begin{aligned} h_1' &= h_1 + g(h_2 - h_1) + \lambda_0 \\ h_2' &= h_2 + \lambda_0 \end{aligned} \tag{2}$$

The master's knowledge increases at a small rate λ_0 which reflects "learning from experience". Apprentice's knowledge increment is much higher and

[4]Generally, any production function with complementary inputs can be used – for example, O-ring production function used by Kremer (1991).

depends on master's knowledge. If an individual does not have a partner, he also increases his human capital at rate λ_0.

I assume the following properties of the learning function $g(\cdot)$:

- $g(0) = 0$ — no learning from an equal partner
- $g'(0) = \lambda$ with $\lambda \in (0, 1)$ — if the master is just slightly smarter than the apprentice, the latter reduces the knowledge gap by fraction λ each period
- $g''(x) < 0$ for all $x \geq 0$ — marginal returns from a smarter master are diminishing

Throughout the paper, I use the following learning function:

$$g(x) = \log(1 + \lambda x) \tag{3}$$

It satisfies all the properties mentioned above.

Note that in the absence of learning from a partner ($\lambda \equiv 0$) the Pareto-optimal allocation is to match individuals of as close as possible skill because of the production complementarity.

2.3 Matching

At the beginning of each period, most individuals are matched to a partner, but some are unmatched. A randomly chosen fraction θ of the unmatched individuals are randomly matched to each other; the remaining fraction $1 - \theta$ stays unemployed this period. The parameter θ serves as a proxy for institutional quality in a country. With higher θ, the unmatched individuals have more frequent opportunities to form a new partnership.

Individuals coupled to each other (both previously matched and newly matched) can decide whether to work together or split and remain unemployed this period. Those who work need to decide how they divide their joint revenue — this decision is made according to Nash bargaining rule (see below). The apprentice's share may be even negative in equilibrium, which implies some sort of tuition for education.

If the two partners decide to stay together, most likely they will be matched to each other again. For most couples, this is beneficial: it enables them to form long-term relationships, the apprentice can acquire most of master's knowledge. Some couples, however, are worse off from being

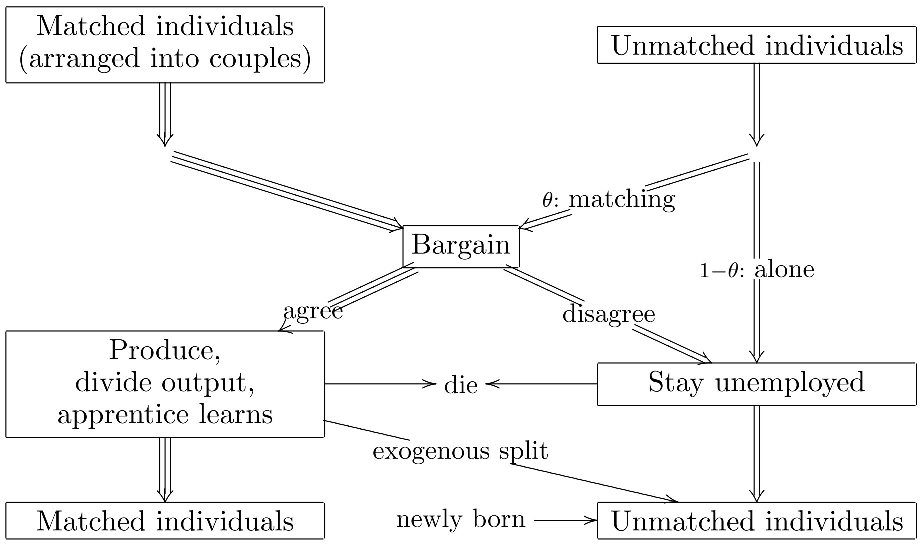

Figure 1: Timing in closed economy. Arrow thickness indicates fraction of population following this path in a typical steady state

matched to each other again: they would prefer to be matched to new randomly chosen partners. By assumption, changing a partner requires at least one period of unemployment; as a result, partnerships last longer than they would in the absence of search frictions. With poor country institutions (low θ), establishing a new partnership takes more time, which makes people reluctant to shut down existing partnerships, and therefore making them less efficient.

Although working partners are usually matched to each other again and again, there exists a small probability that they are forced to join the pool of unmatched people. This happens than one of the partners dies; I also assume that a small fraction of couples are forced to split exogenously, even when both partners survive.[5]

[5]The reason for introducing the exogenous split is technical: when some individuals are forced to enter the job market, the distribution of skill on the job market becomes more stable, which greatly improves the numerical algorithm

2.4 Bargaining

A couple of partners divides their joint output according to Nash bargaining rule. Each potential partner i calculates his reservation wage $\underline{w}_t(h_i, h_j)$, which makes him indifferent between staying with current partner j, and remaining unemployed this period and meeting a new partner tomorrow:

$$\underline{w}_t(h_i, h_j) + \beta[(1-\gamma)V^m_{t+1}(h'_i, h'_j) + \gamma V^u_{t+1}(h'_i)] \equiv \beta V^u_{t+1}(h_i + \lambda_0)$$

where $V^m_t(h_i, h_j)$ is the value of i being matched with j at time t, V^u_t is the value of being unmatched, γ is the probability of exogenous split, h'_i and h'_j are future human capitals of i and j if they work together.

Then, the surplus created by the match is the difference between the joint output of i and j, and the sum of their reservation wages:

$$y(h_i, h_j) - (\underline{w}_t(h_i, h_j) + \underline{w}_t(h_j, h_i))$$

If the surplus is non-negative, i and j stay together; otherwise they split. Since the possibility frontier is linear, Nash bargaining implies that each partner earns his reservation wage plus half of the surplus, hence i's wage when working with j is

$$w_t(h_i, h_j) = \underline{w}_t(h_i, h_j) + \frac{1}{2}\left(y(h_i, h_j) - (\underline{w}_t(h_i, h_j) + \underline{w}_t(h_j, h_i))\right)$$

Since the individuals divide the surplus equally, their accept-reject decisions (whether to stay together or split) are always synchronized: either both partners choose to be together, or both of them choose to split.

2.5 Equilibrium and steady state

The equilibrium in this model consists of the following:

- Distributions of individuals across types, at every moment of time: $f_1, f_2, ..., f_t, ...$ with

  $$f_t = \{f^m_t(h_i, h_j), f^u_t(h_i)\}$$

  where f^m_t describes the density of individuals of type i matched with those of type j, and f^u_t describes the density of unmatched individuals

- Path of wages, or bargaining outcomes, for every potential couple of partners: $w_1, w_2, ..., w_t, ...$

- Values associated with each state, at every moment of time: $V_1, V_2, ..., V_t, ...$ where

$$V_t = \{V_t^m(h_i, h_j), V_t^u(h_i)\}$$

  These values are defined as follows. Define $V_t^{in}(h_i, h_j)$ as the value of i working with j, and $V_t^{out}(h_i)$ as the value of being unemployed:

$$V_t^{in}(h_i, h_j) = w_t(h_i, h_j) + \beta[(1-\gamma)V_{t+1}^m(h_i', h_j') + \gamma V_{t+1}^u(h_i')]$$
$$V_t^{out}(h_i) = \beta V_{t+1}^u(h_i + \lambda_0)$$

  Then, the values of being matched and unmatched are

$$V_t^m(h_i, h_j) = \max\{V_t^{in}(h_i, h_j), V_t^{out}(h_i)\} \quad (4)$$
$$V_t^u(h_i) = \theta \frac{\int V_t^m(h_i, h_j) f^u(h_j) dh_j}{\int f^u(h_j) dh_j} + (1-\theta)V_t^{out}(h_i) \quad (5)$$

In a steady state, all objects mentioned above are time-invariant. The rest of this paper, except the last section, deals with computing and analyzing steady states.

2.6 Results

As many models with heterogenous agents, this model is too complex for analytical analysis. I solve the model numerically, using parameter values described below. I consider two scenarios: the "North" (a closed economy with good institutions) and the "South" (an economy with poor institutions).

Variable	Notation	Value	Comment
Model period, frequency of matching		2 months	
Discount factor	β	0.99	people discount future at about 6% per year
Death rate	δ	0.0025	life expectancy is about 67 years
Probability of exogenous split	γ	0.01	
Speed of learning by experience	λ_0	0.0025	without learning from others, human capital increases by 1 during average lifetime
Speed of learning from others	λ	0.02	knowledge gap between master and apprentice is reduced by at most 12% per year
Probability of being matched with new partner	θ	1 (North); 1/6 (South)	Unmatched people in the South wait on average for 1 year for a match, compared to 2 months in the North

The computational procedure of finding steady states is described in appendix A.

Due to higher entry barriers, it is harder to start a partnership in the South. As a result, Southern individuals are less careful when choosing a partner, and more reluctant to terminate inefficient partnerships, which results in a lower distribution of human capital in the long run. Figure 2 shows the steady-state distribution of human capital in the North and in the South. In both countries, the distribution peaks at zero — there is a mass point of newly born individuals; then, the distribution peaks again near the highest available human capital — because it is relatively easy to reach the frontier of knowledge by learning from others, but very hard to go beyond that frontier. I have no formal proof that the steady state is unique; however, in all my experiments with different initial distributions, the system has converged to the same steady state.

Individual wage $w(h_i, h_j)$, obviously, increases with own human capital:

$$\frac{dw(h_i, h_j)}{dh_i} > 0$$

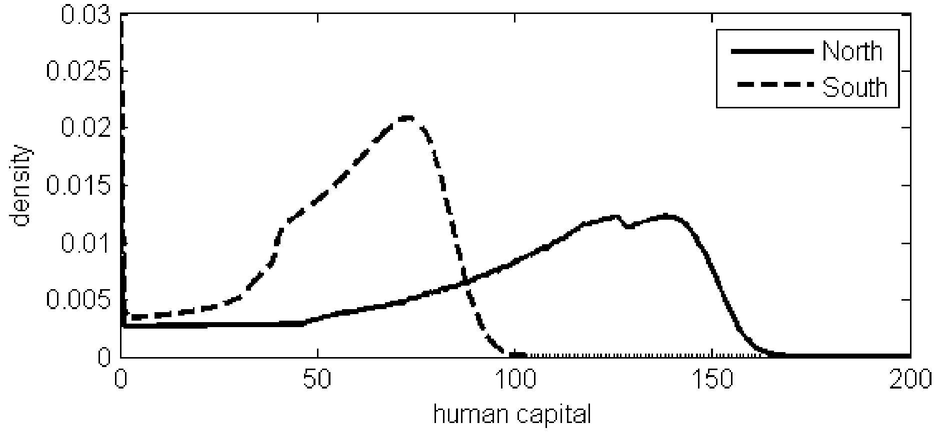

Figure 2: Steady-state distribution of human capital

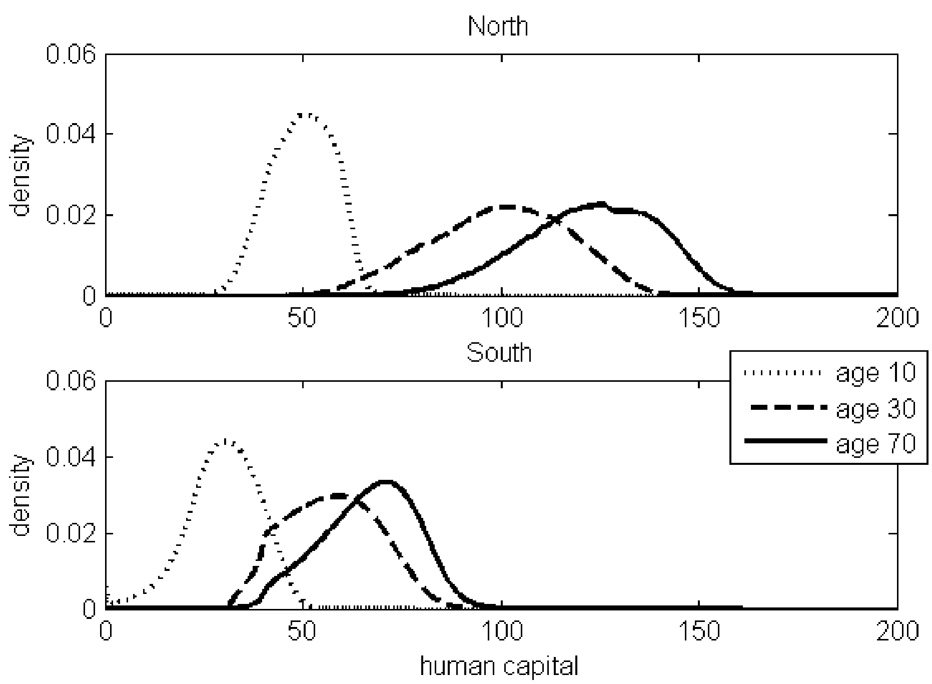

Figure 3: Distribution of human capital, conditional on age of individuals

The dependence of wage on partner's human capital h_j is less trivial. It is always true that the two equal partners ($h_j = h_i$) would divide their joint output equally. Otherwise, the wage depends on model parameters, but some common patterns can be traced. I consider two distinct cases: less skilled partner ($h_j < h_i$) and more skilled partner ($h_j > h_i$).

When j is more skilled, today's output $y(h_i, h_j) = \min\{h_i, h_j\}$ does not depend on h_j, so the amount of wealth to be divided does not change as h_j increases. A higher h_j, however, implies that i would learn faster and be able to earn more tomorrow, therefore i agrees to accept lower wages today. This results in a negatively sloped wage function:

$$\frac{dw(h_i, h_j)}{dh_j} < 0 \qquad \text{when} \qquad h_j > h_i$$

This property, combined with the fact that $w(0,0) = 0$, implies that an individual with zero human capital earns a negative income as long as partner's skill is positive.[6]

When j is less skilled than i, there are two effects of increased h_j. First, since now the output is determined by j's skill, the amount of wealth to be divided increases as h_j increases; both partners, i and j, benefit from it. Second, increased h_j means a smaller knowledge gap between i and j, and therefore less learning occurs, which lowers the reward of the master i. The first effect, increasing i's wage, dominates when h_j is small; the second effect, decreasing i's wage, dominates when h_j gets close to h_i. As a result, for every master with human capital h_i there exists an "optimal" apprentice who provides the highest income for i.

The effect of poor institutions (lower θ) is worse outside opportunities of both bargaining parties. Since the low-skilled individuals have low outside opportunities anyway, they have a relatively stronger bargaining position, and get a higher fraction of output. As a result, learning from a higher partner is cheaper in the South, where institutions are poor; conversely, teaching lower-skilled individuals is rewarded better in the North. This discrepancy creates a basis for South-North high-skilled emigration, when migration be-

[6] I have considered a version of the model with borrowing constraints, when young individuals can only learn if they have enough initial wealth. In this setting individuals are characterized by two state variables: knowledge and wealth, thus the matches of individuals are defined in four-dimensional state space. This model was abandoned due to excessive numerical complexity

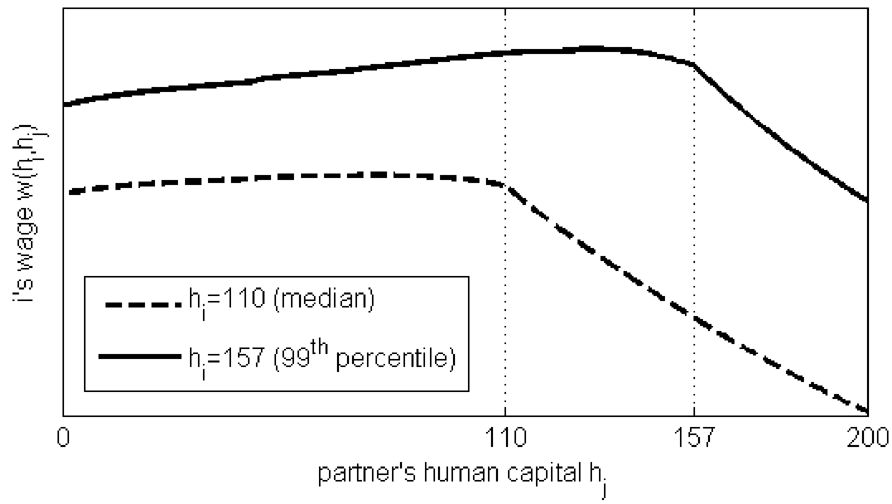

Figure 4: Wage in the North, as a function of partner's human capital

comes available. Figure 5 compares wages in the North and in the South for an individual of a given skill.

Figure 6 shows optimal accept-reject decisions in the steady state, in the North. Individuals do not work together if their human capitals are equally low – because no learning occurs when the partners have equal human capital. Individuals with low human capitals also do not match with those who have very high human capital – because the learning function is concave, it is better to match low-h apprentices with average-h masters. In the South, the opportunity cost of a match is much higher; this means a wider white area on a similar Southern graph.

In the South, the autarky GDP per capita is about 54% of the Northern value.

3 One-way migration and brain drain

When modeling migration between North and South, a number of additional assumptions are made.

- The North is a large country: migration has no effect on its steady state

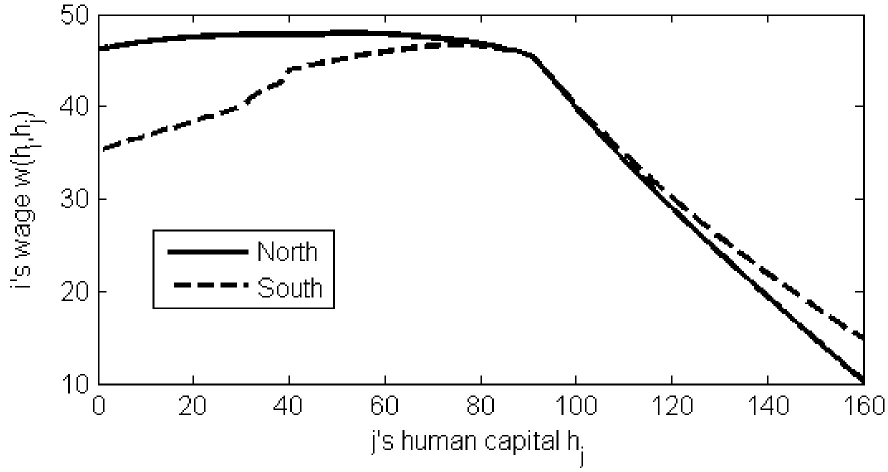

Figure 5: Wage of an individual with $h_i = 91$ (Southern 99^{th} percentile), as a function of partner's human capital

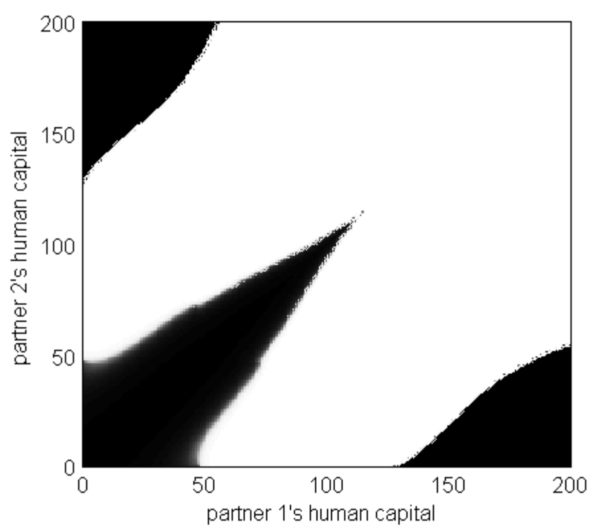

Figure 6: Accept-reject decisions of Northern randomly matched partners: stay together in white area, split in black area

- The Northerners always live in their home country; only Southerners migrate. With this assumption, the South is not flooded with Northerners once the value of living in the South gets high
- Migration is available at the end of each period, and only for unmatched individuals
- Migration is instantaneous; the migrants join the unmatched pool at the new location
- The number of people born each period in the South is constant; it does not depend on migration patterns. This assumption allows me to define a steady state with migration
- The North restricts immigration: the number of Southerners living in the North cannot exceed 5% of Southern population at any time. The Northern government imposes an emigration fee, which makes this restriction incentive compatible

This last assumption is needed to prevent too much South-North migration. In practice, developed countries do restrict immigration, and only a small fraction of the developing world population is able to emigrate.[7]

Assuming that South retains its poor institutions, I show below that migration occurs only in one direction: South to North; emigrants never return. This makes Southerners living in the North identical to Northerners themselves; their value of living in the North is identical to that of Northern-born population. This allows us to introduce migration into the model in a cheap way: emigration simply becomes an outside opportunity; to find the steady state, there is no need to track the history of emigrants.

In the new steady state, all Southerners benefit from emigration. Due to Northern immigration restrictions, only a few of them, those with the highest incentives, emigrate. Surprisingly, the emigrants have high, but not the highest skill; figure 7 shows that emigration incentive peaks around $h = 51$ (90th percentile of Southern human capital), and declines between 51 and 70 (Southern highest human capital). As a result, individuals with human capital around 51 offer the highest emigration fee, and only they actually

[7] According to the study of Docquier and Marfouk (2004), the number of immigrants in the OECD countries does not exceed 60 million people (this number includes migrants from one OECD country to another), which is only about 1% of the world population

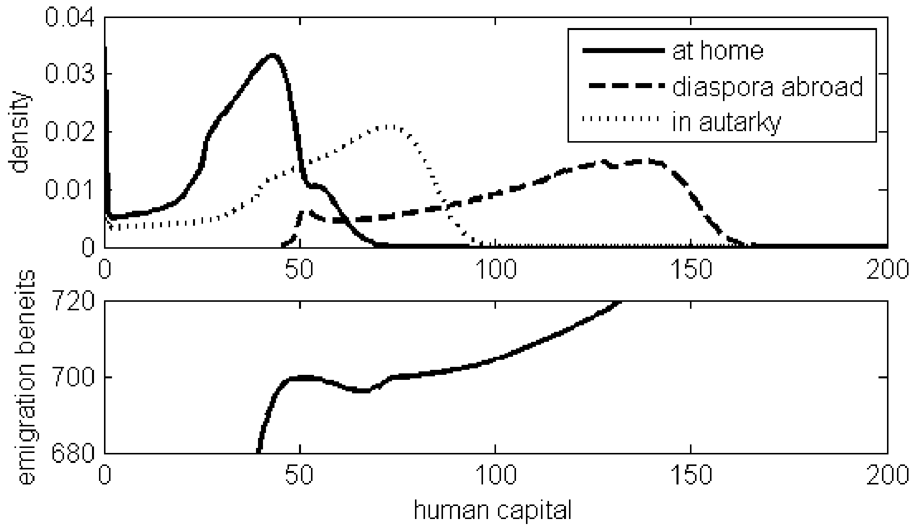

Figure 7: Southern human capital distribution (upper graph) and benefits of emigration (lower graph), in the steady state with one-way migration

emigrate. Their emigration causes a depression of human capital distribution at 51; in the long run, because the emigrants do not pass their knowledge onto young Southerners, the human capital distribution deteriorates compared to the autarky scenario. Figure 7 also confirms that no emigrant wants to return: everyone's emigration benefit is far above zero.

Why is emigration benefit non-monotonic? There are several factors that affect migration incentives. One factor is the difference in bargaining solutions, $w(h_i, h_j)$, between North and South. In the South, everyone has bad outside opportunities, which makes bargaining less dependent on human capitals. High-skilled Southerners generally earn a lower reward for their skill, which makes them more willing to migrate — emigration benefit is generally upward sloping. On the other hand, it pays off to be the "king of the hill" in a country — having slightly more human capital than anybody else in the economy slightly increases the reward because such an individual is basically a monopolist possessing a scarce resource. Consequently, Southerners with a very high (by Southern standards) skill are slightly less willing to emigrate and make a slightly lower bid to purchase the right to do so.

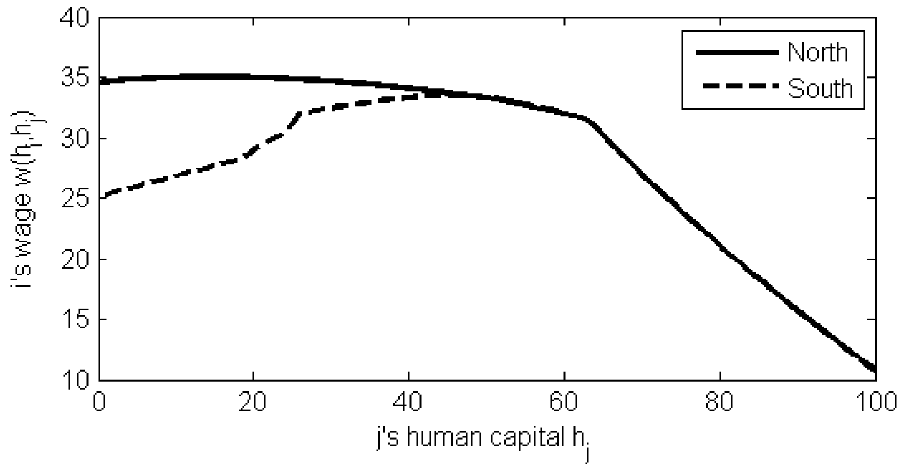

Figure 8: Wage of an individual with $h_i = 63$ (new Southern 99^{th} percentile), as a function of partner's human capital, in the steady state with one-way migration

Figure 8 compares wages at home and abroad, for the most talented Southerners ($h = 63$). Teaching individuals of very low skill is still more beneficial in the North; however, the reward for teaching those only slightly below the top ($h_i \in (50, 63)$) is nearly identical and, in fact, slightly higher in the South. This effect keeps the most smart Southerners at home.

A natural question arises: if individuals of modest skill emigrate, and no one lives forever, then who are the people at the top of human capital distribution and where did they come from? The explanation comes from the fact that people face idiosyncratic shocks: if an individual from the "emigration range" of skills was suddenly left without a partner, he emigrates; if such an individual was learning from a top-skilled master, the former does not emigrate and jumps over the emigration range. Once individual's skill rises above the emigration range of skills, he stays in the South for good.

What happens to Southern emigrants abroad? Because they work with more educated Northerners, the emigrants learn a lot while they live in the North; their skills grow far beyond the Southern knowledge frontier. Their return would have a great effect on the Southern human capital distribution; however, they have no interest in returning.

Due to emigration of the highly skilled, the GDP per capita among Southerners decreases down to 63% of the Southern autarky level. Obviously, the joint income of Southerners at home and Southerners abroad is higher, but still only 70% of its autarky level. Thus, we may conclude that the brain drain hurts the Southern economy. This result confronts Mountford's (1997) idea that emigration possibility increases learning incentives and thus may be beneficial for home country.

I have tested the one-way migration steady state with different values of θ (institutional quality). As long as the Southern institutions are worse than that of the North, one-way emigration is incentive-compatible in the steady state: nobody wants to return.[8] With improved institutions, the "king of the hill" effect gets stronger: the bargaining position of the highly-skilled individuals improves, and they get a better reward for training those below them. Conversely, with worse institutions the "king of the hill" effect weakens until it totally disappears: when institutions are bad enough, the very best people emigrate. I have computed the *selection rate* defined as the ratio of the emigrant's average skill to average skill of all Southerners.[9] In experiments, it is inversely related to the quality of institutions: as the institutional quality improves from $\theta = 1/12$ to $1/6$ (benchmark) to $1/3$, the selection rate decreases from 1.623 to 1.439 to 1.112.

This negative relationship is supported by the data. Docquier and Marfouk (2004) provide data on the stock of migrants from most world countries to the OECD countries, disaggregated by three levels of skill (low, medium, high). Based on this data, I calculate the selection rate, by country of migrants' origin, as the fraction of the highly-skilled among emigrants, divided by the fraction of the highly-skilled among all workforce at home. As a proxy for institutional quality, I use the "government effectiveness" from the cross-country dataset constructed by Kaufmann, Kraay and Mastruzzi (2006). They define the government effectiveness as

> the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the governments commitment to such policies

I regress the selection rate on the government effectiveness and a couple of control variables; the results are shown in the table below.

[8] Return migration only may exist during the transition from one steady state to another

[9] Emigrant's skill is measured at the moment of migration

Dependent variable: selection rate

Explanatory variables	Value	Std.err.	t-stat.
Constant	6.1685	0.9115	6.7677
Govt. effectiveness	-3.5315	0.7823	-4.5145
Workforce at home (mln)	0.0132	0.0117	1.1306
Landlocked country dummy	6.3973	1.9167	3.3377

The effect of the government quality on the selection rate is negative and significant, which supports the predictions of the model.

4 Improved institutions and return migration

4.1 The story

What happens if the South improves its institutions to the Northern level? In the rest of the paper, I study the effects of unexpected institutional improvement on the Southern economy. To isolate the effect of return migration, I compare two scenarios: return migration is allowed and free; return migration is prohibited. In both scenarios, I assume that the Northern government preserves its 5% quota on immigrants at any moment of time.

The long-run effect of the institutional improvement is trivial: under both scenarios, the Southern economy converges to the Northern steady state.[10] The interesting question here is the *speed* of adjustment, which appears to be drastically different. Approximating the speed of adjustment, however, requires to calculate the equilibrium transition path. Appendix B describes the computational strategy.

4.2 Results

No return migration

The institutional improvement has instantaneous effect on bargaining. Now, changing partners becomes easy; high-skilled individuals ask a higher reward for teaching. The benefit of emigration drastically decreases; for people with very high human capital it becomes negative, which means that highly-skilled emigrants would return if they could (figure 9 demonstrates new migration incentives). The emigration pattern changes; the emigrants have lower skill

[10] At least, no "poverty traps" have been discovered

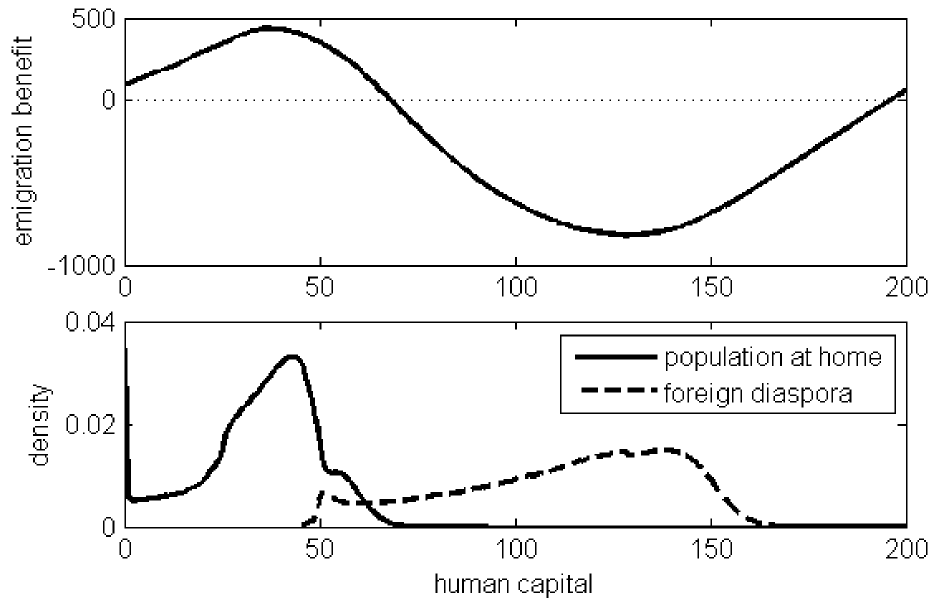

Figure 9: Benefits of emigration from the South, depending on human capital, immediately after the reform. Negative benefit implies willingness to return. Lower figure shows Southern human capital distributions at the moment of reform

than before: about 37 (the Southern median skill), compared to 51 (about 90th percentile) before the reform. The new emigration pattern doesn't hurt the Southern economy as much. The 5% emigration quota is still fully used; the flow of emigrants each period does not change.

Return migration possible

With poor Southern institutions, emigrants left for good; while living in the North, they acquired a lot of human capital. Figure 9 shows that highly-skilled emigrants are better off from returning home, when institutions improve. Intuitively, high skill is scarce in the South; with efficient institutions, highly (by Northern standards) skilled emigrants can earn more by returning and training highly (by Southern standards) skilled locals. Figure 10 compares wages of a smart individual at home and abroad, depending on partner's skill.

As a result, in the first year following the reform, about 10% of emigrants,

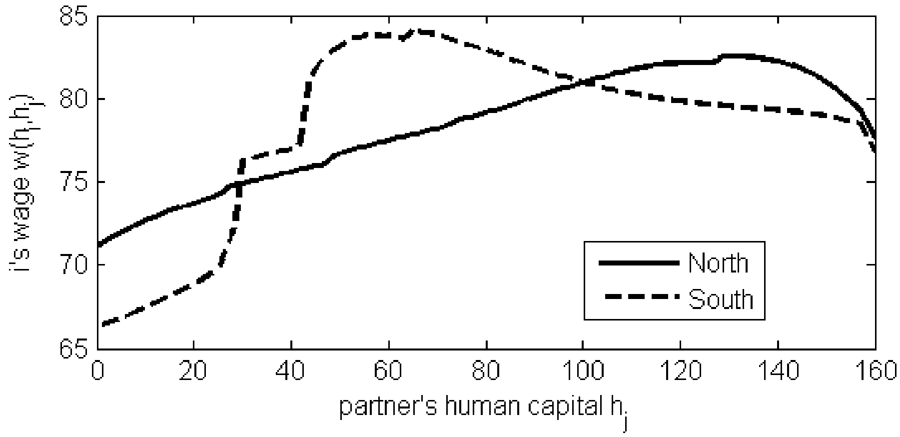

Figure 10: Wage of an individual with $h_i = 157$ (Northern 99th percentile), depending on partner's skill, immediately after the reform

the most skilled ones, return, greatly expanding the frontier of available human capital. In subsequent periods, the following migration pattern arises: some individuals with medium (around 40) human capital emigrate; once they acquire a sufficient amount of knowledge in the North, they return. The average human capital of return migrants is about 120, far above locals' human capital. Overall, about 55% of all emigrants return. Due to high return migration, the North admits more immigrants every period of time which causes higher migration flows.

Still, the flow of return migrants is very small: on average year, the number of return migrants is about 0.1% of total Southern population. Nevertheless, the effect of return migration is tremendous: the returnees bring home knowledge that was previously unavailable; this knowledge is disseminated onto other Southerners. Figure 11 compares the average Southern human capital growth with and without return migration. With return migration, the growth is approximately three times faster. Figure 11 also demonstrates the evolution of per capita GDP under the two scenarios. In the first five years following the reform, both scenarios produce similar results. Immediately after the reform, GDP drops by about 40%: with better institutions, many existing partnerships are terminated, and skills are reallocated in a new, more efficient way. By the end of the second year, GDP is restored to its original level and then continues its growth. After the fifth year, the

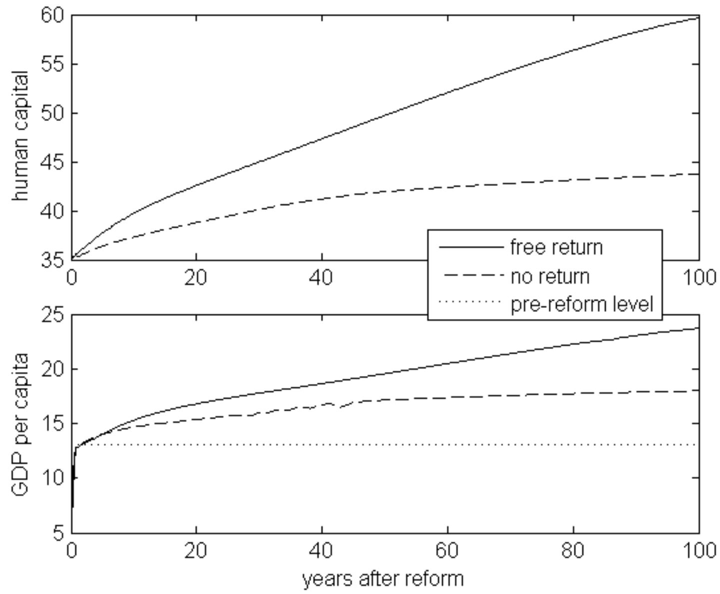

Figure 11: Evolution of economy aggregates under the two scenarios

difference between the two scenarios becomes apparent; the economy grows faster with return migration. Again, GDP growth is about three times faster when return migration is available.

Figure 12 disaggregates the transition path: it shows human capital distributions under the two scenarios, twenty and one hundred years after the reform. After twenty years, the difference between the two scenarios seems small, but return migration brings a long thin tail of highly skilled individuals. Their skill gradually disseminates onto locals through the matching process, and by the year 100 the difference between the two scenarios becomes obvious. Without return migration, most people get close to Southern knowledge frontier (around 70), but expansion of that frontier is a very slow process; without knowledge spillovers from the North, it may take thousands of years to catch up. With return migration, the highest available knowledge in the South (around 150) is just slightly below that of the North; it will probably take another hundred years to get close to Northern human capital distribution.

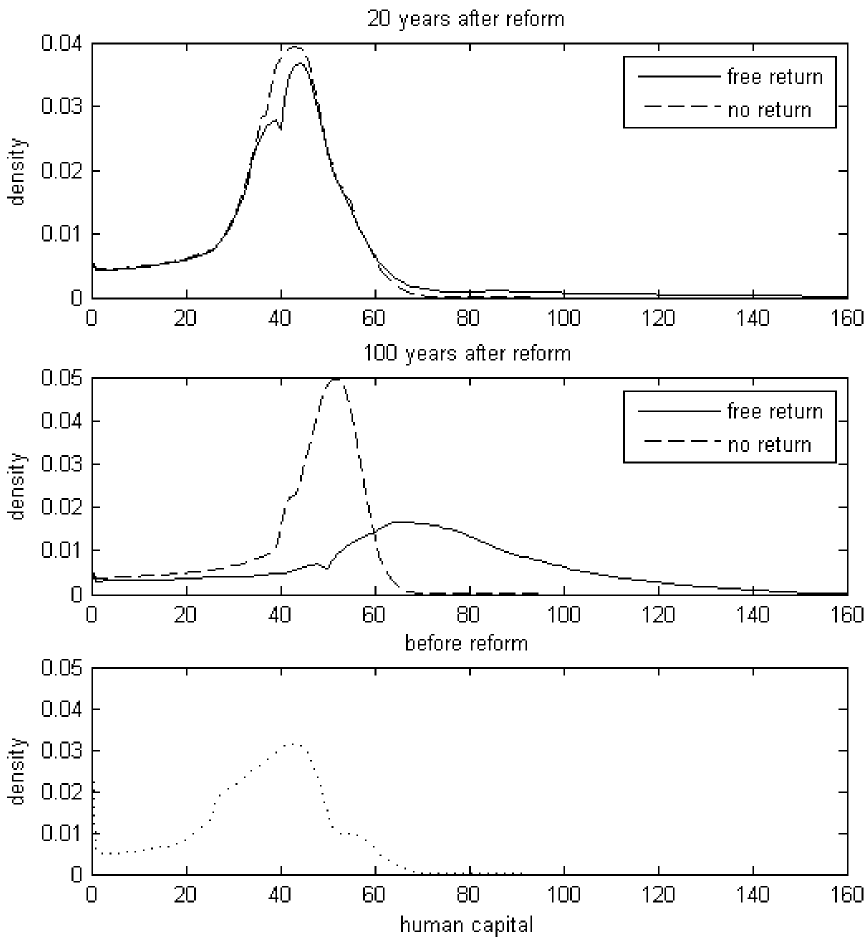

Figure 12: Distribution of human capital before and after the reform

5 Conclusion

This paper constructs a model of "local" knowledge spillovers, in which less skilled individuals learn from more skilled partners; the matching of partners is random. The quality of country institutions, modeled as the degree of matching frictions, greatly affects the accept-reject decisions in the matching process and thus affects the long-run distribution of human capital available on the job market.

When migration becomes available between countries with good (North) and bad (South) institutions, highly-skilled Southerners emigrate for good, leading to a permanent deterioration of the Southern human capital distribution.

When the South improves its institutions to the Northern level, the most highly-skilled emigrants return, because the payoff of being the "king of the hill" in the South now outweighs the payoff of having smarter partners in the North. The return migrants bring home previously unavailable knowledge; local population learns from the return migrants which leads to a rapid human capital growth. Along the equilibrium transition path, the average number of return migrants per year is only about 0.1% of local population; despite their small number, they triple the economy growth after the institutional improvement, compared to no-return-migration scenario.

A Computing steady states

As mentioned above, the concept of the steady state includes the following time-invariant objects:

- Distribution of individuals across types $f = \{f^m(h_i, h_j), f^u(h_i)\}$
- Wages $w(h_i, h_j)$
- Values $V = \{V^m(h_i, h_j), V^u(h_i)\}$

To solve the model, I approximate the continuous distribution across types by 201 discrete types of human capital, ranging from 0 to 200. Theoretically, individual human capital does not have an upper bound because each period individuals increment their knowledge by at least λ_0, a finite positive value, and because individual duration of life is not bounded from above. The calculations, however, show that the probability of reaching beyond some finite threshold of human capital becomes negligibly small, hence a finite grid is a good approximation of human capital distribution.

The total number of individual states is then $201^2 + 201 = 40602$, which renders impossible the precise computation of values and densities at each state. I calculate the unknowns approximately using an iterative algorithm described below. One problem I faced during the computation was discreteness of decisions: individuals accept or reject matches by comparing the two values; they randomize if the values are equal. Since the number of individual types is finite and the mass of individuals of most types is strictly positive, it is highly likely that in a steady state one of the types will be randomizing between the two options. However, because I compute the values approximately using an iterative procedure, the values of accepting or rejecting the matches will never be identical, and no individual will ever randomize. Because of this problem, the system may never fully converge to a steady state: some types of matches will be "blinking", staying together in one iteration but splitting in the next. To override this problem, I force individuals to randomize between the two options: when the values of accepting and rejecting are close, individuals choose the outcome randomly, with probability of accepting sharply increasing as the difference in values increases. This approach is illustrated on figure 13. This randomization was ignored when calculating the values.

Another problem is that future states (human capitals after learning) lie off the grid. I use the standard solution for this problem: when tomorrow

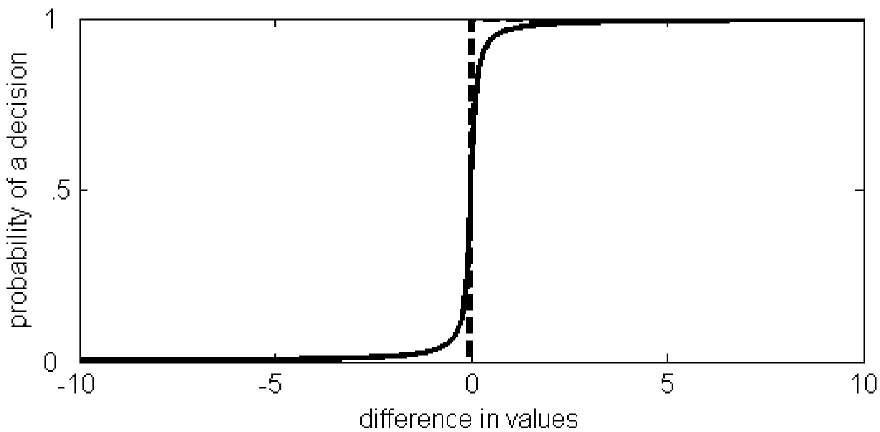

Figure 13: Accept-reject decision randomization. Dashed line represents outcomes in the theoretical model; solid line represents randomizing individuals

human capital falls off the grid, individuals are randomly assigned to one of the nearest grid points. For example, future human capital of 5.2 implies that human capital will be 5 with probability 80%, and 6 with probability 20%.

Below, I describe in detail the steps that I made to find steady states.

Step 0: Initialization

Define an initial distribution of individuals across states, values of each state, and wages at each matched state.

Step 1: Update wages and values

The values are computed using the Bellman equation. First, I compute the wages $w(h_i, h_j)$ as described in section 2.4, using values V^m and V^u taken from the previous iteration. Then, I compute the values V^m and V^u using the newly computed wage w and the distribution of potential partners f^u. I run multiple iterations to update wages and values, given the same distribution f^u, until the change in values becomes small enough.

Step 2: Distribution update

Given wages and values, I simulate individual decisions and compute the resulting distribution across states. The distribution update is performed a fixed number of times[11].

After that, I return to step 1; the whole procedure is repeated until the change in distribution becomes small enough.

B Computing transition dynamics

Because convergence to the steady state takes an infinite number of periods, and because each period in transition is different from the other, the precise transition path cannot be computed; some kind of approximation is required. I use the parametric path approach inspired by Judd (1999). This method exploits the fact that the distribution of individuals across states evolves smoothly over time, and therefore can be approximated by a smooth function of time:

$$f_t^u(h) = \left(\frac{\sum_{k=0}^{K} \phi_k(h) a_k t^k}{\sum_{k=0}^{K} a_k t^k} \right) e^{-\alpha t} + f_\infty^u(h) \left(1 - e^{-\alpha t}\right) \tag{6}$$

where f_∞^u is the steady state distribution of potential partners, α is the speed of convergence to the steady state, $\phi_k(h)$ are ex-ante chosen density functions which resemble f_t^u at different moments of time (notably, $\phi_0(h)$ is the initial distribution), and a_k are the unknown parameters to be estimated. Judd (1999) assumes that α is known ex-ante when solving the model, which is not true in our case: the speed of convergence depends greatly on migration patterns, which in turn depend greatly on individual expectations. Therefore, I estimate the α along with other parameters.

B.1 Computing transition path, given beliefs about future

Values and wages: backward induction

If emigrants are allowed to return, they are no longer identical to Northerners: when bargaining, they have more outside opportunities than Northerners do.

[11] more than once – otherwise the algorithm would be too slow, because the distribution is changing slowly

Therefore, we need to calculate not only the values of living in the South, but also the values of emigrants living in the North.

Using the bargaining rules described in section 2.4, the Bellman equation, and individual beliefs about the evolution of potential Southern partners[12], I compute the path of values $V_0, V_1, ..., V_t, ...$

The backward induction procedure implies that we start from some final date T and from some ex-ante chosen value V_T. This terminal value is likely to be incorrect, but the error should decrease as we move from one iteration to another. My goal was to compute transition dynamics between years 0 and 100; to have accurate values between these dates, I started my backward induction from the terminal year 150.

Distributions: forward induction

Given the path of values, I simulate individual accept-reject decisions in bargaining, and migration decisions. To increase the accuracy of results, I exploit the following property: with good institutions ($\theta = 1$), the value of being unmatched tomorrow is

$$V_{t+1}^u(h_i) = \frac{\int V_{t+1}^m(h_i, h_j) f_{t+1}^u(h_j) dh_j}{\int f_{t+1}^u(h_j) dh_j} \tag{7}$$

As mentioned above, it is impossible to calculate a perfectly accurate V^m because it depends on entire stream of unknown future distributions; however, it is possible to compute rational beliefs about f_{t+1}^u and therefore compute more accurate V^u using (7).

The transition is computed as follows. At the beginning of period t, people have beliefs about V_{t+1}^m and f_{t+1}^u; using (7), they compute the V_{t+1}^u. Then, everybody makes decisions (bargaining, migration), which result in new tomorrow distribution f_{t+1}^u. Using this new distribution and the same V_{t+1}^m, I compute the new V_{t+1}^u, and so on, until beliefs about f_{t+1}^u become perfectly rational. Then, I proceed to the next period.

B.2 Computing beliefs

To use Judd's method described by (6), we need to specify the speed of convergence α, distributions ϕ_k, and weight parameters a_k. I do this in two

[12]The distribution of Northern partners does not change over time and is known precisely

stages: first, I find α and suggest ϕ_k by computing a crude transition path; then, I search for a for more accurate approximation.

Crude approximation

At this stage, I use the simplest version of (6) to model beliefs:

$$f_t^u(h) = f_0^u(h)e^{-\alpha t} + f_\infty^u(h)\left(1 - e^{-\alpha t}\right)$$

I pick up a parameter α and use these beliefs to compute the values, and then the transition path as described in section B.1. Given the transition path, I estimate a new α, and so on, until the change in α becomes small enough.

With this crude approximation, the actual distribution along the path differs quite drastically from individual beliefs about that distribution. I define $\phi_1(h)$ as the "average" distribution over time[13]

$$\phi_1(h) = \frac{1}{T}\sum_{t=1}^{T} f_t^u(h)$$

More accurate approximation

Since we add only one function $\phi_1(h)$ to Judd's formula (6), there is only one unknown parameter a_1.[14] I choose the a_1 to minimize the discrepancy between the actual transition path and individual beliefs about that transition. Given a_1 and new beliefs, I compute the new transition path, then estimate new a_1, and so on, until the change in a_1 becomes small enough. Figure 14 demonstrates actual distributions of potential partners, and beliefs about those distributions, at different points in time.

C If there were no matching frictions

Without matching frictions, when each individual can directly match with the optimal partner, the model simplifies in a number of ways. First, because there is a large number of potential partners of each type, everyone's

[13] We could also pick ϕ_2 and ϕ_3 and so on, but apparently only one extra degree of freedom does a good job in approximation

[14] The a_0 can be normalized to unity

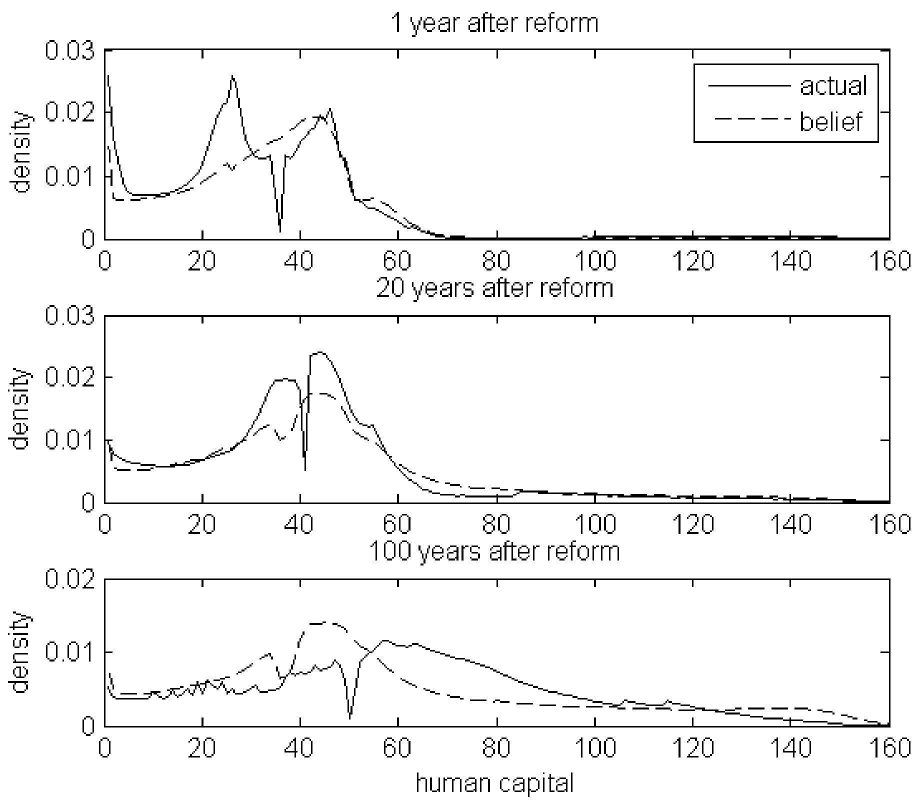

Figure 14: Distribution of potential partners in the South at different points in time

reservation wage exactly equals their actual income. Second, since another partner can be found instantaneously, the state of an individual is described only by his own human capital and does not depend on partner's h:

$$V(h_i) = \max_{h_j} w(h_i, h_j) + \beta V(h_i')$$

where h_i' is the future human capital described by (2). This problem can be split into two subproblems: when i is the master ($h_i \geq h_j$), and when i is the apprentice ($h_i < h_j$). Since master's future state does not depend on apprentice's knowledge, he simply maximizes current income. Therefore, we can define master's reservation wage as

$$\underline{w}^M(h_i) = \max_{h_j \leq h_i} w(h_i, h_j)$$

On the other hand, if i wants to learn from a higher partner j, he realizes that j should earn his reservation wage and i will earn the residual; i's problem is

$$\max_{h_j > h_i} y(h_i, h_j) - \underline{w}^M(h_j) + \beta V(h_i + g(h_j - h_i) + \lambda_0) \tag{8}$$

Since the production function is Leontief, current output does not depend on master's knowledge which simplifies the above optimization problem.[15]

Generally, this problem appears to have no analytical solution; but it can be solved with the following

Assumption Individuals are always indifferent between being masters and apprentices.

Then, the individual value function takes the form

$$V(h_i) = \frac{1}{2}\frac{h_i}{1-\beta} + C \tag{9}$$

where C is a positive constant. With this value, it can be shown that when i is the master, his reservation wage is $\underline{w}^M(h_i) = \frac{1}{2}h_i + C_w$, where C_w is another positive constant.

The apprentice's problem is to solve (8); the first-order condition is

$$-\frac{1}{2} + \frac{1}{2}\frac{\beta}{1-\beta}\frac{\lambda}{1+\lambda(h_j^* - h_i)} = 0$$

[15] That is exactly why the Leontief production function was chosen

As a result, the optimal master-apprentice knowledge gap is

$$h_j^* - h_i = \frac{\beta}{1-\beta} - \frac{1}{\lambda}$$

It is positive as long as $\lambda > \frac{1-\beta}{\beta}$, and it does not depend on apprentice's current state. By calculating apprentice's value, we can confirm that the latter equals (9).

One big problem of frictionless matching is incentive compatibility: the number of individuals with human capital h_1 willing to learn from those with human capital h_2 must be exactly equal to those with h_2 willing to teach h_1. Generally, it cannot be achieved under the assumption made above; finding an equilibrium in such environment might become an arduous task. With random matching, this problem dissolves, at the cost of adding more dimensions to the state space.

References

[1] Acemoglu, D., S. Johnson, and J.A. Robinson (2001) The Colonial Origins of Comparative Development: An Empirical Investigation. American Economic Review, vol. 91, no. 5, pages 1369-1401.

[2] Antras, P., L. Garicano, and E. Rossi-Hansberg (2006) Offshoring in a Knowledge Economy. Quarterly Journal of Economics, vol. 121, no. 1, pages 31-77.

[3] Beine, M. , F Docquier, H Rapoport (2001) Brain drain and economic growth: theory and evidence. Journal of Development Economics, vol. 64, issue 1, page 275.

[4] Borjas, G., and B. Bratsberg (1996) Who leaves? The outmigration of the foreign-born. The Review of Economics and Statistics, vol. 78, issue 1, pages 165–176.

[5] Commander, S., M. Kangasniemi, and L. Alan Winters (2003) The Brain Drain: Curse or Boon? IZA discussion paper no. 809.

[6] Docquier, F., and A. Marfouk (2004). Measuring the international mobility of skilled workers (1990-2000) - Release 1.0. World Bank Policy Research Working Paper, n. 3382.

[7] Dustmann, C. (1996) Return migration: the European experience. Economic Policy (April 1996).

[8] Dustmann, C. (2003) Return migration, wage differentials, and the optimal migration duration. European Economic Review, vol. 47, pages 353–369

[9] Dustmann, C. and O. Kirchkamp (2002) The optimal migration duration and activity choice after re-migration. Journal of Development Economics, vol. 67, pages 351-372.

[10] Gmelch, D. (1980) Return Migration. Annual Review of Anthropology, vol. 9, pages 135-159.

[11] Gould, D.M. (1994) Immigrant links to the home country: empirical implications for U.S. bilateral trade flows. The Review of Economics and Statistics, vol. 76, issue 2, page 302.

[12] Haque, N.U., S.J. Kim (1995) "Human Capital Flight": Impact of Migration on Income and Growth. IMF Staff Papers, vol. 42, No.3, page 577.

[13] Huggett, M. (1993) The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. Journal of Economic Dynamics and Control, vol. 17, pages 953-969.

[14] Jones, B. (2007) The Knowledge Trap: Human Capital and Development Reconsidered (unpublished)

[15] Jovanovic, B. and R. Rob (1989) The Growth and Diffusion of Knowledge. Review of Economic Studies, vol. 56, pages 569-582.

[16] Judd, K., F. Kubler, and K. Schmedders (2003) Computational Methods for Dynamic Equilibria with Heterogeneous Agents. In M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds., Advances in Economics and Econometrics. Cambridge University Press.

[17] Kaufmann, D., A. Kraay, and M. Mastruzzi (2006) Governance matters V: Aggregate and Individual Governance Indicators for 19962005. World Bank Policy Research Working Paper 4012.

[18] Kremer, M. (1991) The O-ring theory of economic development. The Quarterly Journal of Economics, vol. 108, no. 3, pages 551-575.

[19] Ley, D., and A. Kobayashi (2005). Back to Hong Kong: return migration or transnational sojourn? Global Networks, vol. 5 issue 2 (April 2005)

[20] Ljungqvist, L., and T. Sargent (2000) Recursive Macroeconomic Theory. The MIT Press.

[21] Markusen, J. and N. Trofimenko (2007) Teaching Locals New Tricks: Foreign Experts as a Channel of Knowledge Transfers. NBER working paper no. 12872.

[22] Mountford, A. (1997) Can a brain drain be good for growth in the source economy? Journal of Development Economics, vol. 53, no. 2, pages 287-303.

[23] Nunn, N. (2004) Slavery, Institutional Development, and Long-run Growth in Africa, 1400-2000 (unpublished).

[24] Olesen, H. (2002) Migration, Return, and Development: An Institutional Perspective . International Migration, vol. 40, page 125

[25] Rauch, J.E. (1999) Networks versus Markets in International Trade. Journal of International Economics, vol. 48, page 7.

[26] Rios-Rull, J.V. (1999) Computation of Equilibria in Heterogenous Agent Models. In R. Marimon and A. Scott, Eds., Computational Methods for the Study of Dynamic Economies. Oxford University Press.

[27] Santos, M.D.D., and F. Postel-Vinay (2003). Migration as a source of growth: The perspective of a developing country. Journal of Population Economics, vol. 16, pages 161-175.

[28] Saxenian, A. (2006) The New Argonauts: Regional Advantage in a Global Economy. Harvard University Press.

[29] Stark, O. (1991) The Migration of Labor. Blackwell, Oxford.

[30] Stark, O. (1996). On the Microeconomics of Return Migration. In V.N. Balasubramanyam and D. Greenaway, Eds., Trade and Development: Essays in Honor of Jagdish Bhawati, pp.32–41. Basingtoke: Macmillan, 1996.

[31] Yang, D. (2006) Why Do Migrants Return to Poor Countries? Evidence from Philippine Migrants Responses to Exchange Rate Shocks. Review of Economics and Statistics, vol. 88, no. 4, pages 715-735.

[32] Zhao, Y. (2002) Causes and Consequences of Return Migration: Recent Evidence from China. Journal of Comparative Economics, vol. 30, pages 376–394