IHS Economics Series
Working Paper 52
March 1998

Human Capital Formation, Asymmetric Information, and the Dynamics of International Migration

Oded Stark
Nancy H. Chau

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

Impressum

Author(s):
Oded Stark, Nancy H. Chau

Title:
Human Capital Formation, Asymmetric Information, and the Dynamics of International Migration

ISSN: Unspecified

1998 Institut für Höhere Studien - Institute for Advanced Studies (IHS)
Josefstädter Straße 39, A-1080 Wien
E-Mail: office@ihs.ac.at
Web: www.ihs.ac.at

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Reihe Ökonomie / Economics Series No. 52

Human Capital Formation, Asymmetric Information, and the Dynamics of International Migration

Oded Stark, Nancy H. Chau

Human Capital Formation, Asymmetric Information, and the Dynamics of International Migration

Oded Stark, Nancy H. Chau

Reihe Ökonomie / Economics Series No. 52

March 1998

Oded Stark
Department of Economics
University of Oslo
P.O. Box 1095 Blindern
N-0317 Oslo, NORWAY
Phone: ++47/22/85 51 12
Fax: ++47/22/85 79 46
and
University of Vienna
Alser Straße 21/9
A-1080 Vienna, Austria

Nancy H. Chau
Department of Economics
Southern Illinois University at Carbondale
Carbondale, Illinois 62901, USA

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Abstract

We consider the case in which the opening up of an economy to migration results in departure of skilled workers. We point out that while the possibility of migration changes the set of employment opportunities, it also affects the structure of incentives: Higher returns to skills in the foreign country influence decisions about skill acquisition at home. We combine the changing opportunities - changing incentive structure idea with an assumption concerning the information environment: Employers in the foreign country are neither perfectly informed nor equally informed over time about the skill levels of individual migrant workers as employers' experience of employing migrant workers accumulates. Our model gives rise to several interesting results. First, while migration is pursued by the relatively high-skilled, subsequent return migrants are drawn from both tails of the migrant skill distribution. Second, the fraction of the home-country workforce acquiring education in the presence of migration opportunities is higher than the fraction of the home-country workforce acquiring education in the absence of migration opportunities. Third, the intertemporal increase in the probability of discovery of individual skill levels prompts a sequence of migratory moves characterized by a rising average skill level, until the probability of discovery arising from accumulation of migrant employment experience reaches its steady state equilibrium. Finally, under well-specified conditions, per capita output in a country vulnerable to migration of skilled members of its workforce is higher than per capita output in a country that is immune to migration.

Keywords

Human capital, asymmetric information, international migration

JEL-Classifications

J24, F22

Comments

Partial financial support from the Austrian Science Foundation under contract number P10967-SOZ is gratefully acknowledged.

1 Introduction

Whatever workers may take with them when they migrate, they cannot possibly transfer their home country's information structure. Consequently, foreign-country employers are not as informed about home-country workers as are home-country employers. Typically, migration runs across cultures as well as countries. Foreign-country employers who do not share the same culture, background, and language as do home-country employers and, for that matter, as the migrants themselves, lack a common framework for assessing the quality and individual merits of migrant workers. For these reasons, skills of foreign workers cannot be easily discerned, and screening is likely to be imprecise and expensive. Incorporation in mainstream migration research of the natural assumption that migration is inherently associated with a heterogeneous information structure (as opposed to the homogeneous information structure that characterizes nonmigrant employment relationships) has, somewhat surprisingly, been an exception rather than the rule (Kwok and Leland 1982; Katz and Stark 1987, 1989; Stark 1991, 1995). The relative ignorance of foreign employers should not be taken as a constant, however. Exposure breeds familiarity, and increased experience with employing migrants is bound to reduce information asymmetries. Such a change can entail interesting dynamics. For example, the accumulation of information erodes both the pooling of low-skill migrant workers with high-skill migrant workers and the associated wage determination rule – viz., paying each migrant the same wage, based on the average productivity of the entire cohort of migrants. Absent pooling, however, low-skill migrant workers may find it advantageous to return-migrate (Stark 1995).

There is little doubt that, in general, migration gives rise to human capital depletion in the home country. The standard argument holds that, absent migration, the home country would have had available to it a more skilled workforce and, concomitantly, would have enjoyed higher per capita output. Indeed, the "drain-of-brains" view has influenced migration research for at least three decades now (Grubel and Scott 1966), with the associated literature concentrating largely on how to mitigate this adverse consequence (Bhagwati and Wilson 1989). However, that migration induces skill *formation* has essentially escaped analysis. Obviously, workers are not endowed with marketable skills at birth. Skills are acquired, and their level is chosen by optimizing workers who, given their innate learning ability (efficiency in skill formation), weigh the prospective market rewards to enhanced skills, both at home *and* abroad, in addition to the cost of acquiring those skills.

The possibility of migration thus changes the opportunities set, the incentive structure, and the information environment. We study these simultaneous changes and trace their implications. Specifically, we depart from earlier writings by dropping the strong simplifying assumptions that the distribution of migrants' abilities and the monitoring capability of migrants' employers are exogenously given. We endogenize the human capital formation decisions of migrant workers and allow the monitoring capability of the employers to improve over time as their experience with employing migrants accumulates. This allows us to explore the intertemporal interactions among the decision to migrate, the choice to undertake education, and the monitoring capabilities of migrants' employers.

Our framework explains a number of pertinent characteristics of skilled migration (the brain drain). For example, as the experience of employing migrants accumulates, the resulting intertemporal adjustment of the probability of deciphering true skill levels leads to a sequence of migratory moves that progressively selects higher skilled workers. We argue that by raising the likelihood of discovering the true quality of workers, accumulation of experience with migrant employment enhances the incentive of brighter brains to migrate permanently, while it reduces the incentive of low-ability workers to pursue migration. As the probability of ability discovery rises, the ability composition of subsequent migrant cohorts shifts rightward. Whenever the average quality of a migrant cohort exceeds that of a previous cohort, wage offers are bid upward, prompting a subsequent wave of migration involving yet even more able workers. However, this is just a first-round effect. The accumulation of migrant employment experience also implies that both high- and low-ability workers are more likely to be discovered. Accordingly, the probability of permanent migration by high-ability workers, as well as the extent of return migration by low-ability workers, rise simultaneously. The result is a continuing improvement in the average ability of migrant workers remaining in the foreign country. Until the steady-state equilibrium probability of discovery is reached, a virtual cycle of migration of the more able ensues, as wage offers adjust over time in favor of migration of higher ability workers. Meanwhile, the wages of the migrants who stay increase, though not because of an increase in their human capital.

Our model extends earlier work by Katz and Stark (1987). We introduce endogenous human capital formation and examine the dynamics of human capital formation and the corresponding intertemporal pattern of migration and return-migration. We derive several dynamic predictions that are consistent with a considerable body of empirical literature, as reviewed and synthesized by LaLonde and Topel (1997) and Razin and Sadka (1997). Migration is a process, not an event. It is phased and it is sequential: Not all workers who migrate move at the same time. Each cohort of migrants includes workers who stay as migrants and workers who, with a well-defined probability, return-migrate. Ravenstein's century-old "law of migration" (1885, p.199) which predicts that "each main current of migration produces a compensating countercurrent" – often quoted but not generated analytically – turns out to be an implication of our model. Within cohorts, migration is positively selective (Stark, 1995).[1] Cohort by cohort, the average quality of migrants rises.[2] The "cost of migration" is a decreasing function of the stock of previous migrants for some workers but is an *increasing* function of that stock for others (contrary to Carrington, Detragiache, and Vishwanath 1996): Migration of low-skill workers pulls down the average of the marginal products of the contemporaneous group of migrant workers, thereby lowering the wage of high-skill workers. Conversely, the presence of high-skill migrant workers in a pool of low-skill and high-skill workers enables low-skill workers to enjoy a wage higher than their marginal product. As migration proceeds and the cumulative stock of migrants rises, the probability of discovery rises. This favors high-skill would-be migrants but dissuades low-skill would-be migrants. Thus, an increase in the stock of migrant workers confers a positive externality on subsequent migration of high-skill workers but a negative externality on the migration of low-skill workers.

We pay particular attention to the change in the welfare of the home-country population in the wake of international migration. In contrast to the received welfare-theoretic analysis of the brain drain,[3] we show that when potential migrant workers incorporate the feasibility of migration in their education decisions, not only does the level of education acquisition in the home country rise, but national welfare may rise as well if the contribution to national income by educated workers increases. We show that a gain in national welfare generated by migration of educated workers is possible, given a positive probability of return-migration by educated workers once their true productivities are deciphered.

The rest of this paper is organized as follows: In Section 2, we model a home economy not open to migration and determine the extent of education acquisition and the per capita output as benchmarks for subsequent comparisons. In Section 3, we present a two-country framework. The information asymmetry between foreign employers and home-country workers is introduced and the effect of migrant employment experience on the probability of deciphering the true ability of individual migrant workers is incorporated. In addition, we study the education and migration decisions of home-country workers in the presence of asymmetric information. We also compare the resulting level of education with that obtained in the absence of the possibility of migration. Section 4 analyzes the relationship between the dynamic process of skilled migration and the probability of discovery, and the associated steady-state equilibrium probability of discovery. We trace the circumstances under which migration progressively selects

[1] Returnees tend to be less educated than the migrants who stay (DaVanzo 1983; Reilly 1994).

[2] Borjas (1987) provides evidence that the quality of migrant workers from Western Europe to the United States has been increasing over the period 1955 - 1979. However, his measures of quality are different from the one used in this paper.

[3] The primary conclusion of Grubel and Scott (1966) and Berry and Soligo (1969) is that while little migration has no impact on the welfare of those who stay behind, finite levels of migration unambiguously reduce welfare.

higher ability workers. In Section 5, we conduct a welfare analysis and define conditions under which national welfare improves when free migration of skilled workers is permitted. Section 6 summarizes the analysis.

2 An Economy without Migration

2.1 Production

At each time period t, the home economy h produces a single composite good in two sectors: An unskilled sector, u, and a skilled sector, s. Output in the unskilled sector at time t is generated through a simple constant returns to scale production function, $X_t^u = a_u^h L_t$, where L_t denotes the number of workers employed in sector u. Similarly, output in the skilled sector is given by a constant returns to scale production function, $X_t^s = a_s^h E_t$, where E_t is the input of skilled labor measured in efficiency units. Thus, a_u^h is the marginal and average product of a worker in sector u, and a_s^h is the marginal and average product of an efficiency unit of labor in the skilled sector. Without loss of generality, the price of a unit of output is assumed to be unity. There is perfect competition in both output and factor markets. Therefore, the wage paid by profit maximizing employers to a worker in the unskilled sector is $w_u^h = a_u^h$, and the wage paid for an efficiency unit of work in the skilled sector is $w_s^h = a_s^h$.

2.2 Individuals and the Population

In each period N individuals are born. Individuals live for two periods. Thus, population size in any time period is $2N$. Individuals are characterized by endowments and preferences. Each individual is endowed with one unit of physical labor (a pair of hands) and with innate ability (talent) $\theta \in [0, \infty]$. The distribution of θ over the population is summarized by a cumulative distribution function $F(\theta)$, where $F(\theta)$ is continuously differentiable and is associated with a strictly positive density function $f(\theta)$. Assume, in addition, that the expectation of θ ($\int_0^\infty \theta f(\theta) d\theta$) is finite. Denote by y_t the income of the individual in period t. The individual's preferences are summarized by a utility function $u(y_t, y_{t+1})$. To simplify, we take $u(y_t, y_{t+1}) = y_t + \beta y_{t+1}$ where $0 < \beta < 1$ is the time discount rate.

An individual born at any time period t faces the following choice: Remain uneducated and work in the u sector in each of the two periods of his life, or spend the first time period acquiring education and work in the s sector in the second period of his life. Acquiring education involves a direct cost c, which is incurred at the beginning of period t. Having no funds, the individual borrows c in a perfectly competitive credit market where the interest rate is assumed to be zero. The educated individual, whose innate ability is θ, supplies θ efficiency units of labor to the skilled sector. The supply of efficiency units of labor by an uneducated individual in the s sector is zero, irrespective of his level of innate ability. The labor input supplied by a worker in the unskilled sector is independent of his innate abilities and is equal to his physical labor endowment (one unit).

It follows that the discounted lifetime utility of an educated worker is equal to his discounted second period income, net of education costs:

$$Y_t^s(\theta) \equiv \beta(w_s^h\theta - c).$$

The discounted lifetime utility of an uneducated worker is:

$$Y_t^u \equiv (1+\beta)w_u^h.$$

Thus, an individual whose innate ability is θ will decide to acquire education if $Y_t^s(\theta) \geq Y_t^u$, and will choose to remain uneducated otherwise. We thus have:

$$Y_t^s(\theta) \geq Y_t^u \Leftrightarrow \beta(w_s^h\theta - c) \geq (1+\beta)w_u^h,$$

or,

$$\theta \geq \frac{1}{w_s^h}[\frac{(1+\beta)w_u^h}{\beta} + c] \equiv \theta^*.$$

That is, individuals whose $\theta \geq \theta^*$ will become skilled workers, while individuals whose $\theta < \theta^*$ will remain unskilled.

Therefore, the $2N$ individuals from the "young" and the "old" generations are distributed across three activities: Work in the u sector, work in the s sector, and acquisition of education. Since the fraction of uneducated workers per generation is $F(\theta^*)$, the number of uneducated workers in the population is $2NF(\theta^*)$. The fraction of the old generation employed in the s sector is $(1 - F(\theta^*))$. The number of individuals employed in the s sector is thus $N(1 - F(\theta^*))$. Finally, since a fraction $(1 - F(\theta^*))$ of the young generation pursues education, the number of individuals being educated during any time period is $N(1 - F(\theta^*))$. Of course, $2NF(\theta^*) + N(1 - F(\theta^*)) + N(1 - F(\theta^*)) = 2N$. From our previous analysis, it can be confirmed that θ^* is decreasing in w_s^h: The higher the rewards to education, given θ, the larger the fraction of individuals who invest in education, $(1 - F(\theta^*))$, and the larger the number of individuals $N(1 - F(\theta^*))$ who do so.

2.3 Production and Equilibrium

An equilibrium in the economy, at any time, is fully characterized by the parameter θ^*. Once θ^* is known, the allocation of labor across the two employment options and the associated outputs of the two sectors are given. The output of the u sector is $X_t^u = w_u^h 2NF(\theta^*)$. In addition, total labor input (measured in efficiency units) in the skilled sector is $E_t = N\int_{\theta^*}^{\infty}\theta f(\theta)d\theta$.[4] The resulting s sector output is therefore $X_t^s = N\int_{\theta^*}^{\infty} w_s^h\theta f(\theta)d\theta$.

We can now calculate the value of national output, net of education expenditures, and investigate the dependence of national output on θ^*. Denote by $V(\theta^*)$ the time invariant value of national output net of the cost of education. We have

$$V(\theta^*) = N\{2w_u^h F(\theta^*) + \int_{\theta^*}^{\infty} w_s^h\theta f(\theta)d\theta - c(1 - F(\theta^*))\}.$$

Output per capita is thus

$$v(\theta^*) = \frac{1}{2}\{2w_u^h F(\theta^*) + \int_{\theta^*}^{\infty} w_s^h\theta f(\theta)d\theta - c[1 - F(\theta^*)]\}. \tag{1}$$

It follows that[5]

$$\begin{aligned}
\frac{\partial v(\theta^*)}{\partial\theta^*} &= \frac{1}{2}[-w_s^h\theta^* f(\theta^*) + 2w_u^h f(\theta^*) + cf(\theta^*)] \\
&= -\frac{1}{2}f(\theta^*)(w_s^h\theta^* - 2w_u^h - c) \\
&= -\frac{1}{2}f(\theta^*)w_u^h\frac{1-\beta}{\beta} \\
&< 0.
\end{aligned}$$

The value of per capita output is decreasing in θ^*. Recall that an increase in θ^* is equivalent to a reduction in the fraction of the educated workforce. Starting from an equilibrium in which there is no governmental interference in individuals' decisions to acquire education, it follows that per capita output increases as the share of educated workers increases. Note that if in the far right-hand side of $\frac{\partial v(\theta^*)}{\partial\theta^*}$ we have $\beta = 1$, then $\frac{\partial v(\theta^*)}{\partial\theta^*} = 0$. In other words, if individuals do not discount future income, the invisible hand is nicely at work: The level of θ^* chosen by individuals who maximize expected lifetime utility is exactly the same level of θ^* that a social planner will choose to maximize per capita output.

[4]The average number of efficiency units of labor supplied by a skilled worker is $\int_{\theta^*}^{\infty}\theta f(\theta)d\theta / \int_{\theta^*}^{\infty} f(\theta)d\theta$. Since there are $N(1-F(\theta^*))$ skilled workers, their total supply of skilled work is $[\int_{\theta^*}^{\infty}\theta f(\theta)d\theta / \int_{\theta^*}^{\infty} f(\theta)d\theta]N[1 - F(\theta^*)] = N\int_{\theta^*}^{\infty}\theta f(\theta)d\theta$.

[5]To derive the last equality, note that from the definition of θ^*, $w_s^h\theta^* - c = (1+\beta)w_u^h/\beta$. Hence, $w_s^h\theta^* - 2w_u^h - c = (1-\beta)w_u^h/\beta$.

3 A Two-Country World with Migration

3.1 The Foreign Economy

The foreign economy, f, also consists of a u sector and an s sector. Denote by $\tilde{L}_t$ and L_t^m the number of foreign workers and migrant workers employed in the u sector, respectively. The output of the u sector is $\tilde{X}_t^u = a_u^f(\tilde{L}_t + L_t^m)$. The output of the s sector, $\tilde{X}_t^s$, is governed by the production function $\tilde{X}_t^s = a_s^f(\tilde{E}_t + E_t^m)$, where $\tilde{E}_t$ denotes the foreign workforce (measured in efficiency units) employed in the s sector and E_t^m is the input of the migrant workforce, also measured in efficiency units. We shall assume that the foreign country uses superior technologies relative to economy h in both its u and s sectors so that $a_i^f > a_i^h$, $i = u, s$. Perfect competition in both output and factor markets guarantees that the wage paid by profit maximizing employers to a worker in the unskilled sector is $w_u^f = a_u^f > a_u^h = w_u^h$ and the wage paid to an efficiency unit of work in the skilled sector is $w_s^f = a_s^f > a_s^h = w_s^h$.

Foreign employers are assumed to be perfectly aware of the true abilities of indigenous workers. However, the true ability of individual migrant workers is unknown. Each migrant worker can nevertheless be distinguished as belonging to one of the following identifiable groups of workers: Educated or uneducated. Following our specification in Section 2, wage payments to uneducated migrant workers by profit maximizing employers depend only on the sector of employment, not on individual abilities. In particular, an uneducated migrant worker receives zero wages in the s sector since the efficiency labor input of such a worker in this sector is zero. Similarly, an uneducated migrant worker in the u sector receives w_u^f as a wage payment since the physical labor input of such a worker in the u sector is one. The same wage formation procedure no longer applies to educated migrant workers, however, when the educated migrant workforce consists of individuals with heterogeneous abilities.

At any time t, let the wage offer to an educated migrant worker whose true ability is unknown to foreign employers be $w_s^f \theta_t^a$, where θ_t^a denotes the average supply of efficiency labor inputs by the migrant population having unknown individual abilities. In addition, let the total number of migrants at any time τ be $\mathcal{M}_\tau$, and the cumulative number of migrants until time $t-1$ be $M_{t-1} = \sum_{\tau=0}^{t-1} \mathcal{M}_\tau$. We assume that with probability $m_t = m(M_{t-1})$ the actual productivity of a worker who supplies $\theta \neq \theta_t^a$ amount of skilled labor will be discovered. The probability of discovery, m_t, is taken to be strictly positive, increasing in migrant hiring experience, $m'(M_{t-1}) > 0$,[6] and bounded from above with $\lim_{M_{t-1}\to\infty} m(M_{t-1}) = \hat{m} < 1$. Once the true ability of a worker is discovered by one foreign employer, the same information becomes instantly available to all foreign employers (this follows from our assumption of perfect

[6] A prime ($'$) denotes the first derivative with respect to M_{t-1}.

competition in factor markets); hence, the wage payment for such a worker in the s sector of the foreign country is determined by his true ability, θ.

3.2 The Individuals Revisited

Migration entails a per period cost k, which may be perceived as the cost of separation from home. We take this cost to be independent of the level of education acquired and of the stock of migrants. Accordingly, under symmetric information, the per period income net of the separation cost for an educated worker who migrates to the foreign country is just $w_s^f\theta - k$. Figure 1 illustrates the income schedules of an educated worker in the home country and in the foreign country. The value of θ corresponding to the point of intersection, $\bar{\theta} = k/(w_s^f - w_s^h)$, denotes a critical level of innate ability such that any educated home-country worker with an innate ability $\theta \geq \bar{\theta}$ enjoys a higher income in the foreign country, net of the migration cost, than at home. In the absence of asymmetric information, the most talented migrate while skilled workers endowed with ability less than $\bar{\theta}$ remain in the home country because the per-period foreign wage net of the cost of migration $(w_s^f\theta - k)$ is less than the corresponding home-country wage $(w_s^h\theta)$.

Once the prevalence of asymmetric information and the possibility of migration are incorporated into the decision-making calculus of the home-country workers, the problem of a worker born at any time t spans two consecutive periods. In the first period, an individual may acquire education and incur its cost, c. Otherwise, the individual finds employment in the unskilled sector of the home country or the foreign country. In the second period, the uneducated individual reviews his migration decision and chooses to work either in the home country or in the foreign country. For an educated worker, there are four possible, more elaborate second-period employment options:

1. An educated worker of ability θ chooses to migrate. With probability m_t, the true ability of the worker is discovered and the worker return-migrates. With the complementary probability $(1 - m_t)$, the true ability of the worker is not discovered and he remains in the foreign country. The expected income of such a worker, net of the cost of education, y_t^{rd}, is thus
$$y_t^{rd}(\theta) = m_t w_s^h\theta + (1 - m_t)(w_s^f\theta_t^a - k) - c.$$

2. An educated worker of ability θ chooses to migrate. With probability m_t, the true ability of the worker is discovered and the worker remains in the foreign country, receiving $w_s^f\theta$. With the complementary probability $(1 - m_t)$, the true ability of the worker is not discovered and the worker remains in the foreign country, in which case he receives $w_s^f\theta_t^a$.

In this case, the expected income net of the education cost, $y_t^f(\theta)$, is thus

$$y_t^f(\theta) = m_t(w_s^f\theta - k) + (1 - m_t)(w_s^f\theta_t^a - k) - c.$$

3. An educated worker of ability θ chooses to migrate. With probability m_t, the true ability of the worker is discovered and the worker remains in the foreign country, receiving $w_s^f\theta$. With the complementary probability $(1 - m_t)$, the true ability of the worker is not discovered and the worker return-migrates. The expected income of such a worker, net of the cost of education, $y_t^{ru}(\theta)$, is thus

$$y_t^{ru}(\theta) = m_t(w_s^f\theta - k) + (1 - m_t)w_s^h\theta - c.$$

4. An educated worker of ability θ chooses not to migrate and receives a net income, $y_t^h(\theta)$, of

$$y_t^h(\theta) = w_s^h\theta - c$$

with probability one.[7]

Figure 2 depicts these four options and the choice among them. Given θ_t^a and M_{t-1}, the expected income schedules in the four regimes above, $y_t^i(\theta) + c$ $(i = rd, f, ru, h)$, are illustrated by lines R^dR^d, FF, R^uR^u, and HH, respectively. R^dR^d is the income schedule of migrant workers who return upon discovery. R^uR^u respresents the income schedule of migrant workers who return if their true abilities remain undiscovered. HH and FF denote the income schedules of permanent home-country workers and permanent migrants, respectively. Note in particular that R^dR^d and FF coincides with the horizontal income schedule $w_s^f\theta_t^a - k$, while R^uR^u coincides with the home wage schedule HH whenever $m_t = 0$. In addition, R^dR^d coincides with the home wage schedule while R^uR^u and FF coincide with the foreign wage schedule whenever $m_t = 1$ – the case of perfect information elaborated above.

Observe from figure 2 that when $0 < m_t < 1$, the maximum second-period expected income of an educated worker (indicated by the bold segmented line) is demarcated by two critical values of innate abilities: $\bar{\theta}$ and θ_t^f, where the former (latter) denotes the innate ability of a migrant who is indifferent between regimes 1 and 2 (regimes 2 and 3). A comparison of the migration patterns shown in figures 1 and 2 reveals that under asymmetric information, the

[7] In general, there can be two additional migration regimes for educated home-country workers: 5) An educated worker migrates. With probability m_t, the true ability of the worker is discovered and the worker remains in the foreign country to engage in u sector employment. With probability $1 - m_t$, the worker receives $w_s^f\theta_t^a$ in the foreign country. 6) An educated worker migrates. With probability m_t, the true ability of the worker is discovered and the worker return-migrates to engage in u sector employment in the home country. With probability $1 - m_t$, he receives $w_s^f\theta_t^a$ in the foreign country. Later, we show that neither of these options will be pursued by educated migrant workers as long as w_u^h is sufficiently small and k is sufficiently large.

most talented workers (with $\theta > \theta_t^f$) return-migrate with strictly positive probability $(1 - m_t)$. The inability of foreign employers to decipher the true ability of migrant workers thus acts as a tax on the returns to migration for the most talented workers. In particular, the innate ability of a migrant who is indifferent between regimes 2 and 3, θ_t^f, can be found by noting that

$$\begin{aligned}(1 - m_t)(w_s^f \theta_t^a - k) + m_t(w_s^f \theta - k) &= (1 - m_t)w_s^h \theta + m_t(w_s^f \theta - k) \\ \Leftrightarrow \theta &= \frac{w_s^f \theta_t^a - k}{w_s^h} \equiv \theta_t^f.\end{aligned}$$

Note further that since R^dR^d and FF intersect only once, all educated migrants with ability $\theta \leq \theta_t^f$ can be classified into two groups once their true capabilities are detected – return migrants and permanent migrants. This follows from the definition of $\bar{\theta}$, the critical innate ability level at which the home and the foreign wage schedules intersect. Once his true ability is discovered, an educated worker with low-ability $\theta < \bar{\theta}$ chooses to work in the home country, where the per-period return is the highest. To see this, note that a migrant with innate ability θ is indifferent between regimes 1 and 2 if and only if

$$\begin{aligned}(1 - m_t)(w_s^f \theta_t^a - k) + m_t w_s^h \theta &= (1 - m_t)(w_s^f \theta_t^a - k) + m_t(w_s^f \theta - k) \\ \Leftrightarrow w_s^h \theta &= w_s^f \theta - k \\ \Leftrightarrow \theta &= \frac{k}{w_s^f - w_s^h} \equiv \bar{\theta}.\end{aligned}$$

Finally, HH lies below the bold segmented line for all values of θ. As long as $\theta_t^a > \bar{\theta}$ and m_t is strictly between zero and one, the probability that an educated migrant earns a higher wage in the foreign country is strictly positive. In particular, from the definition of $\bar{\theta}$, if $\theta \leq \bar{\theta}$, $w_s^f \theta_t^a - k > w_s^h \theta$. In addition, $w_s^f \theta - k > w_s^h \theta$ if $\theta > \bar{\theta}$. It follows that if return-migration is always an option open to migrant workers, an educated worker will never choose to work only in the home country. We summarize our discussion above by the following proposition:

Proposition 1

If $\theta_t^a > \bar{\theta}$

1. *Educated workers with innate ability* $\theta \leq \bar{\theta}$ *migrate. In addition, return migration yields the maximum second-period income once the true ability of such educated workers is discovered.*

2. *Educated workers with innate ability* $\theta \in (\bar{\theta}, \theta_t^f]$ *migrate. In addition, employment in the foreign country yields the maximum second-period income once the true ability of such educated workers is discovered.*

3. *Educated workers with innate ability $\theta > \theta_t^f$ migrate. In addition, return migration yields the maximum second-period income if the true ability of such educated workers is not discovered.*

Proof: All proofs are relegated to the appendix.

We now proceed to the first-stage education choice by comparing the expected utility of an educated worker and an uneducated worker over the two periods. To focus on the analysis of migration of skilled workers, we assume that k is sufficiently large with $w_u^f - k < w_u^h$: The per-period foreign income net of the migration cost for an uneducated worker is lower than his unskilled wage in the home country. It follows that

$$\begin{aligned} w_u^h > w_u^f - k \quad &\Rightarrow \quad w_u^h(1+\beta) > (w_u^f - k) + \beta w_u^h, \\ &\Rightarrow \quad w_u^h(1+\beta) > w_u^h + \beta(w_u^f - k). \end{aligned}$$

The inequality on the right-hand side in the first line states that lifetime utility from working at home is higher than utility from migrating in the first period of life and utility from working at home in the second period. The inequality on the right-hand side in the second line states that lifetime utility from working at home is higher than utility from working at home in the first period of life and utility from migrating in the second period. Since $(1+\beta)w_u^h > (1+\beta)(w_u^f - k)$ follows from $w_u^h > w_u^f - k$, it follows that migration of the unskilled always yields a lower lifetime utility, irrespective of the timing and duration of migration.

With this in mind, denote the expected lifetime utility $u_{t-1}(\theta)$ of a worker with innate ability θ born at time $t-1$ as

$$u_{t-1}(\theta) = E_{t-1}\{\max[\beta y_t(\theta), (1+\beta)w_u^h]\},$$

where $y_t(\theta) = \max[y_t^{rd}(\theta), y_t^f(\theta), y_t^{ru}(\theta), y_t^h(\theta)]$ and $E_{t-1}(\cdot)$ denotes the expectation operator, with the expectation taken over all possible values of m_t at time $t-1$. The expected lifetime utility $u_{t-1}(\theta)$ can be determined by comparing the expectation of the discounted lifetime income of an educated worker, $E_{t-1}(\beta y_t(\theta))$, and the discounted lifetime income of an uneducated worker, $(1+\beta)w_u^h$. To do so, additional assumptions regarding the determination of the expected future foreign wage offers, θ_t^a, and the probability of discovery, m_t, are required. In what follows, we endow individuals with the faculty of rational expectations, such that $E_{t-1}(x_t) = x_t$.

Consider, then, the lifetime utility of an individual born at time $t-1$ with $\theta < \bar{\theta}$. From proposition 1, such a worker strictly prefers regime 1 if educated, and hence the expectation of his discounted lifetime income is just $\beta y_t^{rd}(\theta)$. Education therefore yields a higher lifetime utility than no education if and only if $\beta y_t^{rd}(\theta) > (1+\beta)w_u^h$, or if and only if

$$\begin{aligned}\beta[m_t w_s^h \theta + (1-m_t)(w_s^f \theta_t^a - k) - c] &> (1+\beta)w_u^h \\ \Leftrightarrow \theta &> [\frac{1+\beta}{\beta}w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c]\frac{1}{m_t w_s^h} \\ &\equiv \theta_t^{er}.\end{aligned}$$

Note that θ_t^{er} is strictly increasing in w_u^h , c, and k. We thus have the following result: All else constant, *the higher the unskilled wage and the higher the cost of education, the smaller the fraction of the home-country population acquiring education* $(1 - F(\theta_t^{er}))$. Interestingly, *an increase in the cost of migration, k, also deters education by home-country workers.* Education not only varies wage earnings at home and abroad, it also renders migration a feasible option. An increase in the cost of migration weakens the migration incentive for acquiring education. Finally, θ_t^{er} is also increasing in m_t whenever $\partial\theta_t^{er}/\partial m_t = (1/w_s^h m_t)(w_s^f \theta_t^a - k - w_s^h \theta_t^{er}) > 0$, which, in turn, holds since $\theta_t^{er} < \bar{\theta}$. An increase in the probability of discovery, m_t, lowers the education incentives of low-ability workers – the probability that these workers will be pooled with high-ability workers is lower; therefore, the expected returns to their acquisition of skills are lower. It follows that the fraction of the home-country workers who remain uneducated rises as m_t rises, all else remaining constant.

Similarly, the lifetime utility of an educated individual with $\theta \in [\bar{\theta}, \theta_t^f)$ is higher than the utility of an unskilled worker if and only if $\beta y_t^f(\theta) > (1+\beta)w_u^h$, or if and only if

$$\begin{aligned}\beta[m_t(w_s^f \theta - k) + (1-m_t)(w_s^f \theta_t^a - k) - c] &> (1+\beta)w_u^h \\ \Leftrightarrow \theta &> [\frac{1+\beta}{\beta}w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c + m_t k]\frac{1}{m_t w_s^f} \\ &\equiv \theta_t^{ef},\end{aligned}$$

where θ_t^{ef} is increasing in w_u^h, c, and k as well as in m_t, provided $\partial\theta_t^{ef}/\partial m_t = (1/w_s^f m_t)(w_s^f \theta_t^a - w_s^f \theta_t^{ef}) > 0$.

With θ_t^{er} and θ_t^{ef} now established, there are two critical levels of innate ability that further divide the home-country population into two groups: Uneducated and educated[8]. Simple

[8]Given θ_t^{er} and θ_t^{ef}, we are now in a position to demonstrate the conditions under which no educated return-migrant will be employed in the u sector of the home country. To this end, note that in equilibrium, $\beta y_t^{rd}(\theta_t^{er}) = (1+\beta)w_u^h$. An educated worker is strictly better off working in the skilled sector if and only if $w_s^h \theta_t^{er} > w_u^h$ as the skill level of all educated workers is no less than θ_t^{er}. From the definition of θ_t^{er}, we have

$$w_s^h \theta_t^{er} = \frac{1}{m_t}[\frac{(1+\beta)w_u^h}{\beta} - (1-m_t)(w_s^f \theta_t^a - k) + c]$$

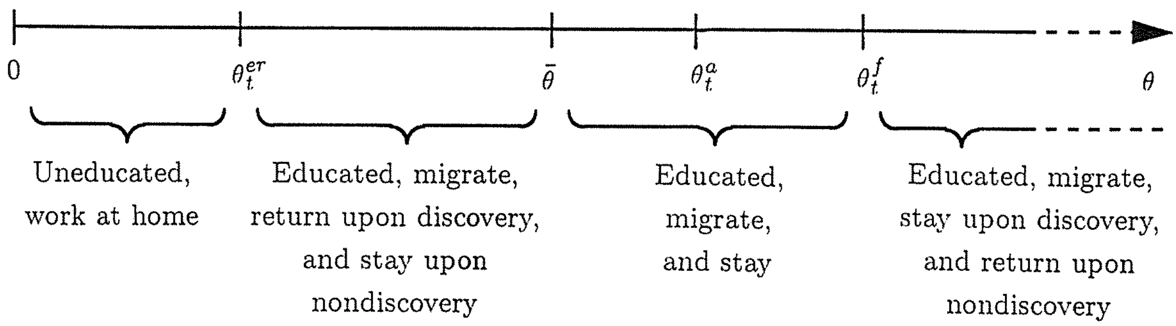
manipulation of the definitions of θ_t^{er} and θ_t^{ef} yields the following result:

Proposition 2

1. *If* $\theta_t^{er} < \bar{\theta}$: *Workers with innate ability* $\theta > \theta_t^{er}$ *are better off acquiring education. The lifetime utility of workers with* $\theta \leq \theta_t^{er}$ *is maximized by remaining uneducated.*

2. *If* $\theta_t^{er} \geq \bar{\theta}$: *Workers with innate ability* $\theta > \theta_t^{ef}$ *are better off acquiring education. The lifetime utility of workers with* $\theta \leq \theta_t^{ef}$ *is maximized by remaining uneducated.*

If $\theta_t^{er} < \bar{\theta}$, θ_t^{er} defines a critical ability level that divides the home-country population into educated and uneducated workers. Now, the home-country population consists of four groups of individuals: Uneducated home-country workers (with $\theta < \theta_t^{er}$), educated workers who migrate and return upon discovery (with $\theta_t^{er} \leq \theta < \bar{\theta}$), educated permanent migrants (with $\bar{\theta} \leq \theta < \theta_t^f$), and educated workers who migrate and return if their true ability is not discovered (with $\theta \geq \theta_t^f$). This partitioning is depicted below.



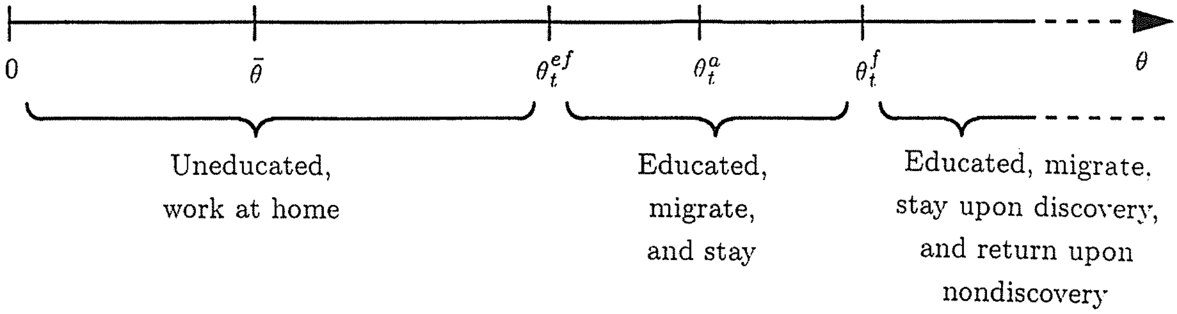
Note again that under asymmetric information, all individuals with $\theta \geq \bar{\theta}$ do not permanently migrate. As noted earlier, asymmetric information penalizes high-ability migrant workers since with probability $1 - m_t$, such migrants do not receive the foreign wage that accords with their abilities. More importantly, upon return-migration, the home-country population consists of individuals with the lowest and highest ability levels.

If $\theta_t^{er} > \bar{\theta}$, then workers with $\theta < \bar{\theta}$, as well as workers with $\bar{\theta} < \theta < \theta_t^{ef}$, remain uneducated. Therefore, the home-country population consists of only three groups: Uneducated

$$\begin{aligned} &= w_u^h + \frac{1}{m_t}\{[\frac{(1+\beta(1-m_t))w_u^h}{\beta}] - (1-m_t)(w_s^f\theta_t^a - k) + c\} \\ &> w_u^h \end{aligned}$$

if w_u^h is sufficiently small. In addition. since $w_u^f - k < w_u^h$, by transitivity, $w_s^h\hat{\theta}_t^{er} > w_u^h > w_u^f - k$. Hence, migration regimes 5 and 6, as discussed in the previous footnote, will not be pursued by any educated migrant worker.

home-country workers (with $\theta \leq \theta_t^{ef}$), educated permanent migrants (with $\theta_t^{ef} \leq \theta < \theta_t^f$), and educated home-country workers who return upon nondiscovery (with $\theta \geq \theta_t^f$). Again, the partitioning is shown below.



As in the previous case, home-country workers consist of individuals with the lowest ability levels and the highest ability levels upon return-migration. The possibility of migration leads to the home country's permanent loss of all migrant workers with skill levels $\theta_t^{ef} < \theta < \theta_t^f$ since, from part 2 of proposition 1, $y_t^f(\theta) > y_t^{rd}(\theta)$ for every $\theta_t^{ef} \leq \theta \leq \theta_t^f$. Note also that since, by definition, θ_t^a is the average skill level of all undiscovered migrant workers at time t, while θ_t^{ef} is the skill level of the lowest ability migrant worker, it must be the case that $\theta_t^a > \theta_t^{ef}$, as shown in the diagram above.

Comparisons of θ^* and θ_t^{er}, and of θ^* and θ_t^{ef}, yield the following:

Proposition 3 *The fraction of the home-country population pursuing education in the presence of migration opportunities is always higher than the fraction of the home-country population pursuing education in the absence of migration opportunities.*

Proposition 3 reveals that the increase in the incentive to pursue education when migration offers a more attractive wage to the educated leads the home country to a higher degree of educational attainment. Yet, it should also be noted that the increase in the fraction of educated workers in the home country due to the prospect of migration does not necessarily imply that the number of educated workers who *stay and work* in the home country increases. To see this, consider the case of $\theta_t^{er} > \bar{\theta}$. From the definitions of θ_t^{ef} and θ^*, we have

$$\begin{aligned} w_s^h\theta^* &= m_t(w_s^f\theta_t^{ef} - k) + (1 - m_t)(w_s^f\theta_t^a - k) \\ \Leftrightarrow w_s^h\theta^* &= m_t(w_s^f\theta_t^{ef} - w_s^f\theta_t^a) + w_s^f\theta_t^a - k \\ \Leftrightarrow w_s^h\theta^* &< w_s^f\theta_t^a - k, \end{aligned}$$

where the last inequality follows as $\theta_t^{ef} < \theta_t^a$. In addition, since, by definition, $w_s^f\theta_t^a - k = w_s^h\theta_t^f$, we have $w_s^h\theta^* < w_s^h\theta_t^f$, or, $\theta^* < \theta_t^f$. Hence, the group of workers who acquire education in response to the prospect of migration (with skill levels $\bar{\theta} < \theta_t^{ef} \leq \theta < \theta^* < \theta_t^f$) belongs to the group of permanent migrants. As a result, the prospect of migration leads not only to a loss for the home country of those educated workers with $\theta > \theta_t^f (> \theta^*)$, who stay in the foreign country upon discovery, it also leads to the preclusion of any increase in the educated workforce in the home country, as a result of the possibility of migration. In what follows, we therefore focus our attention on the case in which $\theta_t^{er} < \bar{\theta}$, where the four "modes of employment" are present simultaneously.[9] As we elaborate below, the possible return-migration of those workers who would not have had the incentive to acquire education in the absence of migration opportunities, enables a possible economy-wide gain in spite of, and along with, a brain drain.

4 The Dynamics of Migration

With θ_t^{er} and θ_t^f defined, we now analyze the process of migration and the evolution of wage offers as experience with employing migrants accumulates over time. Given an initial experience associated with M_0, migration from the home country in subsequent periods can be summarized by the vector $\{\theta_t^a, \theta_t^{er}, \theta_t^f\}$, the elements of which are, in turn, solutions to the following system of simultaneous equations:

$$\theta_t^a = \frac{\int_{\theta_t^{er}}^{\theta_t^f} \theta f(\theta) d\theta}{F(\theta_t^f) - F(\theta_t^{er})} \tag{2}$$

$$\theta_t^f = \frac{w_s^f\theta_t^a - k}{w_s^h} \tag{3}$$

$$\theta_t^{er} = \frac{1}{m_t w_s^h}[\frac{1+\beta}{\beta}w_u^h - (1-m_t)(w_s^f\theta_t^a - k) + c]. \tag{4}$$

On multiplying both sides of equation (2) by w_s^f, the equation can be interpreted as requiring the wage offer to each migrant with unknown ability at time t to be equal to the average ability

[9]For $\theta_t^{er} < \bar{\theta}$, we require that

$$m_t w_s^h\bar{\theta} + (1-m_t)(w_s^f\theta_t^a - k) - c \geq \frac{(1+\beta)w_u^h}{\beta}.$$

Since $\theta_t^a > \bar{\theta}$, the left-hand side of the above is greater than $m_t w_s^h\bar{\theta} + (1-m_t)(w_s^f\bar{\theta} - k) - c = w_s^h\bar{\theta} - c = w_s^h k/(w_s^f - w_s^h) - c$. It follows that

$$m_t w_s^h\bar{\theta} + (1-m_t)(w_s^f\theta_t^a - k) - c > w_s^h\frac{k}{w_s^f - w_s^h} - c \geq \frac{(1+\beta)w_u^h}{\beta},$$

whenever w_u^h is sufficiently small and k is sufficiently large.

of the migrant cohort with unknown individual ability at time t, multiplied by the wage rate per efficiency unit of labor. Equations (3) and (4) require, respectively, that the extent of migration and the education decision follow from the expected utility maximization described in Section 3.[10] From equation (3), we observe further that:

$$\theta_t^a = \frac{w_s^h \theta_t^f + k}{w_s^f}, \tag{5}$$

and on rewriting equation (2),

$$= \frac{\int_{\theta_t^{er}}^{\theta_t^f} \theta f(\theta) d\theta}{F(\theta_t^f) - F(\theta_t^{er})}.$$

The first equality of equation (5) captures the supply side of the migrant labor market, that is, the foreign-country wage of an undiscovered migrant worker at time t, $w_s^f \theta_t^a$, is just sufficient to induce the supply of educated workers with ability $\theta \leq \theta_t^f$ who are willing to stay and work in the foreign country at the wage $w_s^f \theta_t^a$. The second equality of equation (5) holds that if θ_t^f represents the ability of the most able migrant worker who prefers $w_s^f \theta_t^a - k$ to his home wage, and θ_t^{er} represents the ability of the least able migrant worker, $1/w_s^f$ of the wage offer at time t (which reflects the willingness to pay for migrant work) is equal to the average ability of the migrant workforce, with unknown individual abilities, at time t.

Figures 3 and 4 depict the supply (SS) and demand (DD) relationships spelled out in the first and second parts of equation (5), respectively. The intersection points A in figure 3 and B in figure 4 depict equilibrium combinations of θ_t^a and θ_t^f that simultaneously satisfy equations (2) through (4), given m_t. It can be confirmed that both DD and SS are upward sloping.[11] Note also that in general, DD can be flatter or steeper than SS, depending on the exogenous parameters of the model. Consider, for example, the effect of an exogenous increase in the probability of discovery (m_t) when SS is steeper than DD, as in figure 3. An increase in m_t shifts the DD curve upward, while the SS curve remains unchanged.[12]

[10]A natural question is whether a solution to the above system exists. In the appendix, we provide an existence proof and spell out the required assumptions.

[11]From the first equality of equation (5), the slope of the supply relationship $(\partial\theta_t^a/\partial\theta_t^f)|_{SS}$ can be written as:

$$\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{SS} = \frac{w_s^h}{w_s^f} > 0.$$

From the second demand relationship in equation (5), we confirm in the appendix that the slope $(\partial\theta_t^a/\partial\theta_t^f)|_{DD}$ is

$$\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD} = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)/[F(\theta_t^f) - F(\theta_t^{er})]}{1 + \{(\theta_t^a - \theta_t^{er})f(\theta_t^{er})/[F(\theta_t^f) - F(\theta_t^{er})]\}[(1 - m_t)w_s^f/m_t w_s^h]} > 0.$$

[12]To see this, note from the second equality of equation (5) that for any given value of θ_t^f, an increase in m_t

An increase in m_t reduces the number of workers with low ability who acquire education at time $t-1$ $(1 - F(\theta_t^{er}))$ since workers endowed with rational expectations correctly anticipate the future value of m_t in their human capital calculus. As a result, the average ability (and hence the demand price) of migrants rises for any given θ_t^f because an increase in m_t shifts the skill composition of the migrant population in the foreign country to the right.

Note further that an increase in m_t has no direct effect on the supply side of the migrant labor market. θ_t^f divides the home-country population into two subgroups: A subgroup that consists of low-ability individuals (with $w_s^h\theta < w_s^f\theta_t^a - k$) who are better off remaining in the foreign country only if their true abilities are *not* discovered, and a subgroup that consists of individuals (with $w_s^h\theta \geq w_s^f\theta_t^a - k$) who receive a higher home wage than $w_s^f\theta_t^a - k$. It follows that θ_t^a alone determines the value of θ_t^f, given the wage schedules in the home country and the foreign country.

The new equilibrium pair of θ_t^a and θ_t^f is depicted as point A' (in figure 3), where both the average ability of migrants as well as θ_t^f rise as a result of an increase in m_t. In contrast, starting from a point such as B in figure 4, where SS is flatter than DD, an increase in m_t, together with the associated shift of the DD curve, imply a reduction in both θ_t^a and θ_t^f, as depicted by point B'. We denote the solutions to the system of simultaneous equations (2) - (4) as $\theta_t^j(m_t, c, w_u^h)$, $j = a, f, er$. Applying our arguments for the case of an increase in c and for the case of an increase in w_u^h, we obtain the first two parts of the following result; the third part is reasoned momentarily.

Proposition 4

1. *$\theta_t^a(m_t, c, w_u^h)$ is increasing in m_t, c, and w_u^h if and only if SS is steeper than DD or,*

leads to an upward shift of the DD curve since

$$\left.\frac{\partial\theta_t^a}{\partial m_t}\right|_{\theta_t^f\, const.} = \frac{(\theta_t^a - \theta_t^{er})f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})}\frac{\partial\theta_t^{er}}{\partial m_t} = \frac{(\theta_t^a - \theta_t^{er})f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})}\frac{w_s^f\theta_t^a - k - w_s^h\theta_t^{er}}{m_t w_s^h} > 0,$$

where the first equality follows from equation (10) in the appendix, and the second equality follows from equation (12) in the appendix. It follows, therefore, that DD shifts upward when m_t increases, or, $\left.\frac{\partial\theta_t^a}{\partial m_t}\right|_{\theta_t^f\, const.} > 0$. In addition, from the first equality of equation (5),

$$\left.\frac{\partial\theta_t^a}{\partial m_t}\right|_{\theta_t^f\, const.} = 0.$$

Hence, SS is independent of m_t.

equivalently, if and only if

$$1 - \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})}\frac{w_s^f}{w_s^h} + \frac{(\theta_t^a - \theta_t^{er})f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})}\frac{(1-m_t)w_s^f}{m_t w_s^h} > 0. \tag{6}$$

2. $\theta_t^f(m_t, c, w_u^h)$ *is increasing in* m_t, c, *and* w_u^h *if and only if equation (6) is satisfied.*

3. $\theta_t^{er}(m_t, c, w_u^h)$ *is increasing in* m_t *if and only if*

$$\frac{w_s^f\theta_t^a - k - w_s^h\theta_t^{er}}{(1-m_t)w_s^f} > \frac{\partial\theta_t^a}{\partial m_t}. \tag{7}$$

Regarding part 3 of proposition 4, note that from equations (2) - (4), $\partial\theta_t^{er}/\partial m_t = [(w_s^f\theta_t^a - k - w_s^h\theta_t^{er}) - (1-m_t)w_s^f\partial\theta_t^a/\partial m_t]/(m_t w_s^h)$. In general, therefore, an increase in m_t has an ambiguous effect on the incentives of workers with low ability levels to acquire education and migrate. The term $(w_s^f\theta_t^a - k - w_s^h\theta_t^{er}) > 0$, which is equal to the reduction in wages when the true ability of the marginal educated worker (the worker whose skill level is θ_t^{er}) is discovered, captures the negative incentive that an increase in m_t has on the education-cum-migration decision of low-ability workers. This negative incentive, however, coincides with the positive incentive that arises due to the increase in θ_t^a that, contingent on equation (6) holding, occurs as more high-ability workers migrate abroad because of an increased m_t. It follows that the negative incentive effect of an increase in m_t dominates the positive incentive effect whenever the increase in θ_t^a with respect to m_t is sufficiently low, as in equation (7), in which case, $\partial\theta_t^{er}/\partial m_t > 0$.

Proposition 4 completely summarizes the intertemporal variations of θ_t^a, θ_t^f, and θ_t^{er} for any given probability of discovery, m_t. Since migrant employment experience is cumulative, and the probability of discovery at any time $t+1$ depends on the accumulation of migrant employment experience until time $t-1$ ($m^{-1}(m_t)$) plus the increment in the total volume of migration at time t ($\mathcal{M}_t = N\{(1-m_t)[F(\bar{\theta}) - F(\theta_t^{er}(m_t, c, w_u^h))] + F(\theta_t^f(m_t, c, w_u^h)) - F(\bar{\theta}) + m_t[1 - F(\theta_t^f(m_t, c, w_u^h))]\}$), the law of motion governing the process of migration therefore depends only on the evolution of m_t with

$$m_{t+1} = \begin{cases} m[m^{-1}(m_t) + \mathcal{M}_t], & \text{if } m_t < \hat{m}; \\ \hat{m} & \text{otherwise}, \end{cases} \tag{8}$$

where $m^{-1}(m_1) = M_0$ is given.

A steady-state of equation (8) is denoted m^* such that $m_t = m_{t+1} = m^*$. The steady-state values of θ_t^j will be denoted as $\hat{\theta}^j$, $j = a, f, er$. The values of $\hat{\theta}^j$ are determined using equations (2) - (4) once m^* is determined.

Proposition 5 *If equation (6) is satisfied and the initial probability of discovery m_1 is such that $\theta_1^a(m_1, c, w_u^h) > \bar{\theta}$, then the only steady state equilibrium probability of discovery, m^*, is equal to $\hat{m}$.*

This result is straightforward from proposition 4.[13] In essence, the requirement that equation (6) be satisfied guarantees that accumulation of migrant employment experience and hence the probability of discovery, lead to a sequence of migratory moves from the home country over time. In the process, the average productivity of the migrants improves. This is due not only to an increase in the incentive for brighter individuals to migrate as the probability of discovery rises, but also because of the simultaneous decline in the willingness of the lowest ability individuals to acquire education and migrate. Such a cumulative process implies that the only long-run equilibrium consistent with an initial condition that yields a positive rate of migration is such that the probability of discovery no longer improves even when M_t increases.[14]

5 The Possibility of a Welfare Gain

Denote by $\hat{\theta}^{er}$ and $\hat{\theta}^f$ the solutions derived from the system of simultaneous equations (equations (2) - (4)) given $\hat{m}$.

At any time period, the total home-country population ($2N$) is distributed as follows: The N young individuals are divided into $NF(\hat{\theta}^{er})$, those who are uneducated and work and, $N[1 - F(\hat{\theta}^{er})]$, those who acquire education and do not work. The N old individuals are divided into $NF(\hat{\theta}^{er})$, those who are uneducated and work in the home country, and the rest, who engage in migration. These workers, in turn, consist of migrants who, with probability $\hat{m}$, return-migrate, and with probability $(1 - \hat{m})$, remain in the foreign country (consisting of $N[F(\bar{\theta}) - F(\hat{\theta}^{er})]$ individuals); permanent migrants ($N[F(\hat{\theta}^f) - F(\bar{\theta})]$); and migrants who,

[13]We are grateful to Yoram Weiss for pointing out that the steady state assumption could also be supported by an alternative experience accumulation formulation, in which a per-period depreciation rate can be used to capture the fact that recent migrants provide more information on the quality of the current wave of migrants.

[14]It bears emphasizing that proposition 5 also relies on an assumption made in proposition 1, that is, that $\theta_t^a > \bar{\theta}$. Otherwise, from equation (3),

$$\begin{aligned} \theta_t^f - \theta_t^a &= \frac{w_s^f \theta_t^a - k}{w_s^h} - \theta_t^a \\ &= \frac{w_s^f \theta_t^a - k - w_s^h \theta_t^a}{w_s^h} < 0 \end{aligned}$$

as $\theta_t^a < \bar{\theta}$. It follows that equation (2), which requires that θ_t^f be no smaller than θ_t^a, can never be satisfied and, accordingly, migration never takes off.

with probability $1-\hat{m}$, return-migrate and, with probability $\hat{m}$, remain in the foreign country ($N[1-F(\hat{\theta}^f)]$). There are thus $N\{\hat{m}[F(\bar{\theta})-F(\hat{\theta}^{er})]+(1-\hat{m})[1-F(\hat{\theta}^f)]\}$ workers at home who are return migrants, and there are $N\{(1-\hat{m})[F(\bar{\theta})-F(\hat{\theta}^{er})]+[F(\hat{\theta}^f)-F(\bar{\theta})]+\hat{m}[1-F(\hat{\theta}^f)]\} \equiv \hat{\mathcal{M}}^p$ workers who remain abroad.

Therefore, national output accrues from $2NF(\hat{\theta}^{er})$ workers who each produce w_u^h, from $\hat{m}N[F(\bar{\theta})-F(\hat{\theta}^{er})]$ workers who each produce w_s^h times their individual θ, and from $(1-\hat{m})N[1-F(\hat{\theta}^f)]$ workers who each produce w_s^h times their individual θ.

Denote by $V^m(\hat{\theta^{er}},\hat{\theta^f})$ the long run equilibrium value of the per-period national output in the home country, net of the cost of education. It follows that

$$\begin{aligned}
V^m(\hat{\theta^{er}},\hat{\theta^f}) &= 2NF(\hat{\theta^{er}})w_u^h + N\hat{m}\int_{\hat{\theta^{er}}}^{\bar{\theta}}(w_s^h\theta-c)f(\theta)d\theta \\
&\quad +N(1-\hat{m})\int_{\hat{\theta^f}}^{\infty}(w_s^h\theta-c)f(\theta)d\theta \\
&= N[2F(\hat{\theta^{er}})w_u^h + \hat{m}\int_{\hat{\theta}^{er}}^{\bar{\theta}}(w_s^h\theta-c)f(\theta)d\theta \\
&\quad +(1-\hat{m})\int_{\hat{\theta}^f}^{\infty}(w_s^h\theta-c)f(\theta)d\theta] \\
&= N[2F(\hat{\theta^{er}})w_u^h + \int_{\hat{\theta}^{er}}^{\infty}(w_s^h\theta-c)f(\theta)d\theta \\
&\quad -(1-\hat{m})\int_{\hat{\theta}^{er}}^{\bar{\theta}}(w_s^h\theta-c)f(\theta)d\theta - \int_{\bar{\theta}}^{\hat{\theta^f}}(w_s^h\theta-c)f(\theta)d\theta \\
&\quad -\hat{m}\int_{\hat{\theta}^f}^{\infty}(w_s^h\theta-c)f(\theta)d\theta] \\
&= N[2F(\hat{\theta^{er}})w_u^h + \int_{\hat{\theta}^{er}}^{\infty}(w_s^h\theta-c)f(\theta)d\theta - (w_s^h\hat{\theta^p}-c)\frac{\hat{\mathcal{M}^p}}{N}],
\end{aligned}$$

where

$$\hat{\theta^p} = \frac{N}{\hat{\mathcal{M}^p}}[(1-\hat{m})\int_{\hat{\theta^{er}}}^{\bar{\theta}}\theta f(\theta)d\theta + \int_{\bar{\theta}}^{\hat{\theta^f}}\theta f(\theta)d\theta + \hat{m}\int_{\hat{\theta^f}}^{\infty}\theta f(\theta)d\theta]$$

is the average ability of all migrant workers who stay abroad. The term $(w_s^h\hat{\theta^p}-c)\hat{\mathcal{M}^p}$ thus refers to the home-country output, net of the cost of education, that the home country forgoes when $\hat{\mathcal{M}^p}$ of its workers migrate and stay in the foreign country. To recall, $\hat{\mathcal{M}^p}$ is the per-period number of home-country workers employed abroad in a steady-state. Therefore, per capita output at home is

$$v^m(\hat{\theta^{er}},\hat{\theta^f}) = \frac{V^m(\hat{\theta^{er}},\hat{\theta^f})}{2N-\hat{\mathcal{M}^p}}.$$

Thus, $v^m(\hat{\theta^{er}}, \hat{\theta^f}) > v(\theta^*)$ if and only if

$$\frac{1}{2 - \hat{\mathcal{M}^p}/N}[2F(\hat{\theta^{er}})w_u^h + \int_{\hat{\theta}^{er}}^{\infty}(w_s^h\theta - c)f(\theta)d\theta - (w_s^h\hat{\theta^p} - c)\frac{\hat{\mathcal{M}^p}}{N}] >$$
$$\frac{1}{2}[2w_u^hF(\hat{\theta^*}) + \int_{\theta^*}^{\infty}(w_s^h\theta - c)f(\theta)d\theta] \equiv v(\theta^*),$$

or if and only if,

$$\frac{1}{2 - \hat{\mathcal{M}^p}/N}[2F(\hat{\theta^{er}})w_u^h + \int_{\hat{\theta}^{er}}^{\theta^*}(w_s^h\theta - c)f(\theta)d\theta + \int_{\theta^*}^{\infty}(w_s^h\theta - c)f(\theta)d\theta - (w_s^h\hat{\theta^p} - c)\frac{\hat{\mathcal{M}^p}}{N}] >$$
$$\frac{1}{2}[2w_u^hF(\hat{\theta^*}) + \int_{\theta^*}^{\infty}(w_s^h\theta - c)f(\theta)d\theta].$$

On manipulating the above equation, we obtain the following necessary and sufficient condition for $v^m(\hat{\theta^{er}}, \hat{\theta^f}) > v(\theta^*)$:

$$\{\int_{\hat{\theta}^{er}}^{\theta^*}(w_s^h\theta - c - 2w_u^h)f(\theta)d\theta + \frac{\hat{\mathcal{M}^p}}{N}[v(\theta^*) - (w_s^h\hat{\theta^p} - c)]\}\frac{1}{2 - \hat{\mathcal{M}^p}/N} > 0. \quad (9)$$

The first term in the curly brackets on the left-hand side of equation (9) reflects the gain in per capita output when the number of educated workers in the home country increases from $N(1 - F(\theta^*))$ to $N(1 - F(\hat{\theta^{er}}))$ as a result of the prospect of migration. In particular,

$$\begin{aligned}\int_{\hat{\theta}^{er}}^{\theta^*}(w_s^h\theta - c - 2w_u^h)f(\theta)d\theta &\equiv (w_s^h\hat{\theta^d} - c - 2w_u^h)[F(\theta^*) - F(\hat{\theta^{er}})] \\ &> 0\end{aligned}$$

if and only if $w_s^h\hat{\theta^d} - c > 2w_u^h$, where $\hat{\theta}^d$ denotes the average skill level of workers in the range $[\hat{\theta}^{er}, \theta^*]$. Hence, the first term of equation (9) is positive if and only if the average product of the increase in the educated work force in the s sector, net of the cost of education, is higher than the forgone output in the u sector. In particular, a sufficient condition for the above is that $w_s^h\hat{\theta^{er}} - c - 2w_u^h > 0$. From the definition of $\hat{\theta}^{er}$, this requires that

$$(2\hat{m} - \frac{1+\beta}{\beta})w_u^h < (1 - \hat{m})(c + k - w_s^f\hat{\theta^a}),$$

which, for example, is satisfied for sufficiently small $\hat{m}$ and/or sufficiently large c and k. From proposition 3, it follows that $w_s^h\hat{\theta}^{er} - c < w_s^h\theta^* - c$, and from the definition of θ^*, it follows that when $\beta = 1$, $w_s^h\theta^* = 2w_u^h + c$. Therefore, when $\beta = 1$, $w_s^h\hat{\theta}^{er} - c < 2w_u^h$. But if $w_s^h\hat{\theta}^{er} - c - 2w_u^h < 0$, the sufficient condition just referred to may not hold. That is, a gain in per capita income is less likely to occur. Recall our discussion in Section 2.3, in which we pointed out that when $\beta = 1$, individual utility maximization corresponds to the social optimum. Here again, we find that when $\beta = 1$, it is less likely that the migration prospect will lead to an improvement. However, if $\beta < 1$, the smaller the β, the larger the gain that will result from the increase in education prompted by the prospect of migration. This is nicely reflected by the increased likelihood that equation (9) will hold.

The second term in the curly brackets on the left-hand side of equation (9) reflects the change in per capita income resulting from a reduction in total population due to the loss of educated workers. In particular, this term is positive whenever the per capita home-country income of steady-state migrant workers, $w_s^h\hat{\theta}^p - c$, is less than the per capita home-country income in the absence of migration, $v(\theta^*)$. Note that the larger the total number of workers abroad in a steady state ($\hat{\mathcal{M}}^p$), the more significant will be the effect of this source of change in per capita output.

Proposition 6 *The per capita output in a country vulnerable to migration of skilled workers is higher than the per capita output in a country that is immune to migration if and only if equation (9) is satisfied.*

6 Conclusions

When an economy opens up to migration, workers in the economy are presented with a new set of opportunities and a new structure of incentives. While the expansion of opportunities results in human capital depletion, the revised incentives induce human capital formation: Higher returns to skills in the foreign country prompt more skill formation in the home country. We showed that the fraction of the home-country workforce acquiring education in the presence of migration opportunities is higher than the fraction of the home-country workforce undertaking education in the absence of migration opportunities.

Migration is also associated with a changing information environment, implying, in particular, that foreign-country employers are imperfectly informed about the skill levels of individual migrant workers. Consequently, migrants with different skill levels are pooled together and all are paid the same wage, which is based on the average product of the entire cohort of migrants. The imperfect but nonzero capability of employers to decipher true skill levels of individual migrants – captured in the probability of discovery – results in return-migration of both the highest and lowest skilled migrant workers, while permanent migrants are not drawn from the extremes of the skill distribution. Employers nevertheless become less ignorant over time. As their experience with employing migrants builds up, the probability of discovery rises. This progressive rise prompts a sequence of migratory moves characterized by a rising average skill level, until the probability of discovery reaches its steady-state equilibrium.

Accounting for the steady-state goings, comings, and skill formation, we showed that under well-specified conditions, per capita output in the home country is higher than that which would have obtained had the country altogether been immune to migration. An intriguing implication of this is that by allowing (rather than hindering) migration of skilled workers, the home-country population can enjoy higher welfare.[15] A drain of brains and a welfare gain need not be mutually exclusive and, as we have demonstrated, the former can be the very cause of the latter.

Appendix

Proof of proposition 1:

We proceed by stating the conditions under which $y_t^f(\theta) > y_t^{rd}(\theta)$ and $y_t^f(\theta) > y_t^{ru}(\theta)$. Now,

$$\begin{aligned} y_t^f(\theta) - y_t^{rd}(\theta) &= (1-m_t)(w_s^f\theta_t^a - k) + m_t(w_s^f\theta - k) - (1-m_t)(w_s^f\theta_t^a - k) - m_t(w_s^h\theta) \\ &= m_t(w_s^f\theta - k - w_s^h\theta) > 0 \end{aligned}$$

if and only if $\theta > \bar{\theta}$.

Similarly,

$$y_t^f(\theta) - y_t^{ru}(\theta) = (1-m_t)(w_s^f\theta_t^a - k) + m_t(w_s^f\theta - k) - (1-m_t)(w_s^h\theta) - m_t(w_s^f\theta - k) > 0$$

if and only if $\theta < \theta_t^f$ with

$$\theta_t^f = \frac{w_s^f\theta_t^a - k}{w_s^h}.$$

It remains to be shown that $\theta_t^f > \bar{\theta}$ and that $y_t^h(\theta) < \max[y_t^{rd}(\theta), y_t^f(\theta), y_t^{ru}(\theta)]$ for all θ. Now,

$$\begin{aligned} \theta_t^f - \bar{\theta} &= \frac{w_s^f\theta_t^a - k}{w_s^h} - \frac{k}{w_s^f - w_s^h} \\ &= \frac{w_s^f\theta_t^a}{w_s^h} - \frac{kw_s^h + k(w_s^f - w_s^h)}{w_s^h(w_s^f - w_s^h)} \\ &= \frac{w_s^f\theta_t^a}{w_s^h} - \frac{w_s^f}{w_s^h}\frac{k}{(w_s^f - w_s^h)} \\ &= \frac{w_s^f\theta_t^a}{w_s^h} - \frac{w_s^f\bar{\theta}}{w_s^h} \\ &> 0 \end{aligned}$$

[15] Note that this outcome holds independently of migrants remitting either some or none of their higher foreign earnings.

if and only if $\theta_t^a > \bar{\theta}$ (where the first and the fourth equalities follow from the definition of $\bar{\theta} = k/(w_s^f - w_s^h)$).

In addition, for $\theta < \theta_t^f$,

$$\begin{aligned} y_t^{rd}(\theta) - y_t^h(\theta) &= (1-m_t)(w_s^f\theta_t^a - k) + m_t w_s^h\theta - w_s^h\theta \\ &= (1-m_t)(w_s^f\theta_t^a - k) - (1-m_t)w_s^h\theta \\ &= (1-m_t)(w_s^f\theta_t^a - k - w_s^h\theta) > 0. \end{aligned}$$

Hence, $y_t^{rd}(\theta) > y_t^h(\theta)$ for $\theta < \theta_t^f$. Also, since $y_t^f(\theta) > y_t^{rd}(\theta)$, it must also be the case that $y_t^f(\theta) > y_t^h(\theta)$ for $\theta \in [\bar{\theta}, \theta_t^f)$. Finally, for $\theta \geq \theta_t^f$,

$$\begin{aligned} y_t^{ru}(\theta) - y_t^h(\theta) &= (1-m_t)(w_s^h\theta) + m_t(w_s^f\theta - k) - w_s^h\theta \\ &= m_t(w_s^f\theta - k) - m_t w_s^h\theta \\ &= m_t(w_s^f\theta - k - w_s^h\theta) > 0. \end{aligned}$$

It follows, therefore, that for all $\theta < \bar{\theta}$, $y_t(\theta) = y_t^{rd}(\theta)$; $\theta \in [\bar{\theta}, \theta_t^f)$, $y_t(\theta) = y_t^f(\theta)$; otherwise, $y_t(\theta) = y_t^{ru}(\theta)$, where $y_t(\theta)$, recall, is equal to $max[y_t^{rd}(\theta), y_t^f(\theta), y_t^{ru}(\theta), y_t^h(\theta)]$.

Proof of proposition 2:

1. The case of $\theta_t^{er} < \bar{\theta}$: We need to show that $\beta y_t(\theta) > (1+\beta)w_u^h$ for every $\theta > \theta_t^{er}$. From the proof of proposition 1, we have, for all $\theta < \bar{\theta}$, $y_t(\theta) = y_t^{rd}(\theta)$; $\theta \in [\bar{\theta}, \theta_t^f)$, $y_t(\theta) = y_t^f(\theta)$; and otherwise, $y_t(\theta) = y_t^{ru}(\theta)$. Hence, it is sufficient to show that: (A) for $\theta \in [\theta_t^{er}, \bar{\theta})$, $\beta y_t^{rd}(\theta) > (1+\beta)w_u^h$; (B) for $\theta \in [\bar{\theta}, \theta_t^f)$, $\beta y_t^f(\theta) > (1+\beta)w_u^h$ and finally, (C) $\beta y_t^{ru}(\theta) > (1+\beta)w_u^h$ for $\theta \geq \theta_t^f$.

(A) By the definition of θ_t^{er},

$$\begin{aligned} & \frac{1+\beta}{\beta}w_u^h = m_t w_s^h\theta_t^{er} + (1-m_t)(w_s^f\theta_t^a - k) - c \\ \Leftrightarrow \quad & \frac{1+\beta}{\beta}w_u^h < m_t w_s^h\theta + (1-m_t)(w_s^f\theta_t^a - k) - c \\ \Leftrightarrow \quad & (1+\beta)w_u^h < \beta y_t^{rd}(\theta) \end{aligned}$$

for any $\theta \geq \theta_t^{er}$. Clearly, it must also be the case that $(1+\beta)w_u^h < \beta y_t^{rd}(\theta)$ for any $\theta \in [\theta_t^{er}, \bar{\theta})$.
(B) Making use of the definition of θ_t^{er}, suppose that $\theta_t^{er} < \bar{\theta}$, we have

$$\begin{aligned} & [\frac{1+\beta}{\beta}w_u^h - (1-m_t)(w_s^f\theta_t^a - k) + c]\frac{1}{m_t w_s^h} < \bar{\theta} \\ \Leftrightarrow \quad & \frac{1+\beta}{\beta}w_u^h - (1-m_t)(w_s^f\theta_t^a - k) + c < m_t w_s^h\bar{\theta} \end{aligned}$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h < m_t w_s^h \bar{\theta} + (1-m_t)(w_s^f \theta_t^a - k) - c$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h < m_t (w_s^f \bar{\theta} - k) + (1-m_t)(w_s^f \theta_t^a - k) - c$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h < m_t (w_s^f \theta - k) + (1-m_t)(w_s^f \theta_t^a - k) - c = y_t^f(\theta)$$

for every $\theta > \bar{\theta}$. Note that the last but one inequality follows from the definition of $\bar{\theta}$ ($w_s^f \bar{\theta} - k = w_s^h \bar{\theta}$). It follows, therefore, that for every $\theta > \bar{\theta}$, $(1+\beta)w_u^h < \beta y_t^f(\theta)$ whenever $\theta_t^{er} < \bar{\theta}$.

(C) Since $y_t^f(\theta) < y_t^{ru}(\theta)$ for $\theta > \theta_t^f$, it follows from (B) that $(1+\beta)w_u^h < \beta y_t^f(\theta) < \beta y_t^{ru}(\theta)$ for every $\theta > \bar{\theta}$. In the proof of proposition 1, we have that $\theta_t^f > \bar{\theta}$. Hence, for every $\theta > \theta_t^f (> \bar{\theta})$, $(1+\beta)w_u^h < \beta y_t^f(\theta) < \beta y_t^{ru}(\theta)$.

2. The case of $\theta_t^{er} \geq \bar{\theta}$: We need to show that $\beta y_t(\theta) > (1+\beta)w_u^h$ for every $\theta > \theta_t^{ef}$. In particular, we need to show that (D) for all $\theta < \bar{\theta}$, $(1+\beta)w_u^h > y_t^{rd}(\theta)$; (E) for $\theta \in [\bar{\theta}, \theta_t^{ef})$, $(1+\beta)w_u^h > \beta y_t^f(\theta)$; (F) for $\theta \in [\theta_t^{ef}, \theta_t^f]$, $(1+\beta)w_u^h < \beta y_t^f(\theta)$; and finally (G) for $\theta > \theta_t^f$, $(1+\beta)w_u^h < \beta y_t^{ru}(\theta)$.

(D) Suppose that $\theta_t^{er} \geq \bar{\theta}$. By the definition of θ_t^{er}, we have

$$[\frac{1+\beta}{\beta} w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c] \frac{1}{m_t w_s^h} \geq \bar{\theta}$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c \geq m_t w_s^h \bar{\theta}$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h \geq m_t w_s^h \bar{\theta} + (1-m_t)(w_s^f \theta_t^a - k) - c$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h > m_t w_s^h \theta + (1-m_t)(w_s^f \theta_t^a - k) - c = y_t^{rd}(\theta)$$

for $\theta < \bar{\theta}$. It follows, therefore, that for $\theta < \bar{\theta}$, $(1+\beta)w_u^h > \beta y_t^{rd}(\theta)$.

(E) We shall first establish that $\theta_t^{ef} \geq \bar{\theta}$. If $\theta_t^{er} \geq \bar{\theta}$, we have

$$[\frac{1+\beta}{\beta} w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c] \frac{1}{m_t w_s^h} \geq \bar{\theta}$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c \geq m_t w_s^h \bar{\theta}$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h \geq m_t w_s^h \bar{\theta} + (1-m_t)(w_s^f \theta_t^a - k) - c$$

$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h \geq m_t(w_s^f \bar{\theta} - k) + (1-m_t)(w_s^f \theta_t^a - k) - c = y_t^f(\bar{\theta})$$
$$\Leftrightarrow \quad y_t^f(\theta_t^{ef}) = \frac{1+\beta}{\beta} w_u^h \geq y_t^f(\bar{\theta})$$
$$\Leftrightarrow \quad \theta_t^{ef} \geq \bar{\theta}$$

where the last but one line follows from the definition of θ_t^{ef}.

Now, we can make use of the definition of θ_t^{ef} once more, to establish that for $\theta \in [\bar{\theta}, \theta_t^{ef})$, $(1+\beta)w_u^h > \beta y_t^f(\theta)$. From the definition of θ_t^{ef} we have

$$\frac{1+\beta}{\beta} w_u^h = m_t(w_s^f \theta_t^{ef} - k) + (1-m_t)(w_s^f \theta_t^a - k) - c$$
$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h > m_t(w_s^f \theta - k) + (1-m_t)(w_s^f \theta_t^a - k) - c$$

for every $\theta < \theta_t^{ef}$. It follows, therefore, that for every $\theta < \theta_t^{ef}$, $(1+\beta)w_u^h > \beta y_t^f(\theta)$. This includes, of course, all $\theta \in [\bar{\theta}, \theta_t^{ef})$.

(F) Making use of the definition of θ_t^{ef},

$$\frac{1+\beta}{\beta} w_u^h = m_t(w_s^f \theta_t^{ef} - k) + (1-m_t)(w_s^f \theta_t^a - k) - c$$
$$\Leftrightarrow \quad \frac{1+\beta}{\beta} w_u^h < m_t(w_s^f \theta - k) + (1-m_t)(w_s^f \theta_t^a - k) - c$$

for every $\theta > \theta_t^{ef}$. It follows, therefore, that for every $\theta > \theta_t^{ef}$, $(1+\beta)w_u^h < \beta y_t^f(\theta)$. This includes, of course, all $\theta \in [\theta_t^{ef}, \theta_t^f]$.

(G) Recall that (F) states that $\beta y_t^f(\theta) > (1+\beta)w_u^h$ for every $\theta > \theta_t^{ef}$. This includes, as a subset, $\theta > \theta_t^f$ so that for $\theta > \theta_t^f$, $\beta y_t(\theta) > (1+\beta)w_u^h$. But for $\theta > \theta_t^f$, $y_t^{ru}(\theta) > y_t^f(\theta)$ or $\beta y_t^{ru}(\theta) > \beta y_t^f(\theta)$. Therefore, for $\theta > \theta_t^f$, $\beta y_t^{ru}(\theta) > \beta y_t^f(\theta) > (1+\beta)w_u^h$.

Proof of proposition 3:

We need to show that $\theta_t^{er} < \theta^*$. From figure 2, observe that to the left of $\bar{\theta}$, $y_t^{rd}(\theta) > y_t^h(\theta)$. Take $\theta = \theta_t^{er}$. Since

$$\begin{aligned} y_t^{rd}(\theta_t^{er}) &= m_t w_s^h \theta_t^{er} + (1-m_t)(w_s^f \theta_t^a - k) - c, \\ y_t^h(\theta_t^{er}) &= w_s^h \theta_t^{er} - c. \end{aligned}$$

$$w_s^h \theta_t^{er} - w_s^h \theta^* \quad < \quad m_t w_s^h \theta_t^{er} + (1-m_t)(w_s^f \theta_t^a - k) - w_s^h \theta^*$$

$$
\begin{aligned}
&= \frac{w_u^h(1+\beta)}{\beta} + c - w_s^h\theta^* \\
&= 0,
\end{aligned}
$$

where the last but one equality follows from the definition of θ_t^{er}, and the last equality follows from the definition of θ^*. Thus, $w_s^h\theta_t^{er} < w_s^h\theta^*$ or $\theta_t^{er} < \theta^*$.

To show that $\theta_t^{ef} < \theta^*$, observe from figure 2 that to the left of θ_t^f, $y_t^f(\theta) > y_t^h(\theta)$. Taking $\theta = \theta_t^{ef}$, we have,

$$
\begin{aligned}
y_t^f(\theta_t^{ef}) &= m_t(w_s^f\theta_t^{ef} - k) + (1-m_t)(w_s^f\theta_t^a - k) - c, \\
y_t^h(\theta_t^{ef}) &= w_s^h\theta_t^{ef} - c.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
w_s^h\theta_t^{ef} - w_s^h\theta^* &< m_t(w_s^f\theta_t^{ef} - k) + (1-m_t)(w_s^f\theta_t^a - k) - w_s^h\theta^* \\
&= \frac{w_u^h(1+\beta)}{\beta} + c - w_s^h\theta^* \\
&= 0,
\end{aligned}
$$

where the last but one equality follows from the definition of θ_t^{ef}, and the last equality follows from the definition of θ^*. Thus, $w_s^h\theta_t^{ef} < w_s^h\theta^*$ or $\theta_t^{ef} < \theta^*$.

The slope of the curve DD:

Since the DD curve depends on both θ_t^f and θ_t^{er}, we first make use of equation (4) to determine that

$$
\frac{\partial\theta_t^{er}}{\partial\theta_t^a} = -\frac{(1-m_t)w_s^f}{m_t w_s^h}.
$$

Differentiation of the second equality of equation (5) yields:

$$
\begin{aligned}
\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD} &= \frac{\theta_t^f f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} - \frac{\theta_t^{er} f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})}\frac{\partial\theta_t^{er}}{\partial\theta_t^a}(\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD}) \\
&\quad - \frac{\int_{\theta_t^{er}}^{\theta_t^f}\theta f(\theta)d\theta}{(F(\theta_t^f) - F(\theta_t^{er}))^2}[f(\theta_t^f) - f(\theta_t^{er})\frac{\partial\theta_t^{er}}{\partial\theta_t^a}(\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD})] \\
&= \frac{\theta_t^f f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} - \frac{\theta_t^{er} f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})}\frac{\partial\theta_t^{er}}{\partial\theta_t^a}(\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD}) \\
&\quad - \frac{\theta_t^a}{F(\theta_t^f) - F(\theta_t^{er})}[f(\theta_t^f) - f(\theta_t^{er})\frac{\partial\theta_t^{er}}{\partial\theta_t^a}(\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD})]
\end{aligned}
$$

$$= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} - \frac{(\theta_t^{er} - \theta_t^a) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{\partial \theta_t^{er}}{\partial \theta_t^a} \left(\frac{\partial \theta_t^a}{\partial \theta_t^f} |_{DD}\right)$$
$$= \frac{1}{\Delta} \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})},$$

where the second to last equality follows from the definition of θ_t^a in equation (2). Hence, since $\theta_t^f > \theta_t^a$, a necessary and sufficient condition for DD to be upward sloping is that $\Delta > 0$. To see that this is indeed the case, note that

$$\Delta = 1 + \frac{(\theta_t^{er} - \theta_t^a) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{\partial \theta_t^{er}}{\partial \theta_t^a}$$
$$= 1 + \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{(1 - m_t) w_s^f}{m_t w_s^h} > 0.$$

Proof of proposition 4:

1. We need to determine the relationships between θ_t^j, $j = a, f, er$ and the exogenous variables m_t, c and w_u^h, which are implicit in equations (2) - (4). By totally differentiating equation (2), we get

$$\begin{aligned} d\theta_t^a &= \frac{\theta_t^f f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^f - \frac{\theta_t^{er} f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^{er} \\ &\quad - \frac{\int_{\theta_t^{er}}^{\theta_t^f} \theta f(\theta) d\theta}{(F(\theta_t^f) - F(\theta_t^{er}))^2} [f(\theta_t^f) d\theta_t^f - f(\theta_t^{er}) d\theta_t^{er}] \\ &= \frac{\theta_t^f f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^f - \frac{\theta_t^{er} f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^{er} \\ &\quad - \frac{\theta_t^a}{F(\theta_t^f) - F(\theta_t^{er})} [f(\theta_t^f) d\theta_t^f - f(\theta_t^{er}) d\theta_t^{er}] \\ &= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^f + \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} d\theta_t^{er}, \end{aligned} \tag{10}$$

where the last but one line follows from the definition of θ_t^a in equation (2). Since $(\theta_t^f - \theta_t^a) > 0$ and $(\theta_t^{er} - \theta_t^a) < 0$, θ_t^a is increasing in θ_t^f and θ_t^{er}. The derivation above, of course, also confirms our claim in Section 4 that θ_t^a is strictly increasing in θ_t^f.

By totally differentiating equation (3), we obtain

$$d\theta_t^f = \frac{w_s^f}{w_s^h} d\theta_t^a. \tag{11}$$

Hence, θ_t^f is increasing in θ_t^a.

Turning now to the determination of θ_t^{er}, we get, by totally differentiating equation (4),

$$\begin{aligned} d\theta_t^{er} &= \frac{1+\beta}{\beta m_t w_s^h} dw_u^h + \frac{1}{m_t w_s^h} dc - \frac{(1-m_t) w_s^f}{m_t w_s^h} d\theta_t^a + \frac{w_s^f \theta_t^a - k}{m_t w_s^h} dm_t \\ &\quad - \frac{1}{(m_t w_s^h)^2} [\frac{1+\beta}{\beta} w_u^h - (1-m_t)(w_s^f \theta_t^a - k) + c] w_s^h dm_t \\ &= \frac{1+\beta}{\beta m_t w_s^h} dw_u^h + \frac{1}{m_t w_s^h} dc - \frac{(1-m_t) w_s^f}{m_t w_s^h} d\theta_t^a + \frac{w_s^f \theta_t^a - k}{m_t w_s^h} dm_t - \frac{w_s^h \theta_t^{er}}{m_t w_s^h} dm_t \\ &= \frac{1+\beta}{\beta m_t w_s^h} dw_u^h + \frac{1}{m_t w_s^h} dc - \frac{(1-m_t) w_s^f}{m_t w_s^h} d\theta_t^a + \frac{w_s^f \theta_t^a - k - w_s^h \theta_t^{er}}{m_t w_s^h} dm_t. \end{aligned} \tag{12}$$

It follows that θ_t^{er} is increasing in w_u^h and c but decreasing in θ_t^a. Also, since $\theta_t^a > \bar{\theta}$ by the assumption in Section 3.2, $w_s^f \theta_t^a - k > w_s^f \bar{\theta} - k = w_s^h \bar{\theta} > w_s^h \theta_t^{er}$, where the last inequality follows from our assumption that $\theta_t^{er} < \bar{\theta}$ in Section 3.2. Hence, θ_t^{er} is increasing in m_t, all else remaining constant.

We next examine the relationship between θ_t^a and m_t, holding all else constant. By substituting equations (11) and (12) into equation (10), we obtain

$$\begin{aligned} d\theta_t^a &= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{\mathcal{M}_t / N} (\frac{w_s^f}{w_s^h} d\theta_t^a) \\ &\quad - \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} (\frac{(1-m_t) w_s^f}{m_t w_s^h} d\theta_t^a - \frac{w_s^f \theta_t^a - k - w_s^h \theta_t^{er}}{m_t w_s^h} dm_t) \\ &= \frac{1}{\Omega} [\frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{w_s^f \theta_t^a - k - w_s^h \theta_t^{er}}{m_t w_s^h}] dm_t \end{aligned} \tag{13}$$

where

$$\Omega = 1 - \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{er})} \frac{w_s^f}{w_s^h} + \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{(1-m_t) w_s^f}{m_t w_s^h}.$$

The numerator in the last line of equation (13) is positive since $w_s^f \theta_t^a - k - w_s^h \theta_t^{er} > 0$. Therefore, θ_t^a is increasing in m_t if and only if $\Omega > 0$, as stated in proposition 4.

Substituting equation (12) into equation (10), keeping m_t and w_u^h constant, we obtain

$$d\theta_t^a = \frac{1}{\Omega} \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{1}{m_t w_s^h} dc. \tag{14}$$

It follows that θ_t^a is increasing in c if and only if $\Omega > 0$. Finally, holding m_t and c constant, we obtain, on substituting equation (12) into (10),

$$d\theta_t^a = \frac{1}{\Omega} \frac{(\theta_t^a - \theta_t^{er}) f(\theta_t^{er})}{F(\theta_t^f) - F(\theta_t^{er})} \frac{1+\beta}{\beta m_t w_s^h} dw_u^h. \tag{15}$$

Hence, θ_t^a is also increasing in w_u^h if and only if $\Omega > 0$.

2. Turning now to θ_t^f, from equation (11) we obtain

$$d\theta_t^f = \frac{w_s^f}{w_s^h}\frac{\partial \theta_t^a}{\partial m_t} dm_t.$$

Hence, a necessary and sufficient condition for θ_t^f to be increasing in m_t is that $\partial\theta_t^a/\partial m_t > 0$. From equation (13), we have already determined that θ_t^a is increasing in m_t if and only if $\Omega > 0$. It follows that $\Omega > 0$ is necessary and sufficient for θ_t^f to be increasing in m_t.

In a similar fashion, we can determine, using equations (11) and (14), that

$$d\theta_t^f = \frac{w_s^f}{w_s^h}\frac{\partial \theta_t^a}{\partial c} dc.$$

It follows that θ_t^f is also increasing in c under the condition $\Omega > 0$ since $(\partial\theta_t^a/\partial c) > 0$, from equation (14). Finally,

$$d\theta_t^f = \frac{w_s^f}{w_s^h}\frac{\partial \theta_t^a}{\partial w_u^h} dw_u^h.$$

From equation (15), $(\partial\theta_t^a/\partial w_u^h) > 0$ if $\Omega > 0$; hence, θ_t^f is increasing in w_u^h under the condition $\Omega > 0$.

3. To determine the relationship between θ_t^{er} and m_t, note, from equation (12), that all else remaining constant,

$$d\theta_t^{er} = -\frac{(1-m_t)w_s^f}{m_t w_s^h}\left(\frac{\partial \theta_t^a}{\partial m_t}\right) dm_t + \frac{w_s^f\theta_t^a - k - w_s^h\theta_t^{er}}{m_t w_s^h} dm_t$$

From equation (13), we obtain the result that $(\partial\theta_t^a/\partial m_t) > 0$, if and only if $\Omega > 0$. Since $(w_s^f\theta_t^a - k - w_s^h\theta_t^{er}) > 0$, as already pointed out in our discussion following equation (12), we have that θ_t^{er} is increasing in m_t if and only if

$$\frac{w_s^f\theta_t^a - k - w_s^h\theta_t^{er}}{m_t w_s^h} > \frac{(1-m_t)w_s^f}{m_t w_s^h}\left(\frac{\partial \theta_t^a}{\partial m_t}\right),$$

or if and only if

$$\frac{w_s^f\theta_t^a - k - w_s^h\theta_t^{er}}{(1-m_t)w_s^f} > \frac{\partial \theta_t^a}{\partial m_t},$$

as stated in equation (7).

Proof of proposition 5:
Since $\theta_1^a > \bar{\theta}$,

$$\begin{aligned}
\theta_1^f - \theta_1^a &= \frac{w_s^f \theta_1^a - k}{w_s^h} - \theta_1^a \\
&= \frac{w_s^f \theta_1^a - k}{w_s^h} - \frac{w_s^h \theta_1^a}{w_s^h} \\
&= \frac{w_s^f \theta_1^a - k - w_s^h \theta_1^a}{w_s^h} > 0
\end{aligned}$$

and hence, there is positive migration at $t = 1$ with $M_1 = M_0 + \mathcal{M}_1 > M_0$ or, equivalently, $m_2 = m(M_1) > m(M_0) = m_1$. Also, since M_t can be no less than M_0, satisfaction of equation (6) implies that $\theta_t^a(m_t, c, w_u^h) > \theta_1^a(m_1, c, w_u^h)$ and hence $\theta_t^a(m_t, c, w_u^h) > \bar{\theta}$ for all $t = 2, 3, 4...$ In addition, equation (6) also guarantees that $\theta_t^f(m_t, c, w_u^h) > \theta_1^f(m_1, c, w_u^h)$ since θ_t^f is increasing in m_t for any t.

Finally, since $\theta_t^{er} < \bar{\theta}$, we have $\theta_t^f > \theta_t^a > \bar{\theta} > \theta_t^{er}$ and $M_t = M_{t-1} + \mathcal{M}_t > M_{t-1}$ for all t. It follows immediately that $M_{t+i} \geq M_t$, $i = 1, 2, ...$ Hence, the only long run equilibrium probability of discovery must correspond to the upper bound $\hat{m}$.

Existence
To determine whether there exists at least one set of solutions, θ_t^j, $(j = a, f, er)$, to equations (2) - (4) for every m_t, which satisfies the requirement in proposition 1 that $\theta_t^a > \bar{\theta}$, we only need to show that there exists at least one θ_t^f for every m_t at which the SS and DD curves intersect. Once θ_t^f is determined, the first equality of equation (5) can be used to determine θ_t^a. Finally, the value of θ_t^{er} can also be calculated from equation (4) once θ_t^a is determined.

Consider the right-hand side of the first equality in equation (5). Note that as $\theta_t^f \to \infty$, $\theta_t^a \to \infty$ In addition, as $\theta_t^f \to \infty$, the right-hand side of the second equality of equation (5) is finite since, by assumption, θ has a finite expectation. It follows that for sufficiently large θ_t^f, SS lies above DD. By the intermediate value theorem, existence is guaranteed if and only if DD lies above SS for some $\theta_t^f > \bar{\theta}$ or if and only if

$$\frac{\int_{\theta_t^{er}}^{\theta_t^f} \theta f(\theta) d\theta}{F(\theta_t^f) - F(\theta_t^{er})} > \frac{w_s^h \theta_t^f}{w_s^f}.$$

The requirement that there exists a $\theta_t^f > \bar{\theta}$ such that DD lies above SS guarantees that the *equilibrium* value of θ_t^a is strictly greater that $\bar{\theta}$. From the definition of θ_t^f in Section 3,

$\theta_t^f > \bar{\theta}$ implies that

$$\begin{aligned}
& \frac{w_s^f \theta_t^a - k}{w_s^h} > \bar{\theta} \\
\Leftrightarrow \quad & w_s^f \theta_t^a > w_s^h \bar{\theta} + k \\
\Leftrightarrow \quad & w_s^f \theta_t^a > w_s^h \bar{\theta} + (w_s^f - w_s^h)\bar{\theta} \\
\Leftrightarrow \quad & w_s^f \theta_t^a > w_s^f \bar{\theta} \\
\Leftrightarrow \quad & \theta_t^a > \bar{\theta}.
\end{aligned}$$

References

Berry, Albert R. and Soligo, Ronald 1969. "Some Welfare Aspects of International Migration." *Journal of Political Economy* 77, 778-794.

Bhagwati, Jagdish. and Wilson, John D. 1989. *Income Taxation and International Mobility.* Cambridge, MA: MIT Press.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77, 531-553.

Carrington, William J., Detragiache, Enrica and Vishwanath, Tara. 1996. "Migration with Endogenous Moving Costs." *American Economic Review* 86, 909-930.

DaVanzo, Julie. 1983. "Repeat Migration in the United States: Who Moves Back and Who Moves On?" *Review of Economics and Statistics* 65, 552-559.

Grubel, Herbert G. and Scott, Anthony. 1966. "The International Flow of Human Capital." *American Economic Review* 56, 268-274.

Katz, Eliakim and Stark, Oded. 1987. "International Migration Under Asymmetric Information." *Economic Journal* 97, 718-726.

——. 1989. "International Migration under Alternative Informational Regimes: A Diagrammatic Analysis."*European Economic Review* 33, 127-142.

Kwok, Peter V. and Leland, Hayne. 1982. "An Economic Model of the Brain Drain." *American Economic Review* 72, 91-100.

LaLonde, Robert J. and Topel, Robert H. 1997. "Economic Impact of International Migration and the Economic Performance of Migrants." In Rosenzweig, Mark R. and Stark, Oded (eds.) *Handbook of Population and Family Economics.* Amsterdam: North-Holland.

Ravenstein, Ernest George. 1885. "The Laws of Migration." *Journal of the Royal Statistical Society* 48, 167-227.

Razin, Assaf and Sadka, Efraim. 1997. "International Migration and International Trade." In Rosenzweig, Mark R. and Stark, Oded (eds.) *Handbook of Population and Family Economics.* Amsterdam: North-Holland.

Reilly, Barry. 1994. "What Determines Migration and Return? An Individual Level Analysis Using Data for Ireland." University of Sussex. Mimeo.

Stark, Oded. 1991. *The Migration of Labor.* Oxford and Cambridge, MA: Basil Blackwell.

——. 1995. "Frontier Issues in International Migration." *Proceedings of the World Bank Annual Conference on Development Economics, 1994*. Washington, D.C.: The International Bank for Reconstruction and Development, 361-386.

Figure 1

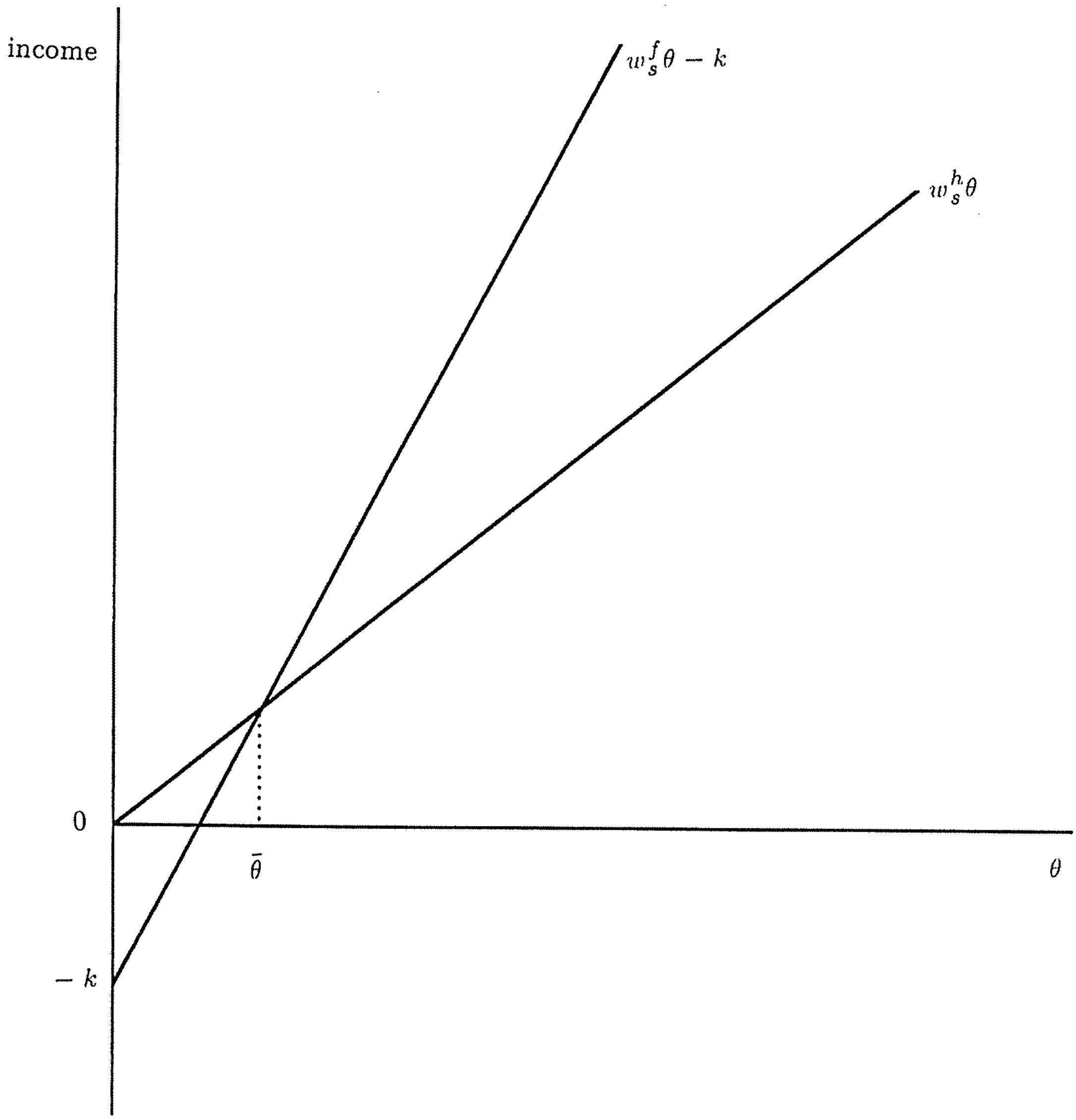


Figure 2

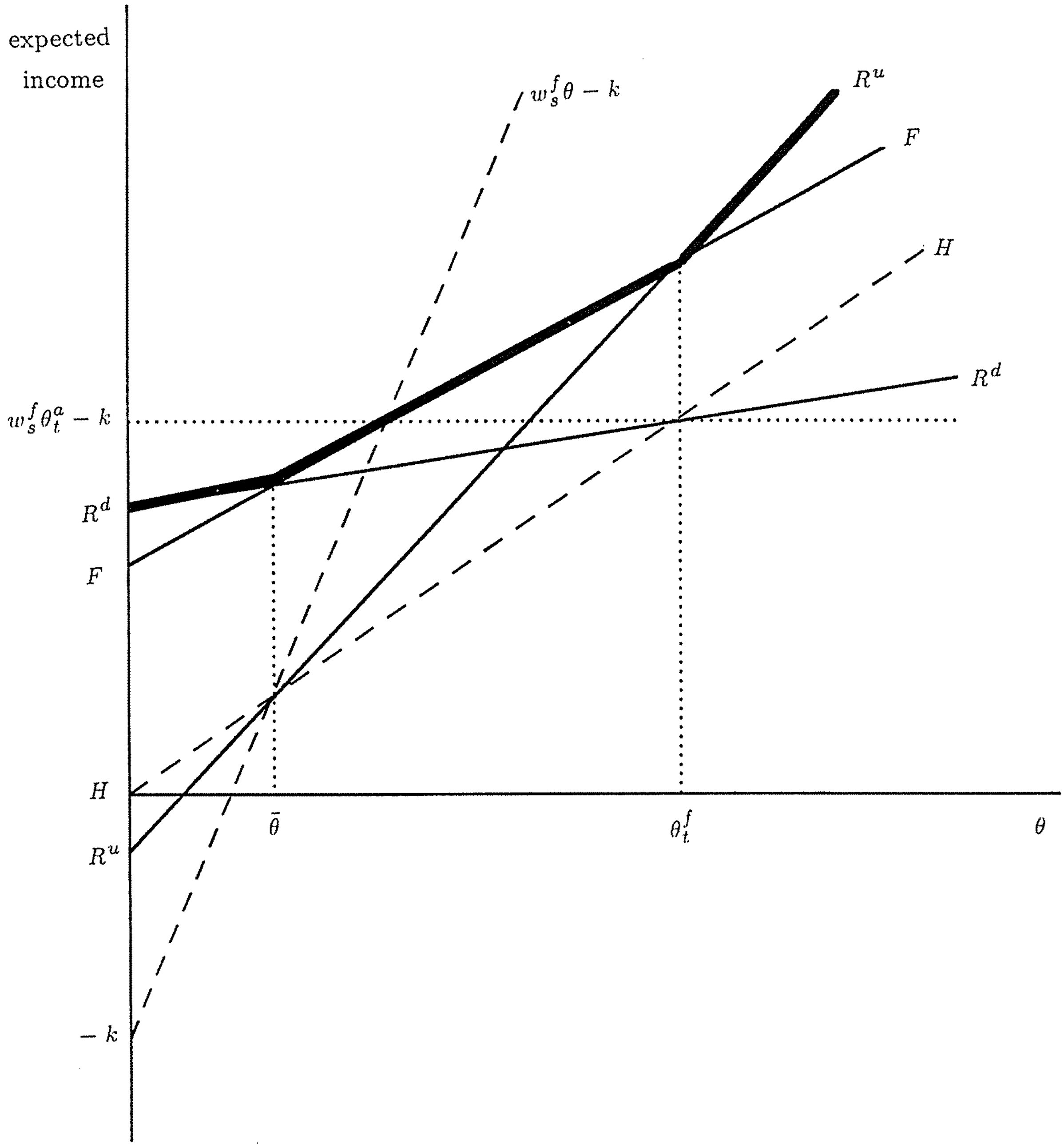

expected
income
$w_s^f\theta - k$
R^u
F
H
R^d
$w_s^f\theta_t^a - k$
R^d
F
H
R^u
$-k$
$\bar{\theta}$
θ_t^f
θ

Figure 3

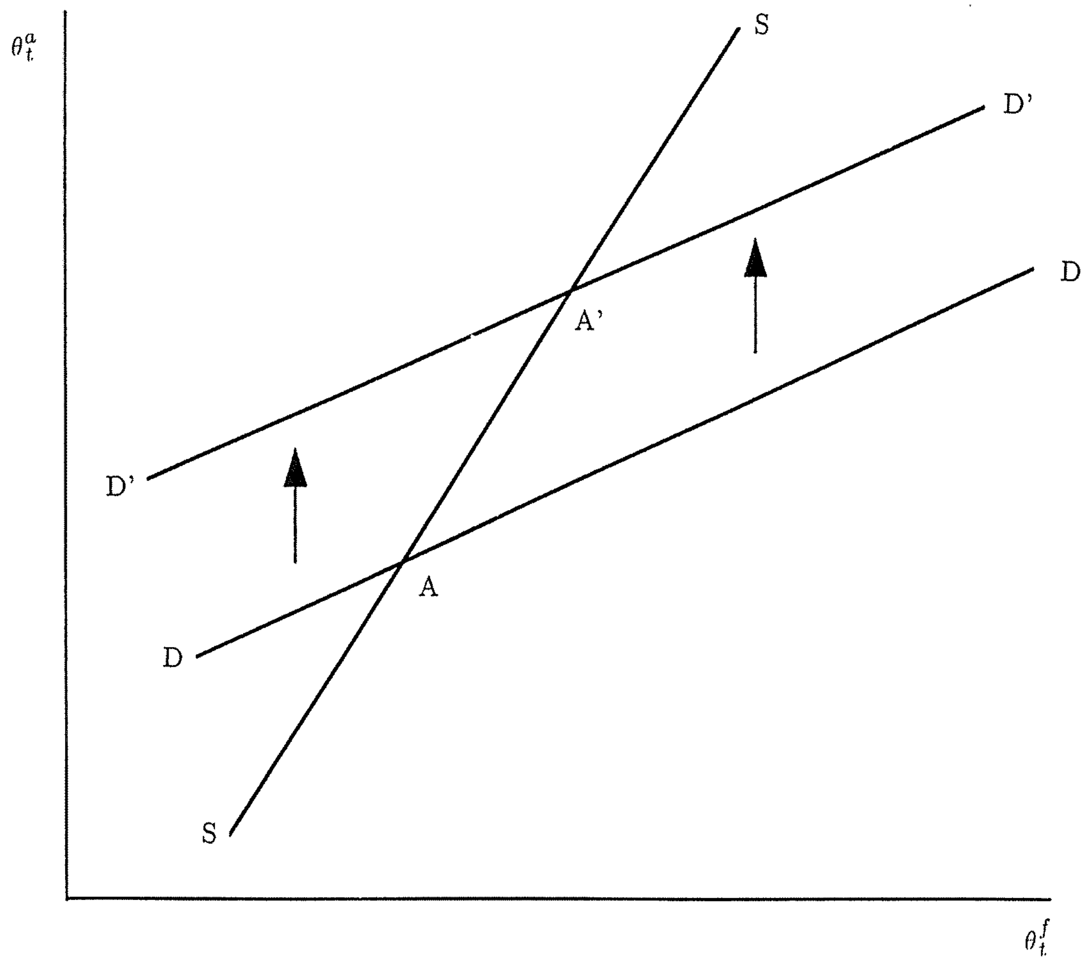
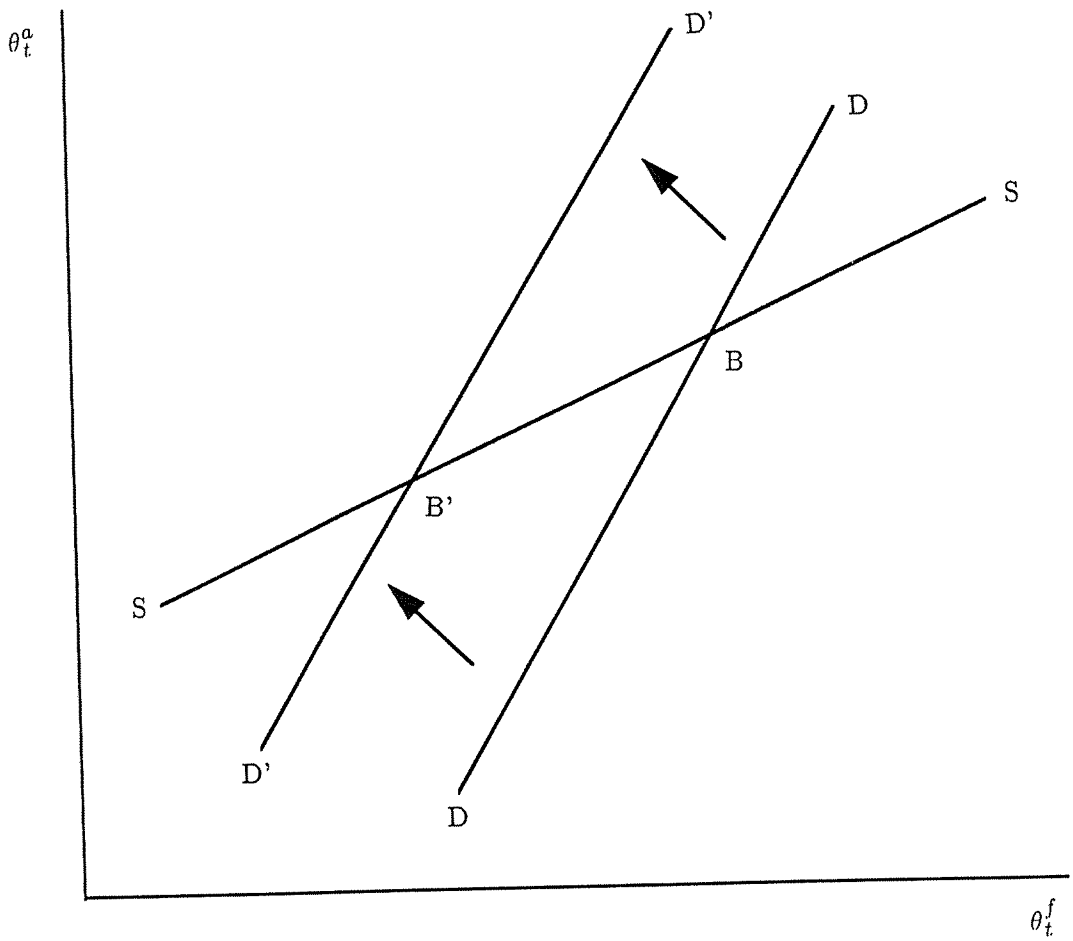


Figure 4



Institut für Höhere Studien
Institute for Advanced Studies
Stumpergasse 56
A-1060 Vienna
Austria

Phone: +43-1-599 91-145
Fax: +43-1-599 91-163