No Brain Gain without Brain Drain? Dynamics of Return Migration and Human Capital Formation under Asymmetric Information

Ozan Hatipoglu*and Serhan Sadıkoglu†

March 10, 2013

Abstract

We study the migration and return migration decisions of skilled workers, along with the impact of migration prospects on human capital formation in the source country under asymmetric information. In addition, we analyze the dynamics of migration and return migration as informational asymmetries and migration costs evolve over time as a result of migration networks. We find that skilled migration is followed by return migration which involves both positive and negative selection of skilled migrants. Furthermore, we show that the possibility of return migration under asymmetric information mitigates the brain draining effect of initial migration prospect and has a positive impact on human capital formation in the source country. Finally, we derive the conditions under which the possibility of migration leads to welfare gain in the source country.

JEL classification: J24, F22, O15, 040

Keywords: Asymmetric Information, Migration, Return Migration, Human Capital.

*Corresponding Author. Department of Economics, Natuk Birkan Hall, Bogazici University, Bebek, Istanbul 34342, Turkey. Email: `ozan.hatipoglu@boun.edu.tr` Phone: +90 212 3597651

† Department of Economics, Room K 641, Tillburg University, PO Box 90153 5000 LE Tilburg, Netherlands. Email: `s.sadikoglu@uvt.nl` Phone: +31 13 466 4220

1 Introduction

International migration of skilled workers has generated substantial amount of research in the discipline of economics and an equally high number of controversial debates among policy circles. The conventional view asserts that migration of skilled workers, the so-called brain drain, is detrimental for the country of origin since in the absence of migration source country would have had a more skilled workforce, and per capita output and national welfare would be higher. Accordingly, the brain drain is considered as a negative externality on the remaining population in the country of origin. Recent studies, however, have been questioning this view by arguing that migration of skilled workers can be associated with some positive effects that can compensate the negative effect of loss of skilled workers.

As a case in point, one branch of the literature on migration of skilled workers studies the effects of migration possibilities on human capital formation in the source country. The key argument presented in this line of research is the following: When an economy opens up to migration, workers would have more incentives for skill acquisition since they have the possibility to migrate to a country where return to skills is higher. As a result, the skilled fraction of the source country would increase and, provided that only a fraction of skilled workers would migrate, the average skill level of the population left in this economy might increase in comparison to autarky. Thus, migration of skilled workers would induce a positive externality on the remaining population[1].

Our paper contributes to this existing body of literature in a couple of ways. Firstly, we introduce the return migration choice into migration decisions of workers to analyze the effects of the possibility of return migration on initial migration choice. Introduction of a return decision choice enables us to analyze the effect of foreign work experience premium on employment and migration choices of domestic workers as well as on human capital formation in the source country. Secondly, we characterize the workers who engage in

[1] This is the so called "Brain Gain with Brain Drain" argument.

migration and return migration under perfect and asymmetric information. Under perfect information, we show that skill acquisition in the source country is higher when return migration is a possibility. We observe that as migration costs decrease, skill acquisition increases and, migration becomes less positively selected while return migration exhibits more negative selection. Under asymmetric information, we derive the set of conditions under which average human capital of migrants rise over time. Finally, we run a welfare analysis, where we derive the conditions where brain drain leads to welfare gain in the source country.

Positive effects of migration have been emphasized by Stark et al. (1998) who show that in an economy populated with homogeneous workforce, a positive probability of migration provides an incentive for higher per worker investment in human capital formation in the source country and the average level of human capital in the economy might rise. Vidal (1998) presents a two-period overlapping generations (OLG) model which incorporates intergenerational externality of average human capital for each generation. The model shows that for a well-defined level of probability of migration, the source country might experience a higher long-run economic growth. Mountford (1997) develops a three-period OLG model which considers a heterogeneous labor force in the sense that the individuals in the source country are endowed with heterogenous latent abilities. It is shown that the migration of skilled workers might be beneficial in terms of a higher growth rate of the source country if the probability of migration is sufficiently low along with sufficiently high wage differential between the source and host countries. Beine et al. (2001) present a two-period dynamic migration model which concentrates on the human capital formation and economic growth in the source country. They argue that migration of skilled workers might lead to a higher steady-state growth rate if the source country's growth rate is already relatively high along with intermediate values for migration probability of skilled workers in the source country. Moreover, Hemmi (2005) extends the model of Beine et al. (2001) by introducing a fixed migration cost to the model. In that context, the transitional dynamics is also explored and it is concluded that although source country might experience a higher steady-state growth rate, it is also possible to exhibit a relatively low growth performance along the transition path. Stark (2004) takes a social planner's perspective and investigates whether there exists an optimum migration probability which maximizes source country's welfare by increasing skill acquisition in the source country. It is argued that migration probability, which yields social optimum outcome, is strictly positive, hence migration of skilled workers might lead to an increase in welfare along with a rise in the average level of human capital in the source country.

All the articles above assume a perfect information setting. Put differently, not only the employers in the home country, but also the employers in the foreign country perfectly observe the skill level of workers and offer wage rates accordingly. The simplifying nature of the perfect information assumption was made clear by Stark et al. (1997) who argue that foreign employers are less capable of assessing the skill level of migrant workers since the home country's information structure differs from the foreign country's. Relaxing the homogeneous information structure, Stark et al. (1997) present a two period-static model on migration decisions of low-skill and high-skill workers. The model assumes that in the first period the migrants are offered a wage rate depending on the average skill level of migrant cohort and in the second period migrants are paid according to their skill levels. Under this framework with exogenously given monitoring capabilities of foreign employers, it is shown that human capital investments increase with the possibility of migration. Moreover, return migration of low-skill workers in the second period is figured out as an additional channel for an increase in the human capital.

Our paper is a contribution to the strand of literature that analyzes dynamics of migration under asymmetric information. It is closely linked to Chau and Stark (1999) who assume that foreign employers' capability of monitoring the skill levels of migrants enhance over time as foreign employers become more experienced with employing migrants. In this setting, they introduce an endogeneous skill acquisition structure and examine how human capital formation decisions of home country workers vary over time. Besides, by the dynamic nature of the model, intertemporal variations in the migration and return migration, which stem from the asymmetric information on the part of foreign employers, are explored. It is argued that as the experience of employing migrant workers accumulates, relatively low-skill workers return-migrate and high-skill workers become permanent migrants. Furthermore, it is shown that when migration is a possibility, average level of human capital in the home country increases along with a welfare gain in the home country under well-specified conditions.

Even though Chau and Stark (1999) present a novel theoretical framework in order to analyze the process of migration, the model does not incorporate some aspects of international labor migration. First is that the model does not capture the role of migrant network in labor migration. Noticing that migration entails a cost, which might be physical such that transportation or initial expenditure for settlement in the foreign country or might be psychological due to such as leaving one's own country, the migrant networks lead to a reduction in the migration cost. The cost-reducing role of migrant networks is theoretically formulated by Carrington et al. (1996), which try to find a plausible explanation to the fact that the migratory flow of the Southern Blacks to the North in the U.S gained momentum while the income differentials diminished. It is concluded that as the migrant network in the foreign country expands over time, even if the wage differential between the home and foreign countries narrows, the flow of migrants increases due to low migration costs. The second feature of labor migration, which Chau and Stark (1999) model does not take into account, is linked to the return migrants. Barrett and O'Connell (2001) argue that Irish return migrants earn a 10-15% higher wage than similar workers, who did not migrate, and Iara (2008) provides evidence for a wage premium for workers who have had work experience in Western Europe. In theoretical grounds, Peri and Mayr (2008), Dos-Santos and Vinay (2003) present models which claim that if there is a productivity premium for return migrants, -thus a wage premium- since they enhance their human capital by acquiring new skills and techniques in the foreign country, which is assumed to be technologically superior to the home country, then return migration would serve as a positive channel which increases the average level of human capital in the home country.

In this context, we extend the model of Chau and Stark (1999) in two main directions. First, we relax constant migration cost assumption and define it as a function of number of migrants in a similar manner by Carrington et al (1996). Second, we construct a three-period model with return migration which yields a foreign work experience premium for return migrants. As a slight modification, we model the human capital formation in a different setting than Chau and Stark's model.

Our findings are as follows. We characterize the workers who engage in migration and return migration under perfect information and asymmetric information. Under perfect information, we show that skill acquisition in the source country is higher when return migration is a possibility. We observe that as migration costs decrease, skill acquisition increases and, migration becomes less positively selected while return migration exhibits more negative selection. Under asymmetric information, we find that average human capital of migrants rise over time under certain conditions and conducting a welfare analysis, we derive the conditions to experience a welfare gain in the source country.

This paper is organized as follows. In the next section, we introduce the building blocks of our model and analyze the dynamics of migration under perfect information. In the third section, we derive our main results under asymmetric information. In the fourth section, we analyze the dynamics of return migration, and in the fifth section we provide a welfare analysis. Section six concludes.

2 The Model

At each period of time, a single composite good is produced according to the constant returns to scale production function $Y_t = AH_t$, where H_t is the labor input which is measured in efficiency units in the home country h. Thus, A, which is total factor productivity, is also the marginal and average product of an efficiency unit of labor. It is assumed that both factor and output markets are perfectly competitive. Hence, the wage paid to a worker is determined by marginal product of labor and the wage paid for an efficiency unit of work is $w = A$.

At each time period t, N individuals are born and individuals live for three periods. Following Chau and Stark (1999), we characterize individuals by their endowments and preferences. Each individual is endowed with one efficiency unit of labor upon being born and innate ability $\theta \in [0, \infty)$. Further, individuals are endowed with different levels of innate ability, which is distributed by a cumulative distribution function $F(\theta)$, over the home country population. $F(\theta)$ is continuously differentiable, has a strictly positive density function $f(\theta)$ and $\int_0^\infty \theta f(\theta) d\theta$ is finite. Moreover, it is assumed that all generations have innate abilities, which are distributed with the same cumulative distribution function $F(\theta)$, and the abilities of younger generations do not depend on the abilities of older generations. Regarding the individuals' preferences, all individuals have identical preferences represented by a utility function $u(y_t, y_{t+1}, y_{t+2})$, where y_t is the income at time period t. For the sake of simplicity, utility function is defined as: $u(y_t, y_{t+1}, y_{t+2}) = y_t + \beta y_{t+1} + \beta^2 y_{t+2}$, where $0 < \beta < 1$ is the time discount rate.

In the first period of life, individuals have the possibility to spend their time and invest resources to acquire education, which increases their supply of efficiency units of labor. Individuals are assumed to incur the fixed cost of e units of output to undertake education and become skilled workers. Since individuals do not have any resources of their own, individuals must borrow e from the credit market and repay their debt in the second period of their lives. In order to simplify, the interest is assumed to be zero. In this context, the education and ability level of an individual are related to his/her human capital level by the following human capital formation function:

$$h_t = (1 + \theta e^{\gamma}), \tag{1}$$

where $\gamma > 0$. Considering the human capital formation described above, the uneducated individual supplies one efficiency unit of labor, which is independent of his innate ability level and is equal to his endowment upon being born. Furthermore, the efficiency units of labor supplied by a skilled worker is given by the expression $(1 + \theta e^{\gamma})$, which depends on worker's innate ability level and the reward to education.

When the economy opens up to migration, skilled workers have a pair of employment options. They might choose to work in the home country or in the foreign country. Not only for employment decision but also for education decision, skilled workers need to compare home country wage with foreign country wage. To define the foreign country wage we assume that skilled workers supply more efficiency units of labor in the foreign country than in the home country. The rationale behind this assumption stems from skill-biased technological progress argument, which provides a framework to understand cross-country wage differences. As Caselli and Coleman (2006) argue, higher income countries use skilled labor more efficiently than lower income countries since they adopt technologies which favor skilled workers by increasing their productivities. Consistent with this argument, human capital of a skilled worker -born at time period $t-1$- of ability θ in the foreign country is:

$$h_t = (1 + \eta\theta e^{\gamma}), \tag{2}$$

where $\gamma > 0$, $\eta > 1$. Upon defining human capital formation in the home country and foreign country by (1) and (2) respectively, we describe how skilled wages are formed in the second and third period of a worker's lifetime.

In the second period of life, if the skilled worker of ability θ does not migrate, he receives a wage:

$$w_t^h = w(1 + \theta e^{\gamma}).$$

Considering the skilled migrant wages in the foreign country, we assume that the migrant workers' education levels can fully be observed by foreign employers while migrant workers' productivity levels can not. In particular, educational attainments are perfectly observed, however individual abilities can not. In that sense, foreign employers can distinguish between uneducated and educated migrant workers. Regarding the skilled migrant workers, (2), which describes the human capital formation in the foreign country, illustrates that the productivity levels of skilled migrant workers depend on innate ability levels, which are not perfectly observable. Thus, there is room for asymmetric information when the wage payments to skilled workers are considered and we elaborate on the skilled wage formation by foreign employers.

Following Chau and Stark (1999), let F_τ be the total number of migrants at time period τ and denote the cumulative number of migrants in the foreign country until time $t-1$ by $M_{t-1} = \sum_{\tau=0}^{t-1} F_\tau$. It is assumed that foreign employers discover the true productivity of a worker with probability $m_t = m(M_{t-1})$ and the following properties for the probability of discovery hold:

i) For each time period t, $m_t > 0$.

ii) $m_t^{'}(M_{t-1}) > 0$.

iii) $\lim_{M_{t-1}\to\infty} m(M_{t-1}) = \widehat{m} < 1$.

In line with migration literature, we assume that there are costs of living abroad to be incorporated into the model as a migration cost k_t which is assumed to decline as the number of permanent migrants increases. Let Z_τ be the flow of permanent migrants working in the foreign country at time period τ and assume that the following properties for k_t hold:

$i)$ For each time period t, $k_t > 0$.

$ii)$ $k_t^{'}(Z_{t-1}) > 0$.

$iii)$ $\lim_{Z_{t-1}\to\infty} k(Z_{t-1}) = k > \widehat{k}$.

It follows that the foreign wage[2], net of migration cost, of the skilled worker of ability θ in the second period of life at time period t when his productivity level is discovered is:

$$w_t^f = w(1 + \eta\theta e^\gamma) - k_t.$$

If the true productivity level of a skilled worker of ability θ is not discovered, foreign employers offer a wage payment which depends on the average productivity of the skilled migrant cohort with unknown ability levels at time period t and net wage of the worker is given by:

$$w_t^{fa} = w\frac{\int_{\theta^l}^{\theta^u}(1 + \eta\theta e^\gamma)f(\theta)d\theta}{F(\theta^u) - F(\theta^l)} - k_t = w(1 + \eta\theta^a e^\gamma) - k_t,$$

where θ^u and θ^l define the ability interval for skilled migrants, whose ability levels are not discovered, θ^a is the average ability level of the migrant population with unknown abilities.

Proceeding with the skilled worker wages in the third period of worker's life, they are formulated as:

If a skilled worker did not migrate in the second period of life, the wage in the home country in third period is:

$$w_{t+1}^h = w(1 + \theta e^\gamma).$$

[2]A foreign wage might be explicitly defined as an additional parameter as well. While only a skilled wage differential between home and foreign countries needed, one more parameter does not change the essence but increases algebraic complexities.

If a skilled worker, who migrated in the second period, decided to stay in the foreign country in the third period, he would receive w^f_{t+1}[3]

$$w^f_{t+1} = w(1 + \eta\theta e^{\gamma}) - k_t.$$

So as to introduce the possibility of return migration of skilled workers in their third period of life, we assume that human capital of a skilled migrant worker, who has worked in the foreign country for one period, has been augmented by learning new skills and techniques thus return-migrant receives foreign experience premium over the home country wage. To capture this idea, if a skilled worker, who migrated in the second period of his life, decided to work in the home country in the third period, he would receive w^r_{t+1}:

$$w^r_{t+1} = w(1 + \mu\theta e^{\gamma}),$$

where $1 < \mu < \eta$.

It is also assumed that only skilled workers migrate with the probability of p which reflects the emigration policies such as quotas, restrictions in the destination country.[4] Furthermore, in all models, we suppose that individuals choose whether to undertake education or not in their first period of life. Only in the second period, they decide whether to migrate or not. This restriction is placed in order to make all the models coherent since in the third period, workers decide to return home country or not, whenever return migration is a possibility.

2.1 An Economy without Migration

In this case, the individual's optimization problem is to decide whether to acquire education in the first period or not. To investigate which individuals acquire education, the discounted

[3] To enhance analytical tractability, we assume that an individual born at time period $t-1$, faces the same migration cost at time period t and $t+1$.

[4] Docquier and Marfouk (2004) document that skilled migration rates are three times higher than unskilled migration rates. This is also a common assumption in the skilled migration litarture.

lifetime utility of acquiring education and becoming a skilled worker should be compared with the discounted lifetime utility of becoming unskilled worker. The discounted lifetime utility of becoming a skilled worker for an individual of ability θ is:

$$y_t^s(\theta) = \beta(w(1 + \theta e^{\gamma}) - e) + \beta^2 w(1 + \theta e^{\gamma}).$$

If the individual does not acquire education and works as an unskilled worker, his discounted lifetime utility is expressed as:

$$y_t^u(\theta) = w(1 + \beta + \beta^2).$$

Hence, the individual optimally decides to acquire education if:

$$y_t^s(\theta) \geq y_t^u(\theta) \Leftrightarrow \beta(w(1 + \theta e^{\gamma}) - e) + \beta^2 w(1 + \theta e^{\gamma}) \geq w(1 + \beta + \beta^2)$$

or if and only if:

$$\theta \geq \frac{w + \beta e}{w\beta(1 + \beta)e^{\gamma}} = \theta^*. \tag{3}$$

Thus, the individuals of ability level greater than θ^* choose to undertake education and become skilled workers, while individuals of ability level lower than θ^* choose to work as unskilled workers. (3) also highlights that as the cost of education, which comprises of foregone earning w in the first period and the discounted direct cost βe, increases, the threshold ability level θ^* rises as well. Obviously, the increase in θ^* implies that the fraction of individuals of the young generation, who decides to acquire education, decreases.

Moreover, by defining θ^*, it is possible to observe how the population of $3N$ individuals in the home country is grouped at each time period t. Clearly, the number of unskilled workers is $3N(F(\theta^*))$ since the number of individuals, who do not acquire education, is $N(F(\theta^*))$ per generation. Since a fraction of $1 - F(\theta^*)$ of each generation gets educated, the number of skilled workers is $2N(1 - F(\theta^*))$ and the number of individuals pursuing education at time period t is $N(1- F(\theta^*))$.

The equilibrium is characterized in the economy at each period of time once the unique threshold ability level θ^* is identified. Not only the allocation of labor as unskilled and skilled labor is determined but also the output level and output per capita are figured out. Since the production in the economy evolves through a simple constant returns to scale production function, the output by unskilled and skilled workers at time period t are given respectively by $Y_t^u(\theta) = 3NwF(\theta^*)$ and $Y_t^s(\theta) = 2Nw\int_{\theta^*}^{\infty}(1+\theta e^{\gamma})f(\theta)d\theta$. Here, Y_t denotes the net value of total output in the economy.

For each time period t, the value of total output per capita, y_t,net of education expenditures, is computed as:

$$y_t(\theta^*) = wF(\theta^*) + \frac{2}{3}w\int_{\theta^*}^{\infty}(1+\theta e^{\gamma})f(\theta)d\theta - \frac{1}{3}e(1-F(\theta^*)) \tag{4}$$

To examine the relationship between the value of output per capita and the threshold ability level θ^*, we differentiate (4) with respect to θ^* :

$$\frac{\partial y_t(\theta^*)}{\theta^*} = wf(\theta^*) - \frac{2}{3}w(1+\theta^* e^{\gamma})f(\theta^*) + \frac{1}{3}ef(\theta^*) = -f(\theta^*)\left[-w + \frac{2}{3}w(1+\theta^* e^{\gamma}) - \frac{1}{3}e\right]$$

$$= -f(\theta^*)\left[\frac{2}{3}\frac{w+\beta e}{\beta(1+\beta)} - \frac{1}{3}(e+w)\right] = \frac{-f(\theta^*)}{3}\left[\frac{w(2-\beta-\beta^2)+\beta e(1-\beta)}{\beta(1+\beta)}\right] < 0. \tag{5}$$

since $0 < \beta < 1$.

By (5), it is inferred that the value of per capita output decreases as θ^* increases. Put differently, since an increase in θ^* implies a decrease in the fraction of skilled workers per generation, the economy experiences a decline in the value of per capita output as the number of skilled workers decreases.

2.2 Migration under Perfect Information

In this case, the individual's optimization problem is formulated as follows: In the first period, individual decides whether to acquire education and become a skilled worker or not, and in the second period, conditional upon being a skilled worker, the individual optimally chooses to migrate or not. As in models, which involve sequential decision-making, the model is solved backwards:

In the second period, a skilled worker has an incentive to migrate if the following condition holds:

$$(1+\beta)[w(1+\eta\theta e^{\gamma})-k_t] \geq (1+\beta)[w(1+\theta e^{\gamma})].$$

From this condition, the threshold ability level θ_t^{mig} for selecting migration is computed:

$$\theta \geq \frac{k_t}{w(\eta-1)e^{\gamma}} = \theta_t^{mig}.$$

In the first period, individuals decide whether to undertake education or not. For $\theta < \theta_t^{mig}$, individuals do not have any incentive to migrate, they acquire education and work in the home country as skilled workers if the following condition holds:

$$(\beta+\beta^2)[w(1+\theta e^{\gamma})] - \beta e \geq (1+\beta+\beta^2)w.$$

From this condition, the threshold ability level θ^{edu-h} is computed:

$$\theta \geq \frac{w+\beta e}{w\beta(1+\beta)e^{\gamma}} = \theta^{edu-h}.$$

For $\theta \geq \theta_t^{mig}$, individuals have incentive to migrate. Therefore, an individual of ability θ such that $\theta \geq \theta_t^{mig}$, acquires education if the following condition holds:

$$p(\beta+\beta^2)[w(1+\eta\theta e^{\gamma})-k_t] + (1-p)(\beta+\beta^2)[w(1+\theta e^{\gamma})] - \beta e \geq (1+\beta+\beta^2)w.$$

From this condition, the threshold ability level θ_t^{edu-f} is computed:

$$\theta \geq \frac{w + p\beta(1+\beta)k_t + \beta e}{w\beta(1+\beta)[p\eta + (1-p)]e^{\gamma}} = \theta_t^{edu-f}.$$

Assuming that migration cost k_1 is sufficiently high, the partitioning of the individuals is presented by Figure 1.

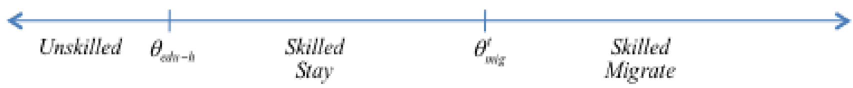

Figure 1: Migration under Perfect Information

Since there exist individuals, who choose to acquire education and work in the home country, for sufficiently high k_1, classical brain gain argument, which states that if individuals have the possibility to migrate, there is a decline in the threshold ability level, which determines the fraction of skilled individuals in the home country population, is not observed. However, since the flow of permanent migrants increase, i.e a decrease in θ_t^{mig}, at each period of time, the migration cost k_t decreases over time. Consequently, all educated individuals have incentive to migrate and the threshold ability level to acquire education is given by θ_t^{edu-f} and brain gain effect is observed[5]. The argument is formalized in the following proposition:

Proposition 1: *Assume that k_1 is sufficiently high and denote $k^* = \frac{(w+\beta e)(\eta-1)}{\beta(1+\beta)}$. Then the partitioning of the individuals in the home country is as follows:*

i) If $\widehat{k} > k^*$, then individuals of ability level $\theta < \theta^{edu-h}$ do not acquire education, of ability level $\theta^{edu-h} \leq \theta < \theta_t^{mig}$ acquire education and stay in the home country, of ability

[5] The threshold ability levels depend on migration cost. This observation motivates the reason that I extend the model of Chau and Stark (1999) by including non-constant migration cost.

level $\theta \geq \theta_t^{mig}$ *acquire education and migrate with probability* p.

ii) If $\widehat{k} \leq k^*$, *then individuals of ability level* $\theta < \theta_t^{edu-f}$ *do not acquire education, of ability level* $\theta \geq \theta_t^{edu-f}$ *acquire education and migrate with probability* p.

Proof See Appendix A.

2.3 Migration and Return Migration under Perfect Information

Taking the previous case one step further, we introduce return migration as a possibility for skilled migrant workers. In this case, the individual's optimization problem is formulated as follows: In the first period, individual decides to acquire education and become a skilled worker or not, in the second period, conditional upon being a skilled worker, the individual chooses to migrate or not. In the third period, a skilled migrant worker decides to return-migrate or not. Similar to the previous case, the model is solved backwards:

In the third period, a skilled migrant worker return-migrates if the following condition holds:

$$[w(1+\mu\theta e^{\gamma})] \geq [w(1+\eta\theta e^{\gamma}) - k_t].$$

From this condition, the threshold ability level θ_t^{ret}, which determines return-migrants among migrant population, is computed:

$$\theta \leq \frac{k_t}{w(\eta-\mu)e^{\gamma}} = \theta_t^{ret}.$$

For the second period decision-making, a skilled worker compares the discounted utility of migrating in the second period and return-migrating in the third period with staying in the home country for both periods:

$$p\{[w(1+\eta\theta e^{\gamma}) - k_t] + \beta[w(1+\mu\theta e^{\gamma})]\} + (1-p)(1+\beta)[w(1+\theta e^{\gamma}) \geq (1+\beta)[w(1+\theta e^{\gamma})].$$

From this condition, the threshold ability level θ_t^{mig-r} for choosing migration in the second period is found:

$$\theta \geq \frac{k_t}{w[(\eta - 1) + \beta(\mu - 1)]e^{\gamma}} = \theta_t^{mig-r}.$$

In the first period, individuals decide whether to acquire education or not. For $\theta < \theta_t^{mig-r}$, individuals do not have any incentive to migrate and they acquire education and work in the home country as skilled workers if the following condition holds:

$$(\beta + \beta^2)[w(1 + \theta e^{\gamma})] - \beta e \geq (1 + \beta + \beta^2)w.$$

From this condition, the threshold ability level θ^{edu-h} is computed:

$$\theta \geq \frac{w + \beta e}{w\beta(1 + \beta)e^{\gamma}} = \theta^{edu-h}.$$

For $\theta_t^{mig-r} \leq \theta < \theta_t^{ret}$, individuals have an incentive to migrate and if he/she migrated in the second period, he/she would return-migrate in third period. Therefore, an individual of ability θ such that $\theta_t^{mig-r} \leq \theta < \theta_t^{ret}$ acquires education if the following condition holds:

$$p\{[\beta w(1 + \eta\theta e^{\gamma}) - k_t] + \beta^2[w(1 + \mu\theta e^{\gamma})]\} + (1 - p)(\beta + \beta^2)[w(1 + \theta e^{\gamma})] - \beta e \geq (1 + \beta + \beta^2)w.$$

From this condition, the threshold ability level θ_t^{edu-r} is computed:

$$\theta \geq \frac{w + p\beta k_t + \beta e}{w[p(\beta\eta + \beta^2\mu) + (1 - p)(\beta + \beta^2)]e^{\gamma}} = \theta_t^{edu-r}.$$

For $\theta \geq \theta_t^{ret}$, individuals have an incentive to migrate and become permanent migrants. Therefore, an individual of ability θ such that $\theta \geq \theta_t^{ret}$ acquires education if the following condition holds:

$$p(\beta + \beta^2)[w(1 + \eta\theta e^{\gamma}) - k_t] + (1 - p)(\beta + \beta^2)[w(1 + \theta e^{\gamma})] - \beta e \geq (1 + \beta + \beta^2)w.$$

From this condition, the threshold ability level θ_t^{edu-f} is computed:

$$\theta \geq \frac{w + p\beta(1+\beta)k_t + \beta e}{w\beta(1+\beta)[p\eta + (1-p)]e^{\gamma}} = \theta_t^{edu-f}.$$

Defining the threshold ability levels and provided that k_1 is sufficiently high, we determine the partitioning of the home country individuals as illustrated by Figure 2. Individuals of ability level $\theta \leq \theta^{edu-h}$ do not acquire education and become unskilled workers, individuals of ability level $\theta^{edu-h} < \theta \leq \theta_t^{mig-r}$ acquire education and stay in the home country as skilled workers, individuals of ability level $\theta_t^{mig-r} \leq \theta < \theta_t^{ret}$ acquire education in the first period, migrate in the second period with probability p and if they migrated, they would return-migrate in the third period and individuals of ability level $\theta \geq \theta_t^{ret}$ acquire education and become permanent migrants with probability p.

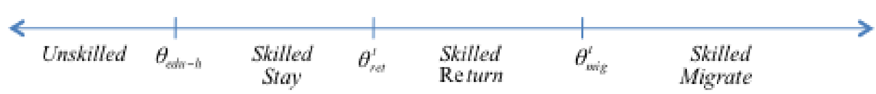

Figure 2: Migration and Return Migration under Perfect Information

Concerning the dynamics of migration when we allow for return migration in the third period, the behavior of migration cost k_t has to be considered. As k_t decreases over time, the partitioning of individuals in the home country changes and for well-defined values of $\widehat{k}$, staying in the home country is not optimal for any skilled worker and further, any skilled worker prefers to return-migrate in the third period. The behavior of k_t yields the following proposition:

Proposition 2: *Assume that k_1 is sufficiently high and denote $k^{**} = \frac{(w+\beta e)[(\eta-1)+\beta(\mu-1)]}{\beta(1+\beta)}$, $k^{***} = \frac{(w+\beta e)(\eta-\mu)}{[A-p\beta(\eta-\mu)]}$, where $A = [p\beta(\eta+\beta\mu) + (1-p)\beta(1+\beta)]$. Then the partitioning of the* *individuals in the home country is as follows:*

i) If $\widehat{k} > k^{**}$, *then individuals of ability level* $\theta < \theta^{edu-h}$ *do not acquire education, of ability level* $\theta^{edu-h} \leq \theta < \theta_t^{mig-r}$ *acquire education and stay in the home country, of ability level* $\theta_t^{mig-r} \leq \theta < \theta_t^{ret}$ *acquire education and become return migrants with probability* p, *of ability level* $\theta \geq \theta_t^{ret}$ *acquire education and become permanent migrants with probability* p.

ii) If $k^{**} \geq \widehat{k} > k^{***}$, *then individuals of ability level* $\theta < \theta_t^{edu-r}$ *do not acquire education, of ability level* $\theta_t^{edu-r} \leq \theta < \theta_t^{ret}$ *acquire education and become return migrants with probability* p, *of ability level* $\theta \geq \theta_t^{ret}$ *acquire education and become permanent migrants with probability* p.

iii) If $\widehat{k} \leq k^{***}$, *then then individuals of ability level* $\theta < \theta_t^{edu-f}$ *do not acquire education, of ability level* $\theta \geq \theta_t^{edu-f}$ *acquire education and become permanent migrants with probability* p.

Proof See Appendix A.

Defining the threshold ability levels for acquiring education in the economy without migration and in the migration models under perfect information, the comparisons of the threshold ability levels for acquiring education, yield the following corollary:

Corollary 1: *Denote the threshold education level in "migration under perfect information" by* θ_m *and "migration and return migration under perfect information" by* θ_r. *For any* $\widehat{k}$, $\theta^{edu-h} \geq \theta_m \geq \theta_r$.

Proof See Appendix A.

3 Migration and Return Migration Under Asymmetric Information

In this section, we first introduce possible scenarios and employment options of skilled workers under the presence of asymmetric information in the second period in a model with return migration. Then we study the main model, which includes asymmetric information along with a possibility of migration in the second period and return migration in the third period.

In the third period of a skilled worker's lifetime, there are apparently same employment options as in the benchmark perfect information case. If the worker migrated in the second period, he/she would have the possibility to return-migrate or stay in the foreign country. If he/she did not migrate in the second period, the worker would stay in the home country in the third period since migration in the third period is restricted by assumption.

In the second period, unlike the benchmark perfect information case, which offers only two employment options such as staying in the home country or migration, the skilled worker has more complex employment options due to asymmetric information. At the beginning of the second period, when whether a skilled worker has the possibility to migrate or not is not determined, some workers may not have incentive to migrate due to migration cost. Such a skilled worker of ability level θ has the expected income, net the cost of education, in the second period:

$$y_t^0(\theta) = w_t^h - e.$$

If a skilled worker had an incentive to migrate and managed to migrate, then the worker would have three more employment options:

1) With probability m_t, the true ability of the worker of ability θ is discovered and the worker return-migrates. With probability $1 - m_t$, the true ability of the worker is not discovered and he/she stays in the foreign country. For such an employment option, the

expected second period income, net the cost of education, of the worker is:

$$y_t^1(\theta) = m_t w_t^h + (1 - m_t) w_t^{fa} - e.$$

2)With probability m_t, the true ability of the worker of ability θ is discovered and the worker stays in the foreign country. With probability $1 - m_t$, the true ability of the worker is not discovered and he/she stays in the foreign country. For such an employment option, the expected second period income, net the cost of education, of the worker is:

$$y_t^2(\theta) = m_t w_t^f + (1 - m_t) w_t^{fa} - e.$$

3) With probability m_t, the true ability of the worker of ability θ is discovered and the worker stays in the foreign country. With probability $1 - m_t$, the true ability of the worker is not discovered and he/she return migrates. For such an employment option, the expected second period income, net the cost of education, of the worker is:

$$y_t^3(\theta) = m_t w_t^f + (1 - m_t) w_t^h - e.$$

Before proceeding to solution of the optimization problem of the skilled workers, we briefly discuss the second period employment options. Intuitively, workers, who prefer employment option 1, would like to benefit from asymmetric information in the foreign country since they do return-migrate once their true abilities are discovered. Hence, such workers are expected to be relatively low-skill workers. Moreover, workers choosing employment option 2 seem to be more skilled than the ones, who choose employment option 1, since they decide to work in the foreign country regardless of the asymmetric information leading to a wage payment depending on the average human capital endowment of the migrant cohort. Consequently, the workers, who prefer employment option 3, are the most-skilled workers in the home country simply because they return-migrate if their true abilities are not discovered by foreign employers.

Turning back to the first period, an individual chooses whether to get educated or not. If an individual did not acquire education, then he/she would become an unskilled worker and would not have the possibility to migrate. Therefore, an individual takes migration possibility into consideration in the first period.

As in sequential-decision making problems, individual's problem is solved backwards. In the third period, a skilled migrant worker of ability θ return-migrates if the following condition holds.

$$[w(1+\mu\theta e^{\gamma})] \geq [w(1+\eta\theta e^{\gamma}) - k_t]$$

From this condition, the threshold ability level θ_t^{ret} is obtained:

$$\theta \leq \frac{k_t}{w(\eta-\mu)e^{\gamma}} = \theta_t^{ret}.$$

Conditional upon migrating in the second period, workers of ability level $\theta \leq \theta_t^{ret}$ return-migrate, while workers of ability $\theta > \theta_t^{ret}$ become permanent migrants.

In the second period, skilled workers compare expected incomes from four employment options, which are defined above. When skilled workers decide on their second period employment options, they do not solely consider the second period expected incomes since their decisions affect their possible choices in the third period. For instance, in the second period, if a worker chooses employment option 3, he/she does not have the opportunity to return-migrate and earn the foreign experience premium μ in the case that his/her ability is not discovered and he/she works in the home country in the second period. Hence, in the second period, a skilled worker of ability θ, who has the opportunity to migrate, has to compare the following expected incomes[6] over two periods t and $t+1$:

$$y^{1j}_{t,t+1} = m_t(w^h_t + \beta w^h_{t+1}) + (1-m_t)[w^{fa}_t + \beta(1^{''}_R(\theta)w^r_{t+1} + (1-1^{''}_R(\theta))w^f_{t+1})] - e,$$

where $j = 1$ if $1^{''}_R(\theta) = 1$ and $j = 2$ if $1^{''}_R(\theta) = 0$.

$$\begin{aligned} y^{2j}_{t,t+1} &= m_t[(w^f_t + \beta(1^{'}_R(\theta)w^r_{t+1} + (1-1^{'}_R(\theta))w^f_{t+1})] \\ &\quad +(1-m_t)[w^{fa}_t + \beta(1^{''}_R(\theta)w^r_{t+1} + (1-1^{''}_R(\theta))w^f_{t+1})] - e, \end{aligned}$$

where $j = 1$ if $1^{'}_R(\theta) = 1$ and if $1^{''}_R(\theta) = 1$, $j = 2$ if $1^{'}_R(\theta) = 1$ and if $1^{''}_R(\theta) = 0$, $j = 3$ if $1^{'}_R(\theta) = 0$ and if $1^{''}_R(\theta) = 1$, $j = 4$ if $1^{'}_R(\theta) = 0$ and if $1^{''}_R(\theta) = 0$.

$$y^{3j}_{t,t+1} = m_t[(w^f_t + \beta(1^{'}_R(\theta)w^r_{t+1} + (1-1^{'}_R(\theta))w^f_{t+1})] + (1-m_t)(w^h_t + \beta w^h_{t+1}) - e,$$

where $j = 1$ if $1^{'}_R(\theta) = 1$ and $j = 2$ if $1^{'}_R(\theta) = 0$.

Considering all the employment options defined above, some employment options are not chosen by any skilled worker. The logic of this result is as follows. In the third period there is negative selection among skilled migrant workers since the returnee group in the third period is defined as the individuals of ability levels lower than θ^{ret}_t and the ones, who stay in the foreign country, are of ability level higher than θ^{ret}_t. Further, in the second period, skilled workers, who choose employment option 1, are expected to be relatively low-skilled while the ones, who select employment option 3, are expected to be the most skilled ones. Hence, the workers ,who pursue employment option 1 in the second period, are expected to be in the returnee group in the third period. The workers, who choose employment option 3 in the second period, are expected to stay in the foreign country in the third period. The following proposition explicitly states the decisions of the skilled workers:

[6] $1^{'}_R(\theta)$ and $1^{''}_R(\theta)$ are indicator functions which take value 0 if the worker return-migrates in the third period and 1 if the worker stays in the foreign country in the third period.

Proposition 3: *Assume that* $\eta(\theta_t^a - \theta_t^{ret}) > \theta_t^{ret}[(1-\mu)(1+\beta)]$ *holds. Skilled migrant workers pursue one of the following employment options:*

i) For the second period, if the true ability of the worker is discovered, then the worker chooses to return-migrate. If the true ability of the worker is not discovered, then the worker works in the foreign country. For the third period, conditional upon staying in the foreign country in the second period, the worker chooses to return-migrate. For such a worker, the expected net income is given by

$$y_{t,t+1}^{11}(\theta) = m_t(w_t^h + \beta w_{t+1}^h) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^r) - e.$$

ii) For the second period, the worker chooses to work in the foreign country regardless of the discovery of his true ability. For the third period, the worker chooses to return-migrate. For such a worker, the expected net income is given by:

$$y_{t,t+1}^{21}(\theta) = m_t(w_t^f + \beta w_{t+1}^r) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^r) - e.$$

iii) For the second period, the worker chooses to work in the foreign country regardless of the discovery of his true ability. For the third period, the worker chooses to work in the foreign country. For such a worker, the expected net income is given by:

$$y_{t,t+1}^{24}(\theta) = m_t(w_t^f + \beta w_{t+1}^f) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^f) - e.$$

iv) For the second period, if the true ability of the worker of ability θ *is discovered, then the worker chooses to work in the foreign country. If the true ability of the worker is not discovered, then the worker return-migrates. For the third period, conditional upon staying in the foreign country in the second period, the worker chooses to work in the foreign country. For such a worker, the expected net income is given by:*

$$y_{t,t+1}^{32}(\theta) = m_t(w_t^f + \beta w_{t+1}^f) + (1-m_t)(w_t^h + \beta w_{t+1}^h) - e.$$

Proof See Appendix A.

After characterizing the decisions of skilled workers who have the opportunity to migrate, we need to consider the possibility of existence of the workers who have incentive to stay in the home country and do not prefer migration. In the previous section, where we analyze migration models under perfect information, we show that for sufficiently large migration costs, there exist skilled workers who do not prefer migration over staying in the home country. In particular, those who stay in the home country are relatively low-skill workers in the whole home country population. However, under asymmetric information all skilled workers have incentive to migrate since low-skilled workers might receive a wage payment depending on the average human capital if their true abilities are not discovered. The following corollary formalizes this argument:

Corollary 2: *All skilled workers have incentive to migrate.*

Proof See the Appendix A

The comparison of θ^{edu-h} and θ_t^{edu} yields the following proposition:

Proposition 4: *The threshold ability level* θ_t^{edu} *is lower than* θ^{edu-h} *where*

$$\theta_t^{edu} = \frac{w - p\beta(1-m_t)(w\eta\theta_t^a - k_t) + \beta e}{w\beta[(1+\beta)(1-p+pm_t)+(1-m_t)p\beta\mu]e^{\gamma}}$$

Proof See the Appendix A

The rationale behind this proposition is as follows: Due to the asymmetric information on the part of foreign employers, any skilled worker chooses to stay in the home country. Recall that θ^{edu-h} is determined by considering the discounted lifetime utilities of working as an unskilled worker and staying in the home country in both periods. Hence, threshold ability level required to undertake education declines when the concept of asymmetric information is introduced to the model.

4 Dynamics of Migration

In this section, we study how the threshold ability levels $\theta_t^{edu}, \theta_t^{mig-r}, \theta_t^{ret}, \theta_t^{f}$, which are derived in the previous section, evolve over time as a result of the changes in the probability of discovery m_t and migration cost k_t. Given $\{M_0, Z_0\}$, migration and return migration decisions of workers are characterized by the vector $\left\{\theta_t^{edu}, \theta_t^{mig-r}, \theta_t^{ret}, \theta_t^{f}, \theta_t^{a}\right\}$ which consists of the solutions to the following system of equations:

$$\theta_t^{edu} = \frac{w - p\beta(1-m_t)(w\eta\theta_t^a - k_t) + \beta e}{w\beta[(1+\beta)(1-p+pm_t) + (1-m_t)p\beta\mu]e^{\gamma}}, \tag{6}$$

$$\theta_t^{mig-r} = \frac{k_t}{w((\eta-1) + \beta(\mu-1))e^{\gamma}}, \tag{7}$$

$$\theta_t^{ret} = \frac{k_t}{w(\eta-\mu)e^{\gamma}}, \tag{8}$$

$$\theta_t^{f} = \frac{w\eta\theta_t^a e^{\gamma} - (1+\beta)k_t}{w(1+\beta-\beta\eta)e^{\gamma}}, \tag{9}$$

$$\theta_t^{a} = \frac{\int_{\theta_t^{edu}}^{\theta_t^{f}} \theta f(\theta) d\theta}{F(\theta_t^{f}) - F(\theta_t^{edu})}. \tag{10}$$

Recalling the optimization problem of skilled workers presented in the previous section, the first four equations (6), (7), (8), (9) show that expected utility maximization of the workers on the set of employment options determine the extent of education, migration and return migration in the home country workforce. (10) is the definition of θ_t^a which is provided in Section 2.

4.1 The Effect of Changes in m_t and k_t on θ_t^a and θ_t^f

To analyze how θ_t^a and θ_t^f responds to changes in m_t and k_t, one has to consider (9) and (10) simultaneously since θ_t^f must satisfy both equations, which include θ_t^f as an argument, to be a solution to the system of equations. Considering (9), it can be rewritten as:

$$\theta_t^a = \frac{\theta_t^f w(1+\beta-\beta\eta)e^\gamma + (1+\beta)k_t}{w\eta e^\gamma}. \tag{11}$$

By simple algebraic manipulations, (11) is equivalent to the following equation:

$$w(1+\eta\theta_t^a e^\gamma) + \beta w(1+\eta\theta_t^f e^\gamma) = (1+\beta)[w(1+\theta_t^f e^\gamma) + k_t], \tag{12}$$

As Chau and Stark (1999) argue, (12) indicates that total wage payment to a permanent skilled worker, whose true ability is not discovered at time period t, must be sufficient to induce the supply of skilled workers of ability level $\theta \leq \theta_t^f$, who find it optimal to work in the foreign country at the total wage given on the left-hand side of (12). Therefore, one can characterize the supply side of the skilled migrant labor market by (11)[7].

Turning to the characterization of the demand side of the skilled migrant labor market, one can consider (10), which yields by rearranging:

$$w(1+\eta\theta_t^a e^\gamma) = w\frac{\int_{\theta_t^{edu}}^{\theta_t^f}(1+\eta\theta e^\gamma)f(\theta)d\theta}{F(\theta_t^f) - F(\theta_t^{edu})}. \tag{13}$$

To interpret (13), as Chau and Stark (1999) argue, one should observe that $1/w$ of the wage offer at time period t must be equal to the average human capital of the migrant workforce, with unknown abilities at time period t. As a consequence, this equation can be considered as an equation describing the demand side of the migrant labor market.

Before following a concise and formal treatment for investigating the effects of m_t and k_t on threshold ability levels, we provide a graphical analysis to grasp the behavior of

[7] Alternatively, one can characterize the supply side of the migrant labor market by (6).

threshold ability levels over time in a more intuitive manner. Plotting the supply and demand relationship by (11) and (13), and denoting the corresponding curves by SS_f and DD_f, Figure 5 and 6 depict the relationship between θ_t^a and θ_t^f. It is confirmed that both curves representing the equations are upward sloping.

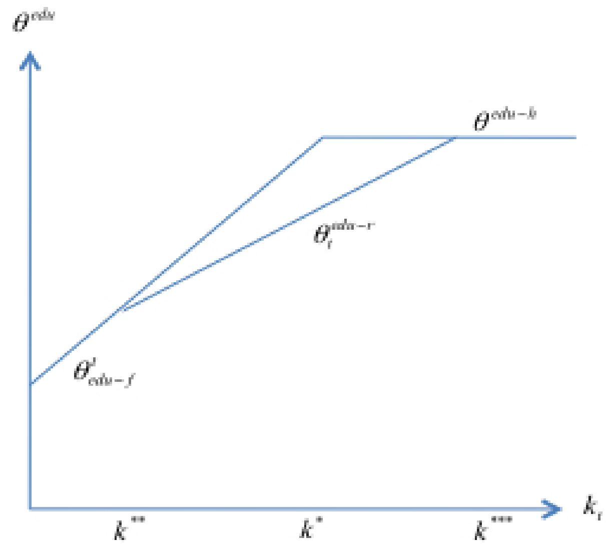

Figure 3: Migration and Return Migration under Perfect Information

The upward slope of the SS_f curve asserts that high values of θ_t^a are associated with high values of θ_t^f. As it is captured by (11), higher values for θ_t^a allow more skilled workers to migrate and stay in the foreign country by raising ability level θ_t^f and thus enlarging the fraction of skilled migrant workers in the population. The slope of the SS_f is:

$$\frac{\partial \theta_t^a}{\partial \theta_t^f}|_{SS_f} = \frac{w(1+\beta-\beta\eta)e^\gamma}{w\eta e^\gamma} = \frac{1+\beta-\beta\eta}{\eta} > 0. \tag{14}$$

Furthermore, the positively sloped DD_f curve seems to be an expected result since an increase in the upper bound θ_t^f of the integral should lead to an increase in θ_t^a. However, there is another channel which determines the positive slope of the DD_f curve. To gain

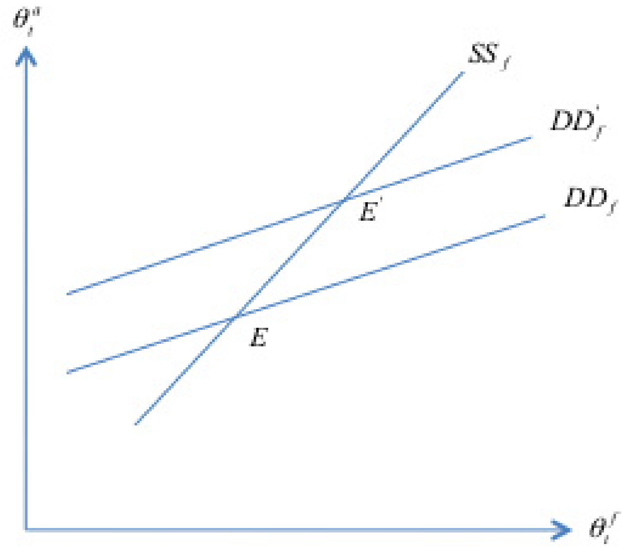

Figure 4: An increase in m_t when SS_f is steeper

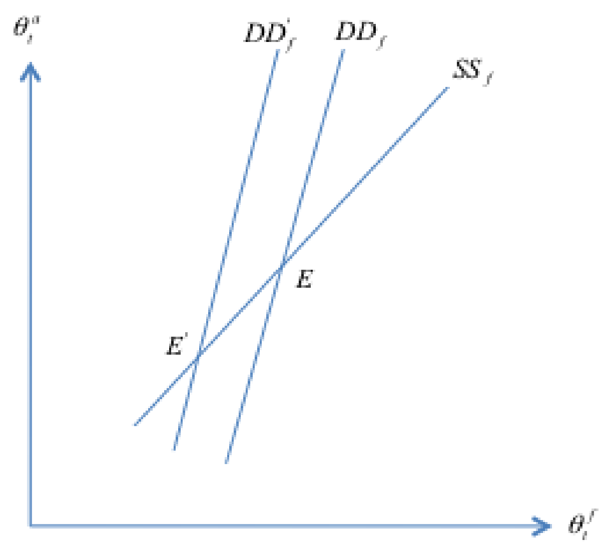

Figure 5: An increase in m_t when DD_f is steeper

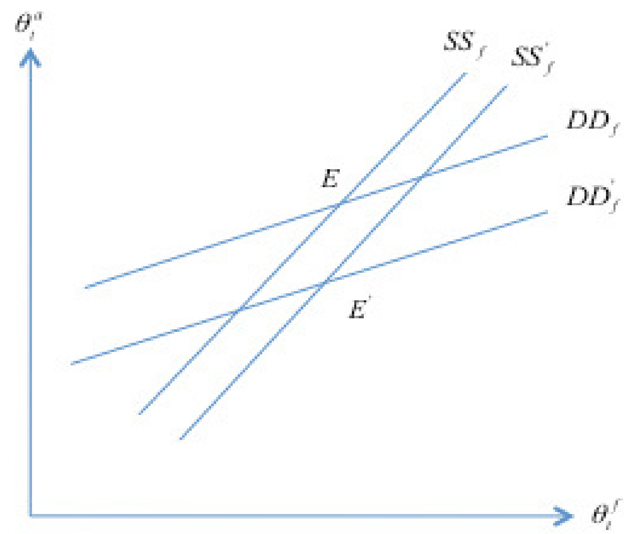

Figure 6: A decrease in k_t when SS_f is steeper

intuition about that channel, consider:

$$\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD_f} = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} + \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]}\frac{\partial\theta_t^{edu}}{\partial\theta_t^a}\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD_f}. \quad (15)$$

(15) states that since $\partial\theta_t^{edu}/\partial\theta_t^a < 0$ by (6), which implies that lower bound for migrant ability distribution responds negatively to an increase in θ_t^a, higher values of θ_t^f should match with higher values of θ_t^a so as to offset the negative effect through θ_t^{edu}. The slope of the DD_f curve is:

$$\frac{\partial\theta_t^a}{\partial\theta_t^f}|_{DD_f} = \frac{1}{\Psi}\frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} > 0, \quad (16)$$

where $\Psi = 1 + \frac{(\theta_t^{edu} - \theta_t^a)f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]}\frac{\partial\theta_t^{edu}}{\partial\theta_t^a} > 0$. Note also that, since the slopes of SS_f and DD_f curves depend on the exogenous parameters of the model, SS_f curve can be steeper or flatter than DD_f curve. Given m_t and k_t, the equilibrium pairs $\left\{\theta_t^f, \theta_t^a\right\}$, which simultaneously satisfy (11) and (13), are determined by the intersection points E in Figure 5 and 6.

Suppose that there is an increase in the probability of discovery m_t. An increase in m_t shifts the DD_f curve upward while the SS_f curve remains unchanged. This result is obtained by differentiating the SS_f and DD_f with respect to m_t while keeping θ_t^f constant. The differentiation of SS_f and DD_f with respect to m_t yields respectively:

$$\frac{\partial \theta_t^a}{\partial m_t}|_{\theta_t^f \text{ constant}} = 0.$$

$$\frac{\partial \theta_t^a}{\partial m_t}|_{\theta_t^f \text{ constant}} = \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})} \frac{\partial \theta_t^{edu}}{\partial m_t} > 0.$$

If DD_f is flatter than SS_f curve as in Figure 5, the new equilibrium pair of θ_t^f and θ_t^a is denoted by the point E', which indicates that there is not only an increase in the average ability of migrants but also a rise in θ_t^f due to the increase in m_t. On the other hand, starting form a point such as E in Figure 6, where SS_f is flatter than DD_f curve, an increase in m_t implies a reduction in both θ_t^a and θ_t^f, as depicted by point E'. The logic of this result is explained by the following transmission mechanism. An increase in m_t invokes a negative incentive for low-ability workers to migrate since $\partial \theta_t^{edu} / \partial m_t > 0$. At the same time, an increase in m_t leads to an upward shift of DD_f curve, which implies an increase in θ_t^a for any given θ_t^f, thus a positive incentive for low-ability workers is spotted. If SS_f curve is steeper than DD_f curve, then an increase in θ_t^a induces a higher θ_t^f and offsets the negative effect of low ability workers on θ_t^a and therefore, θ_t^f and θ_t^a increase as a result of the increase in m_t. In contrast, If SS_f curve is flatter than DD_f curve, then the negative effect of low-ability workers dominate and the new equilibrium pair $\left\{\theta_t^f, \theta_t^a\right\}$ is lower.

Suppose that there is a decrease in the migration cost k_t, when SS_f is steeper than DD_f as in Figure 7. A decrease in k_t leads to a downward shift of the SS_f and the DD_f curves. This result is obtained by differentiating the SS_f and DD_f curves with respect to k_t while keeping θ_t^f constant. The differentiation of SS_f and DD_f with respect to k_t yields respectively:

$$\frac{\partial \theta_t^a}{\partial k_t}|_{\theta_t^f \text{ constant}} = \frac{1+\beta}{w\eta e^\gamma} > 0.$$

$$\frac{\partial\theta_t^a}{\partial k_t}|_{\theta_t^f \text{ constant}} = \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]}\frac{\partial\theta_t^{edu}}{\partial k_t} > 0.$$

Concerning the supply side of the skilled migrant market, as k_t decreases, for any given level of θ_t^f, a decline in θ_t^a takes place. Hence, a downward shift of the SS_f curve is observed. Besides, the effect of a decline in migration cost on the demand side of the migrant labor market operates in the same direction since a decrease in k_t also results in a positive for low-ability workers. Hence, the magnitude of the shifts of the SS_f and DD_f curves determine the new equilibrium pair $\left\{\theta_t^f, \theta_t^a\right\}$. If the magnitude of the former is sufficiently high, then the new equilibrium at a higher $\left\{\theta_t^f, \theta_t^a\right\}$ pair, otherwise the resulting θ_t^f and θ_t^a are lower. Moreover, one should note that there is also a possibility that a higher θ_t^f might match a lower θ_t^a in the new equilibrium when a decrease in k_t is experienced. The following lemma formalizes the discussion about the effects of changes in m_t and k_t on θ_t^a and θ_t^f :

Lemma 1: i) $\theta_t^a(m_t, k_t)$ is increasing in m_t if and only if

$$1 - \frac{A\eta}{(1+\beta-\beta\eta)} + \frac{B[p\beta(1-m_t)w\eta e^\gamma}{C} > 0,$$

where $A = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]}$, $B = \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]}$, $C = w\beta[(1+\beta)(1-p+pm_t)+(1-m_t)p\beta\mu]e^\gamma$.

ii) $\theta_t^f(m_t, k_t)$ is increasing in m_t if and only if $\theta_t^a(m_t, k_t)$ is increasing in m_t.

iii) If $\theta_t^a(m_t, k_t)$ is increasing in m_t, then $\theta_t^a(m_t, k_t)$ is decreasing in k_t if and only if:

$$-A\frac{(1+\beta)}{w(1+\beta-\beta\eta)e^\gamma} + B\frac{p\beta(1-m_t)}{C} < 0.$$

iv) $\theta_t^f(m_t, k_t)$ is decreasing in k_t if and only if

$$1 + \beta > w\eta e^\gamma \left(\frac{\partial\theta_t^a}{\partial k_t}\right).$$

Proof See Appendix B.

4.2 The Effect of Changes in m_t and k_t on θ_t^{mig-r} and θ_t^{ret}

Since θ_t^{mig-r} does not depend on θ_t^a, it is sufficient to consider the impact of changes in m_t and k_t solely on θ_t^{mig-r} without taking the effect on θ_t^a into account. Thus, it is enough to analyze the derivative of θ_t^{mig-r} with respect to m_t and k_t. Further, we only conduct the analysis of a change in k_t on θ_t^{mig-r} since θ_t^{mig-r} is not a function of m_t. Similarly, θ_t^{ret} also does not depend on θ_t^a and m_t and only the effect of a change in k_t is observed. Formally, the derivatives of θ_t^{mig-r} and θ_t^{ret} are respectively given as:

$$\frac{\partial \theta_t^{mig-r}}{\partial k_t} = \frac{1}{w((\eta - 1) + \beta(\mu - 1))e^{\gamma}} > 0,$$

$$\frac{\partial \theta_t^{ret}}{\partial k_t} = \frac{1}{w(\eta - \mu)e^{\gamma}} > 0.$$

Thus, positive derivatives of θ_t^{mig-r} and θ_t^{ret} with respect to k_t imply that if there is a decline in k_t, θ_t^{mig-r} and θ_t^{ret} decrease as well. The intuition behind this result is very clear in the sense that a decline in the migration cost lets more workers to migrate in the second period and less workers to return-migrate in the third period as it is shown by a reduction in the required ability levels. Hence, the following lemma without proof formalizes this argument.

Lemma 2: *θ_t^{mig-r} and θ_t^{ret} are increasing in k_t.*

4.3 The Effect of Changes in m_t and k_t on θ_t^a and θ_t^{edu}

In this section, we study how θ_t^a and θ_t^{edu} adjust to changes in m_t and k_t. Similar to the analysis conducted for θ_t^a and θ_t^f, one can characterize the supply and demand sides of the migrant labor market. While the latter is again represented by (13), the former is obtained by rewriting (6) as:

$$\theta_t^a = \frac{w + p\beta(1 - m_t)k_t - \theta_t^{edu}C + e}{p\beta(1 - m_t)w\eta}. \tag{17}$$

Denoting the corresponding curves to (17) and (13) by SS_e and DD_e, the slope of the SS_e curve is negative since the higher θ_t^a, the higher the number of low-skill workers benefiting from high θ_t^a. The slope of the SS_e curve is:

$$\frac{\partial \theta_t^a}{\partial \theta_t^{edu}}|_{SS_e} = \frac{-C}{p\beta(1-m_t)w\eta} < 0. \tag{18}$$

Regarding the slope of the DD_e curve, first consider the following:

$$\frac{\partial \theta_t^a}{\partial \theta_t^{edu}}|_{DD_e} = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[(F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\partial \theta_t^f}{\partial \theta_t^a} \frac{\partial \theta_t^a}{\partial \theta_t^{edu}}|_{DD_e} + \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[(F(\theta_t^f) - F(\theta_t^{edu})]} \tag{19}$$

Observing (19), it is deduced that it is possible to come up with a negatively sloped DD_e curve despite the fact that θ_t^a is strictly increasing in θ_t^{edu}. Put differently, high values for θ_t^a might be associated with low values for θ_t^{edu} since $\partial\theta_t^f/\partial\theta_t^a > 0$. Formally, the slope of the DD_e curve is:

$$\frac{\partial \theta_t^a}{\partial \theta_t^{edu}}|_{DD_e} = \frac{1}{\Sigma} \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]}, \tag{20}$$

where $\Sigma = 1 - \frac{(\theta_t^{edu} - \theta_t^a)f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})} \frac{\partial \theta_t^f}{\partial \theta_t^a}$. To perform a similar analysis which is discussed to explore the effect of the change in m_t on θ_t^f, one needs to consider the response of θ_t^{edu} by differentiating the SS_e and DD_e curves with respect to m_t. The differentiation of the SS_e and DD_e yields respectively:

$$\frac{\partial \theta_t^a}{\partial m_t}|_{\theta_t^{edu} \text{ constant}} = \frac{p\beta w\eta[w + p\beta(1-m_t)k_t - \theta_t^{edu}C + e}{[p\beta(1-m_t)w\eta]^2} > 0,$$

$$\frac{\partial \theta_t^a}{\partial m_t}|_{\theta_t^{edu} \text{ constant}} = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\partial \theta_t^f}{\partial m_t} = 0.$$

The above expressions illustrate that an upward shift of the SS_f is observed along with an unchanged DD_f curve as a result of an increase in m_t. As it is represented by Figure 8, if the DD_f curve is positively sloped, then the new equilibrium pair $\{\theta_t^{edu}, \theta_t^a\}$ is attained at a higher value. However, as opposed to the co-movement of θ_t^f and θ_t^a as response to a change in m_t, Figure 9 and 10 show that θ_t^{edu} and θ_t^a might move in opposite directions depending on the slope of the DD_f curve.

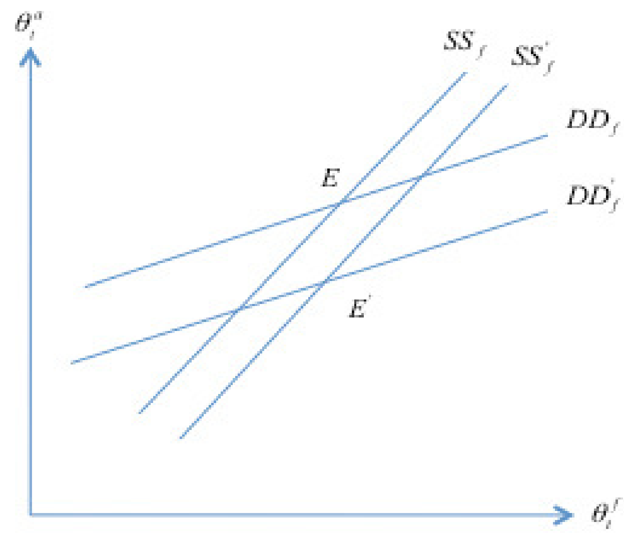

Figure 7: An increase in m_t when SS_e is positively sloped

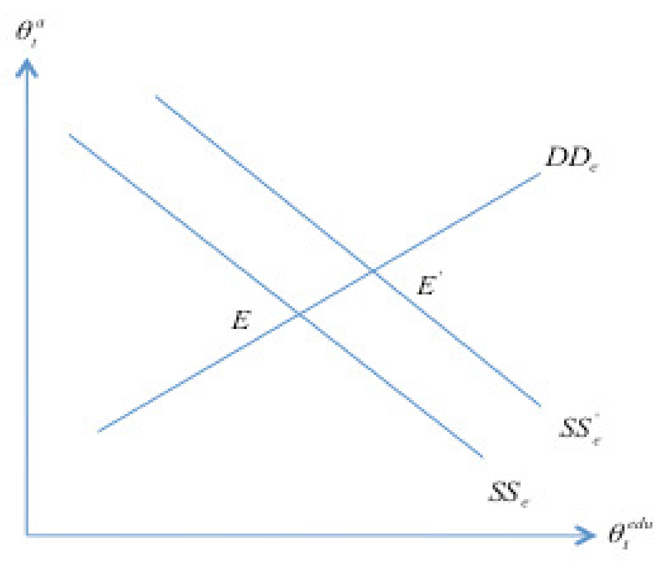

Figure 8: An increase in m_t when SS_e is negatively sloped

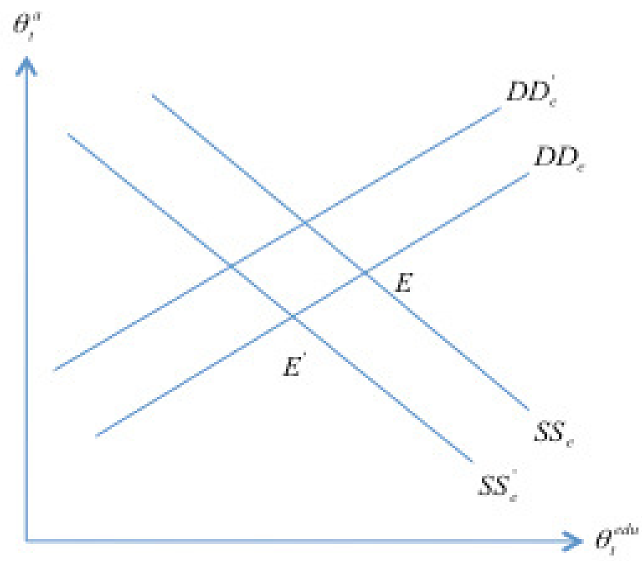

Figure 9: An increase in m_t when SS_e is negatively sloped

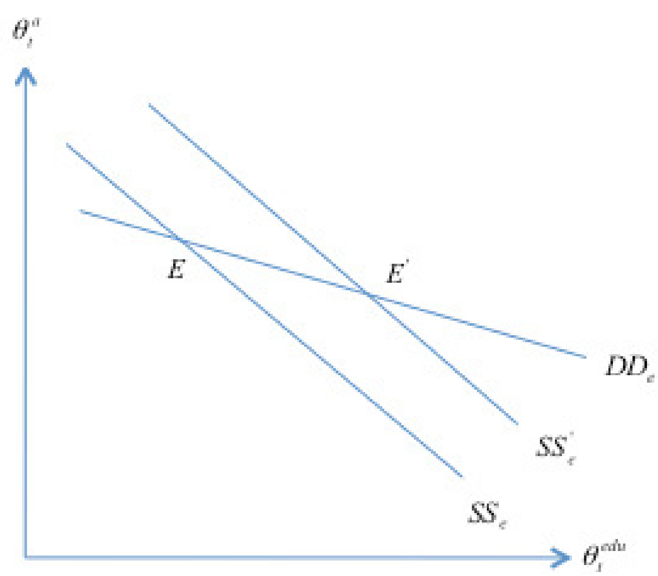

Figure 10: A decrease in k_t when SS_e is negatively sloped

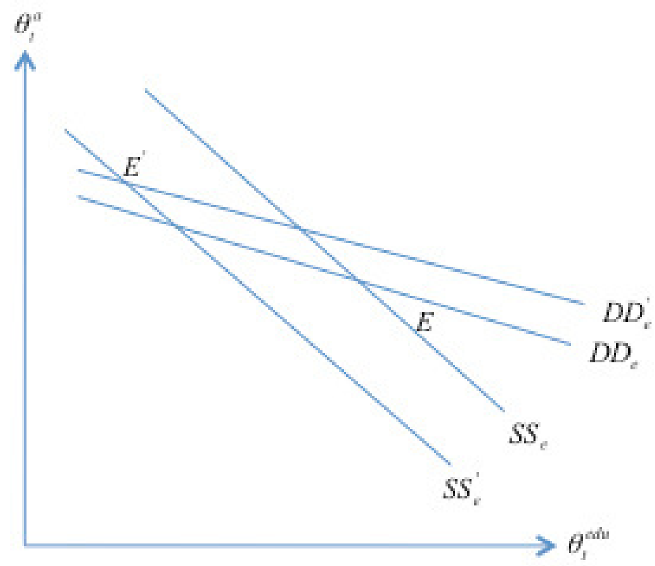

Figure 11: A decrease in k_t when SS_e is negatively sloped

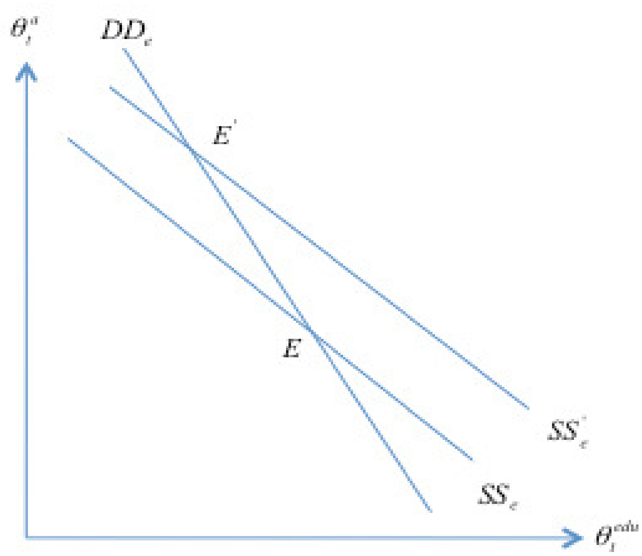

Figure 12: A decrease in k_t when SS_e is negatively sloped

Regarding the effect of a decrease in k_t, a downward shift of the SS_f curve and an upward shift of DD_f are observed since :

$$\frac{\partial \theta_t^a}{\partial k_t}|_{\theta_t^{edu} \text{ constant}} = \frac{1}{w\eta} > 0,$$

$$\frac{\partial \theta_t^a}{\partial k_t}|_{\theta_t^{edu} \text{ constant}} = \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\partial \theta_t^f}{\partial k_t} < 0.$$

Similar to the previous analysis, the effect of a decrease in k_t might lead to different equilibrium pairs $\left\{\theta_t^{edu}, \theta_t^a\right\}$ depending on the slope of the DD_f curve. For instance, as it can be seen by Figure 11 that if the DD_f curve has a positive slope, then we end up with a negative change in θ_t^{edu} along with an ambiguous effect on θ_t^a. Similar conclusions can be drawn upon examining Figure 12 and 13 which graph DD_f curve as negatively sloped.

The following lemma formalizes the analysis on the effects of changes in m_t and k_t on θ_t^{edu} :

Lemma 3: *i)* $\theta_t^{edu}(m_t, k_t)$ *is increasing in* m_t *if and only if*

$$\left[\frac{p\beta(w\eta\theta_t^a e^\gamma - k_t) - p(1+\beta-\beta\mu)w\theta_t^{edu} e^\gamma}{p\beta(1-m_t)w\eta e^\gamma}\right] > \left(\frac{\partial \theta_t^a}{\partial m_t}\right).$$

ii) $\theta_t^{edu}(m_t, k_t)$ *is increasing in* k_t *if and only if*

$$1 > w\eta e^\gamma \left(\frac{\partial \theta_t^a}{\partial k_t}\right).$$

Proof See Appendix B.

Upon analyzing the intertemporal variations of the threshold ability levels, we follow with the discussion on whether there exist equilibrium values for m_t and k_t which govern the process of migration and return migration over time. Since m_t and k_t depend on the cumulative number of migrants and flow of permanent migrants respectively, m_{t+1} and k_{t+1}

are defined as follows.

$$m_{t+1} = m_t[m^{-1}(m_t) + F_t], \text{ if } m_t < \widehat{m}\widehat{m}, \qquad \text{otherwise,} \tag{21}$$

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

$$k_{t+1} = k_t(Z_t), \text{ if } k_t > \widehat{k}\widehat{k} \quad \text{otherwise,} \tag{22}$$

where $k^{-1}(k_1) = Z_0$ is given.

Denote the equilibrium values for (21) and (22) by m^* and k^* such that $m_t = m_{t+1} = m^*$ and $k_t = k_{t+1} = k^*$. The equilibrium values of θ_t^j are denoted by $\widehat{\theta_t^j}$, for $j = edu, mig-r, ret, f, a$. Once m^* and k^* are evaluated, one can easily compute the equilibrium values of $\widehat{\theta_t^j}$ by (6), (7), (8), (9) and (10). Further, denote the effect of m_t on θ_t^a by Σ_m and the effect of k_t on θ_t^a by Σ_k.

Proposition 5: *If* $\Sigma_m dm_t + \Sigma_k dk_t > 0$ *and* m_1 *and* k_1 *are such that* $\eta(\theta_1^a - \theta_1^{ret}) > \theta_1^{ret}[(1-\mu)(1+\beta)]$ *holds, then the only equilibrium values for* m_t *and* k_t *are* $\widehat{m}$ *and* $\widehat{k}$.

Proof See Appendix B.

This proposition states that even if the decline in k_t reduces θ_t^a, it is still possible to observe an increase in θ_t^a if the effect of increasing m_t offsets the negative effect by k_t[8]. Also, if the condition $\Sigma_m dm_t + \Sigma_k dk_t > 0$ is satisfied, then θ_t^f rises as well. This result indicates that return migration in the upper tail of the ability distribution becomes more positively selected over time. Moreover, since there is no restriction on the behavior of θ_t^{edu}, the fraction of skilled workers in the home country population might rise during the process of migration. Finally, since θ_t^{mig-r} and θ_t^{ret} show a downward trend, return migration both

[8]Note that I rule out the possibilities that if $\Sigma_k > 0$, k_t reaches $\widehat{k}$ after m_t reaches $\widehat{m}$ and if $\Sigma_m < 0, m_t$ reaches $\widehat{m}$ after k_t reaches $\widehat{k}$.

in the second period and third period demonstrates more negative selection over time.

5 Welfare Analysis

In this three-period setting,at any time period t, the $3N$ individuals are distributed as follows: From each generation t, $t-1$ and $t-2$, there are $N(F(\widehat{\theta^{edu}})$ unskilled workers. From the generation t, $N(1-F(\widehat{\theta^{edu}}))$ individuals pursue education. From the generation $t-1$, $p(1-\widehat{m})N[F(\widehat{\theta^{mig-r}})-F(\widehat{\theta^{edu}})]+pN[F(\widehat{\theta^{f}})-F(\widehat{\theta^{mig-r}})]+p\widehat{m}N[(1-F(\widehat{\theta^{f}})]=M_{-1}$ workers stay in the foreign country. From the generation $t-2$, $pN[F(\widehat{\theta^{ret}})-F(\widehat{\theta^{mig-r}})]+p\widehat{m}N[(1-F(\widehat{\theta^{f}})]$ workers work in the foreign country while $p(1-\widehat{m})N[F(\widehat{\theta^{mig-r}})-F(\widehat{\theta^{edu}})]+pN[F(\widehat{\theta^{ret}})-F(\widehat{\theta^{mig-r}})]$ workers return-migrated along with an augmentation in their human capital by μ.

The equilibrium value of per-period national output, net of education expenditures, is:

$$\begin{aligned} Y_t(\widehat{\theta}) &= 3N(F(\widehat{\theta^{edu}})w + N(1-p)\int_{\widehat{\theta^{edu}}}^{\infty}(w(1+\theta e^{\gamma})-e)f(\theta)d\theta \\ &+N(1-p)\int_{\widehat{\theta^{edu}}}^{\infty}w(1+\theta e^{\gamma})f(\theta)d\theta + Np\widehat{m}\int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}(w(1+\theta e^{\gamma})-e)f(\theta)d\theta \\ &+Np(1-\widehat{m})\int_{\widehat{\theta^{f}}}^{\infty}(w(1+\theta e^{\gamma})-e)f(\theta)d\theta \\ &+Np(1-\widehat{m})\int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}w(1+\mu\theta e^{\gamma})f(\theta\theta)d\theta + Np\int_{\widehat{\theta^{mig-r}}}^{\widehat{\theta^{ret}}}w(1+\mu\theta e^{\gamma})f(\theta)d\theta. \end{aligned} \tag{23}$$

Rewriting (23) yields the following:

$$\begin{aligned} Y_t(\widehat{\theta}) &= 3N(F(\widehat{\theta^{edu}})w + N\left[\int_{\widehat{\theta^{edu}}}^{\infty}(w(1+\theta e^{\gamma})-e)f(\theta)d\theta - (w(1+\theta_{-1}e^{\gamma})-e)\frac{M_{-1}}{N}\right] \\ &+N\left[\int_{\widehat{\theta^{edu}}}^{\infty}w(1+\theta e^{\gamma})f(\theta)d\theta - (w(1+\theta_{-2}e^{\gamma})\frac{M_{-2}}{N}\right]+ \\ &N\left[p(1-\widehat{m})\int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}(w(\mu-1)\theta e^{\gamma})f(\theta)d\theta + Np\int_{\widehat{\theta^{mig-r}}}^{\widehat{\theta^{ret}}}(w(\mu-1)\theta e^{\gamma})f(\theta)d\theta\right], \end{aligned}$$

$$\text{where } \theta_{-1} = \frac{pN}{M_{-1}}\left[(1-\widehat{m})\int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}\theta f(\theta)d\theta + \int_{\widehat{\theta^{mig-r}}}^{\widehat{\theta^{f}}}\theta f(\theta)d\theta + \widehat{m}\int_{\widehat{\theta^{f}}}^{\infty}\theta f(\theta)d\theta\right],$$

$$\theta_{-2} = \frac{pN}{M_{-2}}\left[\int_{\widehat{\theta^{ret}}}^{\widehat{\theta^{f}}}\theta f(\theta)d\theta + \widehat{m}\int_{\widehat{\theta^{f}}}^{\infty}\theta f(\theta)d\theta\right]$$

Defining the equilibrium per capita output $\frac{Y_t(\widehat{\theta})}{3N-M_{-1}-M_{-2}} = y_t(\widehat{\theta})$, the comparison of $y_t(\widehat{\theta})$ and $y_t(\theta^*)$ yields the following result.

Denote the gain from return migration in the third period by Π[9]. Then, $y_t(\widehat{\theta}) > y_t(\theta^*)$ if and only if:

$$\frac{1}{3N-M_{-1}-M_{-2}}\left\{\begin{array}{c} \int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}[(w(1+\theta e^{\gamma})-e)-w]f(\theta)d\theta \\ +\int_{\widehat{\theta^{edu}}}^{\widehat{\theta^{mig-r}}}[(w(1+\theta e^{\gamma})-w]f(\theta)d\theta \\ +\frac{M_{-1}}{N}[(y_t(\theta^*)-(w(1+\theta_{-1}e^{\gamma})-e)] \\ +\frac{M_{-2}}{N}[(y_t(\theta^*)-(w(1+\theta_{-2}e^{\gamma})]+\Pi \end{array}\right\} > 0.$$

By the above expression, the first two strictly positive terms in the integral represent the gain from reduction in the threshold ability level for acquiring education due to the possibility of migration. The third and fourth terms refer to the possible change in output per capita due to loss of skilled workforce. Hence, if the positive incentive effect reflected by the first two terms and gain by return migration represented by Π offset the human capital depletion as a result of migration, then the home country experiences a welfare gain by opening up to migration.

6 Conclusion

In this paper, we constructed a dynamic migration model which embodies endogenous skill acquisition, return migration and non-constant migration cost. Allowing heterogeneous abil-

[9] $\Pi = Np(1-\widehat{m})\int_{\widehat{\theta}^{edu}}^{\widehat{\theta}^{mig-r}} w(\mu-1)\theta e^{\gamma})f(\theta)d\theta + Np\int_{\widehat{\theta}^{mig-r}}^{\widehat{\theta}^{ret}} w(\mu-1)\theta e^{\gamma})f(\theta)d\theta$

ity levels for individuals, we analyzed the model under different settings and information structures. We first characterized the extent of skill acquisition in the source country and determined the level of output in a closed economy. Then, migration and employment opportunity in the destination country have been introduced to the model and individual behavior concerning skill acquisition and migration decisions under perfect information have been explored. We have shown that when an economy opens up to migration, skill acquisition in the source country increases as migration cost decreases over time. We further incorporated return migration into the model under perfect information and we have found that skill acquisition in the source country is higher compared to the model that does not involve return migration. This result indicates that return migration might increase further the average level of human capital in the source country. Then, we studied the model under asymmetric information on the part of employers and found that regardless of the level of migration cost, skill acquisition in this setting is higher than closed economy setting. Finally, we examined the process of migration over time and conducted a welfare analysis, which concluded that if skill acquisition effect due to migration possibility and human capital gain from return migration were sufficiently high, then the source country would experience a welfare gain.

The analysis above demonstrates that migration of skilled workers might lead to a rise in the average level of human capital. Moreover, return migration along with human capital augmentation increases the probability of an increase in the average level of human capital and welfare in the source country. Hence, our model model supports the Brain Gain with Brain Drain argument.

Since the model takes a source-country perspective by focusing on the effects of migration on the skill acquisition and welfare of the source country, migration probability of unskilled workers is assumed to be zero whereas the migration probability of skilled workers is strictly positive and exogenously given. This simplifying assumption contradict with some stylized facts and causes a counterfactual migration pattern in the sense that the migration of unskilled workers can not be observed by the model. However, the literature on labor migration stresses that the illegal migration shows an upward trend and mostly involves unskilled labor. Hence, the model does not account for unskilled labor migration induced by illegal migration. Furthermore, migration probability of skilled workers might be endogeneized and it might negatively depend on the number of migrants since the natives lose their jobs and unemployment among natives might increase with migration. Consequently, natives might put pressure on the immigration authorities for stricter immigration policies.

Our model can be extended in two directions. Firstly, if an appropriate functional form for the probability of discovery and migration cost can be found, then simulation with data can be conducted to explore how the welfare of the source country changes as migration probability varies. Secondly, intergenerational externality of human capital might be introduced, and growth effects of migration and return migration can be studied under various informational structures

References

Barrett, A. and O'Connell P.J., (2001), "Is there a wage premium for returning Irish migrants?" Economic and Social Review, 32(1): 1-21.

Beine, M., Docquier, F., Rapoport, H., (2001), " Brain Drain and Economic Growth,"Journal of Development Economics 64(1): 275-289.

Benhabib, J. and B. Jovanovic (2012), "Optimal migration: A world perspective," International Economic Review 53, 321-348.

Borjas, G.J. (1987), 'Self-selection and the earnings of immigrants, American Economic Review, 77, 531-53.

Carrington,W., Detragiache, E.,Vishwanath, T., (1996), "Migration with Endogenous Moving Costs." American Economic Review 86 (3): 909-930.

Caselli, F. and Coleman, W.J., (2006), "The World Technology Frontier." American Economic Review 96 (3): 499-522.

Chau, N., and O. Stark (1999), "Migration under Asymmetric Information and Human Capital Formation" Review of International Economics, 7(3): 455-83.

Cortes, P. (2008), "The effect of low-skilled immigration on US prices: Evidence from CPI data," Journal of Political Economy 116, 381-422.

Mattoo,A., I.C.Neagu, and C.Ozden(2008),"Brainwaste? Educated immigrants in the US labor market," Journal of Development Economics 87, 255-269.

Docquier, F. and Marfouk, A., (2004), "Measuring the International Mobility of Skilled Workers,"World Bank Policy Research Working Paper No. 3381.

Dos Santos, M. and Postel Vinay F., (2003), 'Migration as a source of growth: the perspcetive of a developing country." , Journal of Population Economics 16: 161-175.

Hemmi, N., (2005), "Brain drain and economic growth: theory and evidence: a comment " , Journal of Development Economics 77: 251-256

Iara A., (2008) "Skill diffusion by temporary Migration?" WIIW working paper, July 2008.

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

Mayr, K. and G. Peri (2008), "Return Migration as Channel for Brain Gain," CReAM DP. 04/08.

Stark, O., (2004), "Rethinking the brain drain." World Development 32 (1), 15-22.

Stark, O., Helmenstein C., Prskawetz A., (1997), "A brain drain with a brain gain" Economics Letters 55: 227-234.

Stark, O., Helmenstein C., Prskawetz A., (1998), "Human capital depletion, human capital formation, and migration: a blessing or a "curse?" Economics Letters 60: 363-367.

Vidal, J.-P., (1998), "The effect of emigration on human capital formation."Journal of Population Economics 11: 589-600

APPENDIX A

Proof of Proposition 1: For k_1, there exist skilled workers, who have an incentive to migrate and the threshold ability level for migration is given by $\theta_1^{mig} = \frac{k_1}{w(\eta-1)e^{\gamma}}$. Since the workers of ability level $\theta \geq \theta_1^{mig}$ migrate with probability p at time period $t = 1$ and the number of permanent migrants increases, $k_2 < k_1$.

In particular, since $\frac{\partial \theta_t^{mig}}{\partial k_t} = \frac{1}{w(\eta-1)e^{\gamma}} > 0$, θ_t^{mig} is an increasing function of k_t. As k_t declines over time, θ_t^{mig} decreases until it reaches its lower bound $\widehat{k}$.

If $\theta_t^{mig} \leq \theta^{edu-h}$, then all educated individuals find migration more optimal than staying in the home country. To find the migration cost level such that $\theta_t^{mig} = \theta^{edu-h}$:

$$\frac{k_t}{w(\eta-1)e^{\gamma}} = \frac{w+\beta e}{w\beta(1+\beta)e^{\gamma}}$$

Solving for k_t yields k^* :

$$k^* = \frac{(w+\beta e)(\eta-1)}{\beta(1+\beta)}.$$

Hence, if $\widehat{k} > k^*$, then there are unskilled workers, skilled workers staying in the home country and permanent skilled migrants, and the corresponding threshold ability levels are $\theta^{edu-h}, \theta_t^{mig}$. If $\widehat{k} \leq k^*$, then the home country population consists of two groups. One group is composed of unskilled workers and the other is composed of the skilled workers, who have an incentive to migrate, and the corresponding ability level is θ_t^{edu-f}.

Proof of Proposition 2: Since $\theta_t^{ret} = \frac{k_t}{w(\eta-\mu)e^{\gamma}} > 0$ and $\frac{\partial \theta_t^{ret}}{\partial k_t} = \frac{1}{w(\eta-\mu)e^{\gamma}}$, there exist permanent migrants for all t and k_t decreases over time until it reaches its lower bound $\widehat{k}$.

If k_t has a lower bound such that $\theta_t^{mig-r} \leq \theta^{edu-h}$, then all individuals prefer migration to staying in the home country. To find the migration cost level such that $\theta_t^{mig-r} = \theta^{edu-h}$:

$$\frac{k_t}{w[(\eta - 1) + \beta(\mu - 1)]} = \frac{w + \beta e}{w\beta(1+\beta)e^{\gamma}}$$

Solving for k_t yields k^{**} :

$$k^{**} = \frac{(w + \beta e)[(\eta - 1) + \beta(\mu - 1)]}{\beta(1+\beta)}.$$

If k_t has a lower bound such that $\theta_t^{ret} \leq \theta_t^{edu-r}$, then all skilled migrants choose to become permanent migrants. To find the migration cost level such that $\theta_t^{ret} = \theta_t^{edu-r}$:

$$\frac{k_t}{w(\eta - \mu)e^{\gamma}} = \frac{w + pk_t + \beta e}{w[p\beta(\eta + \beta\mu) + (1-p)\beta(1+\beta)]e^{\gamma}}$$

Solving for k_t yields k^{***} :

$$k^{***} = \frac{(w + \beta e)(\eta - \mu)}{[A - p\beta(\eta - \mu)]},$$

where $A = [p\beta((\eta + \beta\mu) + (1-p)\beta(1+\beta)]$.

Hence, similar to the previous proposition, if $\widehat{k} > k^{**}$, then the corresponding threshold ability levels are $\theta^{edu-h}, \theta_t^{mig-r}, \theta_t^{ret}$. If $k^{***} < \widehat{k} \leq k^{**}$, then the corresponding ability levels are $\theta_t^{edu-r}, \theta_t^{ret}$. If $\widehat{k} \leq k^{***}$, then the corresponding threshold ability level is θ_t^{edu-f}.

Proof of Corollary 1: Clearly, for $\widehat{k} \geq k^{*}$, $\theta^{edu-h} = \theta_m$ and for $\widehat{k} < k^{*}$, $\theta^{edu-h} > \theta_t^{edu-f} = \theta_m$. Hence, the first inequality is established.

In order to establish the second inequality, we consider separate intervals of k_t :

For $\widehat{k} \geq k^{**}$, $\theta^{edu-h} = \theta_m = \theta_r$ and for $\widehat{k} \leq k^{***}$, $\theta_m = \theta_r = \theta_t^{edu-f}$.

For $k^{**} > \widehat{k} \geq k^{*}$, $\theta_m = \theta^{edu-h} > \theta_t^{edu-r} = \theta_r$.

For $k^{*} > \widehat{k} > k^{***}$, $\theta_m > \theta_r$ since $\frac{\partial \theta_t^{edu-r}}{\partial k_t} < \frac{\partial \theta_t^{edu-f}}{\partial k_t}$ and $\theta_t^{edu-r}, \theta_t^{edu-f}$ are linear in k_t. Therefore, the result follows.

Proof of Proposition 3: To prove this proposition, we search for pairs of employment options and consider the choices by workers between two employment options, and show that for some employment options, there does not exist any interval $\left(\theta^i, \theta^j\right)$, $i \neq j$ such that individuals of ability level in the interval $\left(\theta^i, \theta^j\right)$ choose those employment options.

For any k_t, consider the ability level which equates $y_{t,t+1}^{12}(\theta)$ and $y_{t,t+1}^{11}(\theta)$:

$$m_t(w_t^h + \beta w_{t+1}^h) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e = m_t(w_t^h + \beta w_{t+1}^h) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^r) - e.$$

$$w_{t+1}^f = w_{t+1}^r \Leftrightarrow w(1 + \eta\theta e^{\gamma}) - k_t = w(1 + \mu\theta e^{\gamma}) \Rightarrow \theta_t^{ret} = \frac{k_t}{w(\eta - \mu)e^{\gamma}}.$$

Hence, $y_{t,t+1}^{11}(\theta) \geq y_{t,t+1}^{12}(\theta)$ for $\theta \leq \theta_t^{ret}$ and workers of ability level $\theta \leq \theta_t^{ret}$ choose employment option 1.1. Further, $y_{t,t+1}^{12}(\theta) > y_{t,t+1}^{11}(\theta)$ for $\theta > \theta_t^{ret}$ and workers of ability level $\theta > \theta_t^{ret}$ choose employment option 1.2.

For any k_t, consider the ability level which equates $y_{t,t+1}^{12}(\theta)$ and $y_{t,t+1}^{24}(\theta)$:

$$m_t(w_t^h + \beta w_{t+1}^h) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e = m_t(w_t^f + \beta w_{t+1}^f) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e.$$

$$w_t^h + \beta w_{t+1}^h = w_t^f + \beta w_{t+1}^f \Leftrightarrow w(1 + \theta e^{\gamma}) = w(1 + \eta\theta e^{\gamma}) - k_t \Rightarrow \theta_t^{mig} = \frac{k_t}{w(\eta - 1)e^{\gamma}}.$$

Hence, $y_{t,t+1}^{12}(\theta) \geq y_{t,t+1}^{24}(\theta)$ for $\theta \leq \theta_t^{mig}$ and workers of ability level $\theta \leq \theta_t^{mig}$ choose employment option 1.2. Further, $y_{t,t+1}^{24}(\theta) > y_{t,t+1}^{12}(\theta)$ for $\theta > \theta_t^{mig}$ and workers of ability level $\theta > \theta_t^{mig}$ choose employment option 2.4.

Noting that $\theta_t^{mig} < \theta_t^{ret}$, when employment options $1.1, 1.2$ and 2.4 are compared on the ability interval, $y_{t,t+1}(\theta) = \max[y_{t,t+1}^{11}(\theta), y_{t,t+1}^{12}(\theta), y_{t,t+1}^{24}(\theta)] = \max[y_{t,t+1}^{11}(\theta), y_{t,t+1}^{24}(\theta)]$, therefore any worker chooses employment option 1.2.

For any k_t, consider the ability level which equates $y_{t,t+1}^{24}(\theta)$ and $y_{t,t+1}^{22}(\theta):$

$$m_t(w_t^f + \beta w_{t+1}^f) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e = m_t(w_t^f + \beta w_{t+1}^r) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e.$$

$$w_{t+1}^f = w_{t+1}^r \Leftrightarrow w(1 + \eta\theta e^\gamma) - k_t = w(1 + \mu\theta e^\gamma) \Rightarrow \theta_t^{ret} = \frac{k_t}{w(\eta - \mu)e^\gamma}.$$

Hence, $y_{t,t+1}^{22}(\theta) \geq y_{t,t+1}^{24}(\theta)$ for $\theta \leq \theta_t^{ret}$ and workers of ability level $\theta \leq \theta_t^{ret}$ choose employment option 2.2. Further, $y_{t,t+1}^{24}(\theta) > y_{t,t+1}^{22}(\theta)$ for $\theta > \theta_t^{ret}$ and workers of ability level $\theta > \theta_t^{ret}$ choose employment option 2.4.

For any k_t, consider the ability level which equates $y_{t,t+1}^{21}(\theta)$ and $y_{t,t+1}^{22}(\theta):$

$$m_t(w_t^f + \beta w_{t+1}^r) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^r) - e = m_t(w_t^f + \beta w_{t+1}^r) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e.$$

$$w_{t+1}^r = w_{t+1}^f \Leftrightarrow w(1 + \mu\theta e^\gamma) = w(1 + \eta\theta e^\gamma) - k_t \Rightarrow \theta_t^{ret} = \frac{k_t}{w(\eta - \mu)e^\gamma}.$$

Hence, $y_{t,t+1}^{21}(\theta) \geq y_{t,t+1}^{22}(\theta)$ for $\theta \leq \theta_t^{ret}$ and workers of ability level $\theta \leq \theta_t^{ret}$ choose employment option 2.1. Further, $y_{t,t+1}^{22}(\theta) > y_{t,t+1}^{21}(\theta)$ for $\theta > \theta_t^{ret}$ and workers of ability level $\theta > \theta_t^{ret}$ choose employment option 2.2.

When employment options $2.1, 2.2$ and 2.4 are compared on the ability interval, $y_{t,t+1}(\theta) = \max[y_{t,t+1}^{21}(\theta), y_{t,t+1}^{22}(\theta), y_{t,t+1}^{24}(\theta)] = \max[y_{t,t+1}^{21}(\theta), y_{t,t+1}^{24}(\theta)]$, therefore there does not exist any interval $\left(\theta^i, \theta^j\right)$, $i \neq j$ such that individuals of ability level in the interval

$\left(\theta^i, \theta^j\right)$ choose employment option 2.2.

For any k_t, consider the ability level which equates $y^{24}_{t,t+1}(\theta)$ and $y^{23}_{t,t+1}(\theta)$:

$$m_t(w^f_t + \beta w^f_{t+1}) + (1-m_t)(w^{fa}_t + \beta w^f_{t+1}) - e = m_t(w^f_t + \beta w^f_{t+1}) + (1-m_t)(w^{fa}_t + \beta w^r_{t+1}) - e.$$

$$w^f_{t+1} = w^r_{t+1} \Leftrightarrow w(1+\eta\theta e^\gamma) - k_t = w(1+\mu\theta e^\gamma) \Rightarrow \theta^{ret}_t = \frac{k_t}{w(\eta-\mu)e^\gamma}.$$

Hence, $y^{23}_{t,t+1}(\theta) \geq y^{24}_{t,t+1}(\theta)$ for $\theta \leq \theta^{ret}_t$ and workers of ability level $\theta \leq \theta^{ret}_t$ choose employment option 2.3. Further, $y^{24}_{t,t+1}(\theta) > y^{23}_{t,t+1}(\theta)$ for $\theta > \theta^{ret}_t$ and workers of ability level $\theta > \theta^{ret}_t$ choose employment option 2.4.

For any k_t, consider the ability level which equates $y^{23}_{t,t+1}(\theta)$ and $y^{21}_{t,t+1}(\theta)$:

$$m_t(w^f_t + \beta w^f_{t+1}) + (1-m_t)(w^{fa}_t + \beta w^r_{t+1}) - e = m_t(w^f_t + \beta w^r_{t+1}) + (1-m_t)(w^{fa}_t + \beta w^r_{t+1}) - e.$$

$$w^f_{t+1} = w^r_{t+1} \Leftrightarrow w(1+\eta\theta e^\gamma) - k_t = w(1+\mu\theta e^\gamma) \Rightarrow \theta^{ret}_t = \frac{k_t}{w(\eta-\mu)e^\gamma}.$$

Hence, $y^{21}_{t,t+1}(\theta) \geq y^{23}_{t,t+1}(\theta)$ for $\theta \leq \theta^{ret}_t$ and workers of ability level $\theta \leq \theta^{ret}_t$ choose employment option 2.1. Further, $y^{23}_{t,t+1}(\theta) > y^{21}_{t,t+1}(\theta)$ for $\theta > \theta^{ret}_t$ and workers of ability level $\theta > \theta^{ret}_t$ choose employment option 2.3.

When employment options 2.1, 2.3 and 2.4 are compared on the ability interval, $y_{t,t+1}(\theta) = \max[y^{21}_{t,t+1}(\theta), y^{23}_{t,t+1}(\theta), y^{24}_{t,t+1}(\theta)] = \max[y^{21}_{t,t+1}(\theta), y^{24}_{t,t+1}(\theta)]$, therefore there does not exist any interval $\left(\theta^i, \theta^j\right)$, $i \neq j$ such that individuals of ability level in the interval $\left(\theta^i, \theta^j\right)$ choose employment option 2.3.

For any k_t, consider the ability level which equates $y^{32}_{t,t+1}(\theta)$ and $y^{31}_{t,t+1}(\theta)$:

$$m_t(w^f_t + \beta w^f_{t+1}) + (1 - m_t)(w^h_t + \beta w^h_{t+1}) - e = m_t(w^f_t + \beta w^r_{t+1}) + (1 - m_t)(w^h_t + \beta w^h_{t+1}) - e.$$

$$w^f_{t+1} = w^r_{t+1} \Leftrightarrow w(1 + \eta\theta e^\gamma) - k_t = w(1 + \mu\theta e^\gamma) \Rightarrow \theta^{ret}_t = \frac{k_t}{w(\eta - \mu)e^\gamma}.$$

Hence, $y^{31}_{t,t+1}(\theta) \geq y^{32}_{t,t+1}(\theta)$ for $\theta \leq \theta^{ret}_t$ and workers of ability level $\theta \leq \theta^{ret}_t$ choose employment option 3.1. Further, $y^{32}_{t,t+1}(\theta) > y^{31}_{t,t+1}(\theta)$ for $\theta > \theta^{ret}_t$ and workers of ability level $\theta > \theta^{ret}_t$ choose employment option 3.2.

For any k_t, consider the ability level which equates $y^{21}_{t,t+1}(\theta)$ and $y^{31}_{t,t+1}(\theta)$:

$$m_t(w^f_t + \beta w^r_{t+1}) + (1 - m_t)(w^{fa}_t + \beta w^r_{t+1}) - e = m_t(w^f_t + \beta w^r_{t+1}) + (1 - m_t)(w^h_t + \beta w^h_{t+1}) - e.$$

$$\begin{aligned} w^{fa}_t + \beta w^r_{t+1} &= w^h_t + \beta w^h_{t+1} \Leftrightarrow [w(1 + \eta\theta^a_t e^\gamma) - k_t] + \beta w(1 + \mu\theta e^\gamma) \\ &= (1 + \beta)w(1 + \theta e^\gamma) \Rightarrow \overline{\theta_t} = \frac{w\eta\theta^a_t e^\gamma - k_t}{w(1 + \beta - \beta\mu)e^\gamma}. \end{aligned}$$

Hence, $y^{21}_{t,t+1}(\theta) \geq y^{31}_{t,t+1}(\theta)$ for $\theta \leq \overline{\theta_t}$ and workers of ability level $\theta \leq \overline{\theta_t}$ choose employment option 2.1. Further, $y^{31}_{t,t+1}(\theta) > y^{21}_{t,t+1}(\theta)$ for $\theta > \overline{\theta_t}$ and workers of ability level $\theta > \overline{\theta_t}$ choose employment option 3.1.

Thus, it remains to be shown that $\overline{\theta_t} > \theta^{ret}_t$:

$$\begin{aligned} \overline{\theta_t} - \theta^{ret}_t &= \frac{w\eta\theta^a_t e^\gamma - k_t}{w(1 + \beta - \beta\mu)e^\gamma} - \frac{k_t}{w(\eta - \mu)e^\gamma} \\ &= \frac{w\eta\theta^a_t e^\gamma}{w(1 + \beta - \beta\mu)e^\gamma} - \frac{k_t}{w(1 + \beta - \beta\mu)e^\gamma} - \frac{k_t}{w(\eta - \mu)e^\gamma} \end{aligned}$$

$$= \frac{\eta\theta_t^a}{1+\beta-\beta\mu} - \frac{k_t}{we^\gamma}\frac{(\eta-\mu)+(1+\beta-\beta\mu)}{(1+\beta-\beta\mu)(\eta-\mu)} = \frac{\eta\theta_t^a}{1+\beta-\beta\mu} - \theta_t^{ret}\frac{(\eta-\mu)+(1+\beta-\beta\mu)}{(1+\beta-\beta\mu)}$$

$$= \frac{1}{1+\beta-\beta\mu}[(\eta\theta_t^a - \eta\theta_t^{ret}) - \theta_t^{ret}(1-\mu)(1+\beta)] > 0.$$

Noting that $\overline{\theta_t} > \theta_t^{ret}$,when employment options $2.1, 3.1$ and 3.2 are compared on the ability interval, $y_{t,t+1}(\theta) = \max[y_{t,t+1}^{21}(\theta), y_{t,t+1}^{31}(\theta), y_{t,t+1}^{32}(\theta)] = \max[y_{t,t+1}^{21}(\theta), y_{t,t+1}^{32}(\theta)]$, therefore there does not exist any interval $\left(\theta^i, \theta^j\right)$, $i \neq j$ such that individuals of ability level in the interval $\left(\theta^i, \theta^j\right)$ choose employment option 3.1.

Proof of Corollary 2: For any k_t, consider the ability level which equates $y_{t,t+1}^0(\theta)$ and $y_{t,t+1}^{32}(\theta)$:

$$m_t(w_t^h + \beta w_{t+1}^h) + (1-m_t)(w_t^h + \beta w_{t+1}^h) - e = m_t(w_t^f + \beta w_{t+1}^f) + (1-m_t)(w_t^h + \beta w_{t+1}^h) - e.$$

$$w_t^h + \beta w_{t+1}^h = w_t^f + \beta w_{t+1}^f \Leftrightarrow w(1+\theta e^\gamma) = w(1+\eta\theta e^\gamma) - k_t \Rightarrow \theta_t^{mig} = \frac{k_t}{w(\eta-1)e^\gamma}.$$

Hence, $y_{t,t+1}^0(\theta) \geq y_{t,t+1}^{32}(\theta)$ for $\theta \leq \theta_t^{mig}$ and workers of ability level $\theta \leq \theta_t^{ret}$ choose employment option 0. Further, $y_{t,t+1}^{32}(\theta) > y_{t,t+1}^0(\theta)$ for $\theta > \theta_t^{mig}$ and workers of ability level $\theta > \theta_t^{mig}$ choose employment option 3.2.

For any k_t, consider the ability level which equates $y_{t,t+1}^{11}(\theta)$ and $y_{t,t+1}^0(\theta)$:

$$m_t(w_t^h + \beta w_{t+1}^h) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^r) - e = m_t(w_t^h + \beta w_{t+1}^h) + (1-m_t)(w_t^h + \beta w_{t+1}^h) - e.$$

$$\begin{aligned} w_t^{fa} + \beta w_{t+1}^r &= w_t^h + \beta w_{t+1}^h \Leftrightarrow [w(1+\eta\theta_t^a e^\gamma) - k_t] + \beta w(1+\mu\theta e^\gamma) \\ &= (1+\beta)w(1+\theta e^\gamma) \Rightarrow \overline{\theta_t} = \frac{w\eta\theta_t^a e^\gamma - k_t}{w(1+\beta-\beta\mu)e^\gamma}. \end{aligned}$$

Hence, $y_{t,t+1}^{11}(\theta) \geq y_{t,t+1}^{0}(\theta)$ for $\theta \leq \overline{\theta_t}$ and workers of ability level $\theta \leq \overline{\theta_t}$ choose employment option 1.1. Further, $y_{t,t+1}^{0}(\theta) > y_{t,t+1}^{11}(\theta)$ for $\theta > \overline{\theta_t}$ and workers of ability level $\theta > \overline{\theta_t}$ choose employment option 0.

Since $\theta_t^{mig} < \theta_t^{ret} < \overline{\theta_t}$, when employment options $0, 1.1$ and 3.1 are compared on the ability interval, $y_{t,t+1}(\theta) = \max[y_{t,t+1}^{0}(\theta), y_{t,t+1}^{11}(\theta), y_{t,t+1}^{31}(\theta)] = \max[y_{t,t+1}^{11}(\theta), y_{t,t+1}^{32}(\theta)]$, therefore all skilled workers have incentive to migrate.

Proof of Proposition 4: Since there does not exist any skilled worker who chooses to stay in the home country, only employment options $(i), (ii), (iii)$ and (iv) are left and the threshold values for the remaining employment strategies $(i), (ii), (iii)$ and (iv) are defined below:

In particular, the ability of a skilled worker, who is indifferent between (i) and (ii), can be found by:

$$y_{t,t+1}^{11}(\theta) = y_{t,t+1}^{21}(\theta),$$

$$m_t(w_t^h + \beta w_{t+1}^h) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^r) - e = m_t(w_t^f + \beta w_{t+1}^r) + (1-m_t)(w_t^{fa} + \beta w_{t+1}^r) - e.$$

Rearranging the above expression and substituting the explicit forms for the corresponding wages, it is obtained that:

$$\theta_t^{mig-r} = \frac{k_t}{w[(\eta - 1) + \beta(\mu - 1)]e^{\gamma}}.$$

Proceeding with the identification of the ability of a skilled worker, who is indifferent between (*ii*) and (*iii*), it can be found by:

$$y_{t,t+1}^{21}(\theta) = y_{t,t+1}^{24}(\theta),$$

$$m_t(w_t^f + \beta w_{t+1}^r) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^r) - e. = m_t(w_t^f + \beta w_{t+1}^f) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e.$$

Rearranging the above expression and substituting the explicit forms for the corresponding wages, it is obtained that:

$$\theta_t^{ret} = \frac{k_t}{w(\eta - \mu)e^{\gamma}}.$$

Finally, the ability of a skilled worker, who is indifferent between (*iii*) and (*iv*) can be found by:

$$y_{t,t+1}^{24}(\theta) = y_{t,t+1}^{32}(\theta),$$

$$m_t(w_t^f + \beta w_{t+1}^f) + (1 - m_t)(w_t^{fa} + \beta w_{t+1}^f) - e = m_t(w_t^f + \beta w_{t+!}^f) + (1 - m_t)(w_t^h + \beta w_{t+1}^h) - e.$$

Rearranging the above expression and substituting the explicit forms for the corresponding wages, it is obtained that:

$$\theta_t^f = \frac{w\eta\theta_t^a e^{\gamma} - (1 + \beta)k_t}{w(1 + \beta - \beta\eta)e^{\gamma}}.$$

Hence, referring to the employment options stated in Proposition 3, the partitioning of home country individuals is as follows[10]: individuals of ability level θ such that $\theta < \theta_t^{mig-r}$ choose employment option (i), of ability level $\theta_t^{mig-r} < \theta \leq \theta_t^{ret}$ choose employment option (ii), of ability level $\theta_t^{ret} < \theta \leq \theta_t^{f}$ choose employment option (iii), of ability level $\theta > \theta_t^{f}$ choose employment option (iv).

To obtain the threshold ability level, which determines educated individuals, one needs to consider the lifetime utility of the unskilled and skilled workers. More specifically, individuals in the home country compare the discounted lifetime utility of being an unskilled worker with choosing employment option (i) defined in the previous proposition[11]. Assuming that the threshold ability level acquiring education is lower than the ability level of individuals, who choose employment option (i), if $(1+\beta+\beta^2)w \geq p\beta y_{t,t+1}^{11}(\theta) + (1-p)\beta y_{t,t+1}^{0}(\theta)$ for an individual of ability θ, then the individual decides to become an unskilled worker. Hence, the threshold ability level for acquiring education is defined as:

$$\theta \geq \theta_t^{edu} = \frac{w - p\beta(1-m_t)(w\eta\theta_t^a - k_t) + \beta e}{w\beta[(1+\beta)(1-p+pm_t) + (1-m_t)p\beta\mu]e^{\gamma}}.$$

By Corollary 2, $y_{t,t+1}^{11}(\theta) \geq y_{t,t+1}^{0}(\theta)$ for $\theta \leq \overline{\theta_t}$, hence $y_{t,t+1}^{11}(\theta_t^{edu}) > y_{t,t+1}^{0}(\theta_t^{edu})$. Rewriting

[10] It should be ensured that $\theta_t^{mig-r} < \theta_t^{ret} < \theta_t^{f}$. The first part of the inequality has already been established. For the second part it suffices to show that $\theta_t^{f} > \overline{\theta_t}$ which holds by $\eta(\theta_t^a - \theta_t^{ret}) > \theta_t^{ret}[(1-\mu)(1+\beta)]$.

[11] We implicitly assume that $\theta_t^{edu} < \theta_t^{mig-r}$.This assumption is made in order to keep all employment options. For instance, one can assume that $\theta_t^{edu} < \theta_t^{ret}$, in that case there are no skilled workers choosing employment option 1.1.

$y^{11}_{t,t+1}(\theta^{edu}_t) > y^{0}_{t,t+1}(\theta^{edu}_t)$ explicitly and subtracting $(1+\beta)w(1+\theta^* e^\gamma)$ from both sides:

$$\begin{aligned} & (1+\beta)w(1+\theta^{edu}_t e^\gamma) - (1+\beta)w(1+\theta^* e^\gamma) \\ < \; & p[m_t(1+\beta)w(1+\theta^{edu}_t e^\gamma) + (1-m_t)(w(1+\eta\theta^a_t e^\gamma) - k_t) + \beta(w(1+\mu\theta^{edu}_t e^\gamma)] \\ & +(1-p)(1+\beta)w(1+\theta^{edu}_t e^\gamma) - (1+\beta)w(1+\theta^* e^\gamma) \end{aligned}$$

Substituting the definition of θ^{edu}_t and simplifying, we get:

$$(1+\beta)w(1+\theta^{edu}_t e^\gamma) - (1+\beta)w(1+\theta^* e^\gamma) < \frac{w}{\beta} + e - (1+\beta)w\theta^* e^\gamma = 0$$

Therefore, $\theta^{edu}_t < \theta^*$.

APPENDIX B

Derivation of the slope of the SS_f curve:

$$\frac{\partial\theta^a_t}{\partial\theta^f_t}|_{ss} = \frac{w(1+\beta-\beta\eta)e^\gamma}{w\eta e^\gamma} = \frac{1+\beta-\beta\eta}{\eta} > 0$$

Derivation of the slope of the DD_f curve

Following Stark and Chau (1999), one can find the slope of DD_f curve by differentiating (10) :

$$\frac{\partial\theta^a_t}{\partial\theta^f_t}|_{DD_f} = \frac{[\theta^f_t f(\theta^f_t) - \theta^{edu}_t f(\theta^{edu}_t)\frac{\partial\theta^{edu}_t}{\partial\theta^a_t}\frac{\partial\theta^a_t}{\partial\theta^f_t}][F(\theta^f_t) - F(\theta^{edu}_t)]}{[F(\theta^f_t) - F(\theta^{edu}_t)]^2} \quad (24)$$

$$-\frac{\int_{\theta^{edu}_t}^{\theta^f_t}\theta f(\theta)d\theta(f(\theta^f_t) - \theta^{edu}_t f(\theta^{edu}_t)\frac{\partial\theta^{edu}_t}{\partial\theta^a_t}\frac{\partial\theta^a_t}{\partial\theta^f_t}}{[F(\theta^f_t) - F(\theta^{edu}_t)]^2} \quad (25)$$

$$= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} - \frac{(\theta_t^{edu} - \theta_t^a) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\partial \theta_t^{edu}}{\partial \theta_t^a} \frac{\partial \theta_t^a}{\partial \theta_t^f} \text{ ,since } \theta_t^a = \frac{\int_{\theta_t^{edu}}^{\theta_t^f} \theta f(\theta) d\theta}{F(\theta_t^f) - F(\theta_t^{edu})}$$

$$= \frac{1}{\Psi} \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]}$$

Since $\theta_t^f > \theta_t^a$, if $\Psi > 0$, then it is confirmed that DD_f is upward sloping.

$$\Psi = 1 + \frac{(\theta_t^{edu} - \theta_t^a) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\partial \theta_t^{edu}}{\partial \theta_t^a}$$

$$= 1 + \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{p(1 - m_t)}{w[(1+\beta)(1 - p + pm_t) + (1 - m_t)p\beta\mu]e^\gamma} > 0$$

Proof of Lemma 1: To determine the relationships between $\theta_t^j = a, f$, and the variables k_t and m_t which are implicit in the equations. By totally differentiating (10):

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

$$= \frac{\theta_t^f f(\theta_t^f)}{F(\theta_t^f) - F(\theta_t^{edu})} d\theta_t^f - \frac{\theta_t^{edu} f(\theta_t^{edu})}{F(\theta_t^f) - F(\theta_t^{edu})} d\theta_t^{edu} - \theta_t^a [f(\theta_t^f) d\theta_t^f - f(\theta_t^{edu}) d\theta_t^{edu}]$$

$$= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} d\theta_t^f + \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} d\theta_t^{edu} \qquad (26)$$

Since $\theta_t^f - \theta_t^a > 0$ and $\theta_t^a - \theta_t^{edu} > 0, \theta_t^a$ is strictly increasing in θ_t^f and θ_t^{edu}.

By totally differentiating (9), $d\theta_t^f$ is obtained:

$$d\theta_t^f = \frac{\eta}{(1+\beta-\beta\eta)} d\theta_t^a - \frac{(1+\beta)}{w(1+\beta-\beta\eta)e^\gamma} dk_t \tag{27}$$

Hence, all else remaining constant θ_t^f is increasing in θ_t^a and decreasing in k_t.

To determine $d\theta_t^{edu}$, one needs to take the differential of (6) :

$$\begin{aligned} d\theta_t^{edu} &= \left[\frac{p\beta(w\eta\theta^a e^\gamma - k_t)}{C} - \frac{p(1+\beta-\beta\mu)w\theta_t^{edu}e^\gamma}{C} \right] dm_t \\ &\quad + \frac{p\beta(1-m_t)}{C} dk_t - \frac{p\beta(1-m_t)w\eta e^\gamma}{C} d\theta_t^a \end{aligned} \tag{28}$$

where $C = w\beta[(1+\beta)(1-p+pm_t) + (1-m_t)p\beta\mu]e^\gamma$

To examine the relationship between θ_t^a and m_t keeping all else constant, substitute (27) and (28) into (26) :

$$\begin{aligned} d\theta_t^a &= \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{\eta}{(1+\beta-\beta\eta)} d\theta_t^a \\ &\quad + \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \left\{ \begin{array}{c} \left[\frac{p\beta(w\eta\theta^a e^\gamma - k_t)}{C} - \frac{p(1+\beta-\beta\mu)w\theta_t^{edu}e^\gamma}{C} dm_t \right] - \\ \frac{p\beta(1-m_t)w\eta e^\gamma}{C} d\theta_t^a \end{array} \right\} \end{aligned}$$

$$= \frac{1}{\Delta} \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \left[\frac{p\beta(w\eta\theta^a e^\gamma - k_t)}{C} - \frac{p(1+\beta-\beta\mu)w\theta_t^{edu}e^\gamma}{C} \right] dm_t = \Sigma_m dm_t \tag{29}$$

$$\begin{aligned} \text{where } \Delta &= 1 - \frac{A\eta}{(1+\beta-\beta\eta)} + \frac{B[p\beta(1-m_t)w\eta e^\gamma]}{C}, A = \frac{(\theta_t^f - \theta_t^a)f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]}, \\ B &= \frac{(\theta_t^a - \theta_t^{edu})f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \end{aligned}$$

Therefore, by (29), the necessary and sufficient condition for θ_t^a to be increasing in m_t is

$\Delta > 0$

To analyze the relationship between θ_t^a and k_t keeping all else constant, substitute (27) and (28) into (26):

$$
\begin{aligned}
d\theta_t^a &= \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} \left(\frac{\eta}{(1+\beta-\beta\eta)} d\theta_t^a - \frac{(1+\beta)}{w(1+\beta-\beta\eta)e^{\gamma}} dk_t \right) \\
&\quad + \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \left[\frac{p\beta(1-m_t)}{C} dk_t - \frac{p\beta(1-m_t) w\eta e^{\gamma}}{C} d\theta_t^a \right]
\end{aligned}
$$

$$
= \frac{1}{\Delta} \left[- \frac{(\theta_t^f - \theta_t^a) f(\theta_t^f)}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{(1+\beta)}{w(1+\beta-\beta\eta)e^{\gamma}} + \frac{(\theta_t^a - \theta_t^{edu}) f(\theta_t^{edu})}{[F(\theta_t^f) - F(\theta_t^{edu})]} \frac{p\beta(1-m_t)}{C} \right] dk_t
$$

$$
= \frac{1}{\Delta} \left[-A \frac{(1+\beta)}{w(1+\beta-\beta\eta)e^{\gamma}} + B \frac{p\beta(1-m_t)}{C} \right] dk_t = \Sigma_k dk_t \tag{30}
$$

Hence, by (30) θ_t^a is decreasing in k_t if and only if $-A\frac{(1+\beta)}{w(1+\beta-\beta\eta)e^{\gamma}} + B\frac{p\beta(1-m_t)}{C} < 0$, provided that $\Delta > 0$.

To analyze the relationship between θ_t^f and m_t, one needs to consider (27) which yields:

$$
d\theta_t^f = \frac{\eta}{1+\beta-\beta\eta} \frac{\partial \theta_t^a}{\partial m_t} dm_t \tag{31}
$$

Thus, by (31), it is clear that if $\theta_t^a/\partial m_t > 0$, then θ_t^f is increasing in m_t. From (29), it is determined that θ_t^a is increasing in m_t if and only if $\Delta > 0$. Thus, θ_t^f is increasing in m_t provided that $\Delta > 0$.

To determine the relationship between θ_t^f and k_t, from equation (27), all else remaining constant it is obtained that:

$$
d\theta_t^f = \frac{\eta}{1+\beta-\beta\eta} \left(\frac{\partial \theta_t^a}{\partial k_t} \right) dk_t - \frac{(1+\beta)}{w(1+\beta-\beta\eta)e^{\gamma}} dk_t
$$

$$\frac{1}{1+\beta-\beta\eta}\left[\eta\left(\frac{\partial\theta_t^a}{\partial k_t}\right)-\frac{(1+\beta)}{we^{\gamma}}\right]dk_t \tag{32}$$

Hence, we have that θ_t^f decreasing in k_t if and only if $\eta\left(\frac{\partial\theta_t^a}{\partial k_t}\right)<\frac{(1+\beta)}{we^{\gamma}}$

Proof of Lemma 3: To determine the relationship between θ_t^{edu} and m_t, from (28), it is obtained that:

$$d\theta_t^{edu}=\left[\frac{p\beta(w\eta\theta^a e^{\gamma}-k_t)-p(1+\beta-\beta\mu)w\theta_t^{edu}e^{\gamma}}{C}\right]dm_t-\frac{p\beta(1-m_t)w\eta e^{\gamma}}{C}\left(\frac{\partial\theta_t^a}{\partial m_t}\right)dm_t$$

Thus, we have that θ_t^{edu} is increasing in m_t if and only if;

$$\left[\frac{p\beta(w\eta\theta^a e^{\gamma}-k_t)-p(1+\beta-\beta\mu)w\theta_t^{edu}e^{\gamma}}{C}\right]>\frac{p\beta(1-m_t)w\eta e^{\gamma}}{C}\left(\frac{\partial\theta_t^a}{\partial m_t}\right)$$

or if and only if

$$\left[\frac{p\beta(w\eta\theta^a e^{\gamma}-k_t)-p(1+\beta-\beta\mu)w\theta_t^{edu}e^{\gamma}}{p\beta(1-m_t)w\eta e^{\gamma}}\right]>\left(\frac{\partial\theta_t^a}{\partial m_t}\right)$$

Turning to the relationship between θ_t^{edu} and k_t, that all else remaining constant, from (28) :

$$d\theta_t^{edu}=\frac{p\beta(1-m_t)}{C}dk_t-\frac{p\beta(1-m_t)w\eta e^{\gamma}}{C}\left(\frac{\partial\theta_t^a}{\partial k_t}\right)dk_t$$

We have that θ_t^{edu} is increasing in k_t if and only if,

$$\left[\frac{p\beta(1-m_t)}{C}\right]>\frac{p\beta(1-m_t)w\eta e^{\gamma}}{C}\left(\frac{\partial\theta_t^a}{\partial k_t}\right)$$

or if and only if

$$1 > w\eta e^{\gamma}\left(\frac{\partial\theta_t^a}{\partial k_t}\right)$$

Proof of Proposition 5: Since $\eta(\theta_1^a - \theta_1^{ret}) > \theta_1^{ret}[(1-\mu)(1+\beta)]$, there exist individuals who migrate at $t = 1$. Thus, $m_2 = m(M_1) > m(M_0) = m_1$ and $k_2 = k(Z_1) < k(Z_0) = k_1$.Since $M_t \geq M_0$ and $Z_t \geq Z_0$ by $\frac{\partial\theta_t^{ret}}{\partial k_t} > 0$, $\Sigma_m dm_t + \Sigma_k dk_t > 0$, it is ensured that $\theta_t^a > \theta_1^a$ and hence, $\eta(\theta_t^a - \theta_t^{ret}) > \theta_t^{ret}[(1-\mu)(1+\beta)]$ holds $\forall t = 2, 3, ...$ Moreover, the condition $\Sigma_m dm_t + \Sigma_k dk_t > 0$ imposes that $\forall t$, $\theta_t^f > \theta_1^f$ by (31) and (32). Since $\theta_t^{edu} < \theta_t^{mig-r}$, we have $\theta_t^{edu} < \theta_t^{mig-r} < \theta_t^{ret} < \theta_t^f$ and $M_{t+i} \geq M_t, Z_{t+i} \geq Z_t, i = 1, 2, ...$Therefore, $m^* = \widehat{m}$ and $k^* = \widehat{k}$. *q.e.d*

.