DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

MIGRATION OF SKILLED WORKERS: POLICY INTERACTION BETWEEN HOST AND SOURCE COUNTRIES

Slobodan Djajic, Michael S. Michael and Alexandra Vinogradova

Discussion Paper 12-2012

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: +357-22893700, Fax: +357-22895028
Web site: http://www.econ.ucy.ac.cy

Migration of Skilled Workers: Policy Interaction between Host and Source Countries

Slobodan Djajić[*], Michael S. Michael[†]
and Alexandra Vinogradova[‡]

June 22, 2012

Abstract

This paper examines the interaction between policies of the host and source countries in the context of a model of skilled-worker migration. The host country aims to provide low-cost labor for its employers while also taking into consideration the fiscal burden of providing social services to migrant workers and their dependants. It optimizes by setting a time limit on the duration of a guest-worker's permit. The source country seeks to maximize its own welfare by optimally choosing the amount of training it offers to its citizens, some of whom may end up working abroad. Within this framework, we solve for the Nash equilibrium values of the policy instruments and compare them with the case where both countries cooperate to maximize joint welfare.

JEL Classification: F22

Key Words: Temporary Migration, Skilled Labor, Immigration Policy

[*]Corresponding Author: Slobodan Djajić, 11A, Avenue de la Paix, CH-1202 Geneva, Switzerland, Tel.: (41) 22 908 5934, Fax: (41) 22 733 3049,e-mail: slobodandjajic@yahoo.com. This paper was initiated when Djajić was visiting the Department of Economics of the University of Cyprus. I wish to thank the Department for its hospitality and support. We also wish to thank a co-editor of this Journal, James Hines, and three anonymous referees for very helpful comments on the first draft of this paper.

[†]Department of Economics, University of Cyprus, email: m.s.michael@ucy.ac.cy

[‡]Center of Economic Research, CER-ETH, Zurich, Switzerland. Tel: +41-44-632-49-49, email: avinogradova@ethz.ch

1 Introduction

Migration of skilled workers from the developing to the advanced countries has attracted considerable attention ever since Jagdish Bhagwati brought the brain-drain problem into focus in the 1970s. By recruiting skilled professionals from the developing countries, where education is heavily subsidized by the public sector, the advanced countries were widely viewed as pursuing policies detrimental to the source countries.[1] When migration of skilled workers is permanent, the bulk of the potential benefits stemming from public expenditures on training are lost from the perspective of the taxpayers.[2] When it is temporary, there is more scope for gains, especially if the returnees bring with them productive human capital accumulated while working abroad [see, e.g., Wong (1997), Dustmann (2001), Domingues Dos Santos and Postel-Vinay (2003), Meyr and Peri (2009), Dustmann et al. (2011), and Docquier and Rapoport (forthcoming)].

The vast majority of skilled migrants come from the developing and transition economies with the main poles of attraction being the U.S.A. and Canada, but also the economies of Western Europe [see Lucas (2005)]. Recent efforts to measure the magnitudes of these flows, including the works of Salt (1997), Carrington and Detragiache (1998), Docquier and Marfouk (2006), and Beine et al. (2007), reveal that the brain drain is a particularly

[1]It is well recognized that the problem is not only fiscal in nature. The presence of skilled workers in an economy is thought to generate positive externalities at various levels, including technological, social, political and economic. If we take the example of an important sector such as health care, massive emigration of professionals can have a devastating impact on the health status of the population in the short run and a strong negative influence on productivity and welfare in the long run.

[2]Note that even permanent migration can generate benefits for the source country through network effects, by developing business links at home, and through remittance flows. See, e.g., Grubel and Scott (1966), Bhagwati and Hamada (1974), McCulloch and Yellen (1977), Djajić (1986), Lopez and Schiff (1998), Rauch and Casella (2003), Kugler and Rapoport (2007), and Javorcik et al. (2011). In addition, a number of papers examine how the prospect of emigration can contribute to the accumulation of human capital in the source country by inducing individuals to invest more in their education [see, e.g., Mountford (1997), Wong (1997), Stark et al. (1997), Vidal (1998), Beine et al. (2001), and Mountford and Rapoport (2011)]. In an important recent study of this relationship, Beine et al. (2008) analyze data for 127 developing economies and find that doubling the emigration rate of the highly skilled induces the population of the source country to increase its human capital formation on the average by 5%.

acute problem for the relatively small developing countries. In terms of regions, island economies of the Caribbean and the Pacific, as well as countries in Central America, Sub-Saharan Africa, and South-East Asia have the highest skilled-emigration rates in proportion of their skilled populations.[3]

In the 21st century, emigration of skilled workers from the less developed parts of the world continues with a growing number of advanced countries offering fast-track labor-market access for skilled migrants through special temporary visa programs, such as the H1-B visa in the U.S.A. or the "Blue Card" in the EU.[4] In response to a severe shortage of health-care workers, Japan has entered into bilateral agreements with Indonesia, the Philippines, and Vietnam to recruit a certain number of nurses on the basis of three-year contracts.[5] Other countries aim to increase their stocks of highly trained workers by means of *permanent* immigration programs. The Canadian points system is a prominent example of this policy, also followed in slightly different forms by Australia, New Zealand and, more recently, Great Britain. In the U.S.A., special permanent residence visas for highly talented individuals have been available for decades.

These practices and policies clearly have an impact on the flows of highly trained migrants from the developing economies. The outflows of skilled workers reduce, in turn, the incentive for the authorities to provide public subsidies for higher education [see Justman and Thisse (1997)]. In an important recent paper, Docquier et al. (2008) examine this question both theoretically and empirically. On the basis of a sample of 108 middle-income and low-income countries they find a negative relationship between

[3]See Commander et al. (2004) and Docquier and Rapoport (2008) for very useful surveys of the various issues and evidence related to the brain drain.

[4]In the case of the European Blue Card initiative, highly-skilled non-EU nationals are granted renewable 2 year work permits. In addition, a holder of such a permit, who returns back to his/her country of origin after having worked in the EU for an extended period of time, has the possibility to reenter and work in the EU in the future without going through the application procedure over again (Council Directive 2009/50/EC).

[5]In theory, the foreign nurses can stay longer if they pass a Japanese nursing exam within the three-year period. As fluency in the Japanese language is difficult to achieve for these foreign workers within such a limited period of time, only one Philippino and two Indonesians out of a total of 251 managed to pass the exam in 2010 (see Asahi Shimbun (2010)).

education subsidies and skilled emigration rates. An obvious consequence is that the level of training and human capital possessed by the graduates (and thus skilled emigrants) is likely to be lower than it would be otherwise. Lower skills of migrants, in turn, affect the relationship between the costs and benefits of immigration from the perspective of the host countries. This can and does influence their immigration policies. The points systems of Canada, Australia and New Zealand are designed to filter out those with low training and skills. In the U.S.A., whether an H1-B worker can renew her temporary three-year visa depends on the willingness of the employer to sponsor a renewal, which depends to a large extent on the worker's training and ability.

The purpose of this study is to examine the brain-drain problem within a game-theoretic framework, where both the immigration policy of the host country and the optimal provision of higher education and training in the source country are endogenously determined. The analysis is conducted in the context of a simple two-country model developed in Section 2. The host country's objective is to support the profitability of enterprises employing skilled labor while also taking into account the fiscal impact of immigration. The latter consists of the immigration-induced increase in tax revenues minus the cost of public services absorbed by the skilled immigrants and their dependents. The policy instrument at the disposal of the host country is assumed to be the duration of time it allows migrants to work in the economy. The source country is assumed to provide education free of charge to its citizens, with the objective of maximizing its GDP. How much education is optimally provided depends on whether or not its citizens work abroad and, if they do, how long they stay.

Within this simple framework, Section 3 solves for the Nash equilibrium values of the policy instruments of both countries and examines how they respond to changes in the model's parameters. It is found that the host countries with relatively higher tax rates on income, where the authorities attach a relatively larger weight to employers' interests in their objective function, and where the public sector provides individuals with lower levels of social services, are countries that have stronger incentives to allow their skilled immigrants to work in the economy for a longer period of time. Whether a longer duration of stay raises or lowers the optimal level of training provided by the source country depends primarily on the rate at which immigrants accumulate skills while working abroad and the valuation of those skills after return. It is also found that an increase in the cost of providing public education reduces the equilibrium level of training and the amount of time immigrants are allowed to work in the host country. An increase in the home-country valuation of skills acquired by migrant workers abroad has the opposite effects on the two policy instruments: The source country provides more training and the host country allows migrants to stay longer. Finally, if the host country chooses to increase its *stock* of immigrants, this will either lower or increase the level of training provided by the source country, depending on the parameters of the model.

Section 4 extends the analysis to a setting where both countries set their policies to maximize joint welfare. In that case the level of training provided by the source country is higher in comparison with its Nash equilibrium value, while the duration of stay of immigrants in the host country may be either higher or lower. Section 5 looks at the equilibrium with permanent migration and Section 6 concludes the paper with a summary of the main findings.

2 The Analytic Framework

We consider a world consisting of two countries: An advanced labor-importing country and a less-developed country of emigration. The latter provides higher education and training to its citizens so as to maximize its GDP, net of training costs. Because potential earnings of skilled workers are higher abroad, some of the graduates will choose to migrate and thereby contribute to the GDP of the foreign rather than the home country. Migration opportunities may be temporary or permanent, depending on immigration policy of the host country, to which we now turn.

2.1 Host Country

The authorities of the host country, F, are typically concerned with two key issues when choosing the structure of their immigration policy. One of them is the fiscal impact of immigration: While employment of immigrants increases the economy's output and revenues of the fiscal authority, immigration also implies greater absorption of services provided by the public sector. This is a particular concern in the case of low-skilled workers (especially in economies that rely heavily on foreign sources of unskilled labor), although the issue is also important in the case of skilled workers in economies with generous social programs.[6]

Another key issue is the impact of immigration on the distribution of income between the native workers and their employers. Immigration allows employers to enjoy larger rents by hiring foreign workers. If the demand for labor expands, immigration prevents wages of natives from rising as much as they otherwise would, serving to redistribute income from native workers (and immigrants) to their employers. Broadly speaking,

[6] The various versions of the "points" system used in Canada, Australia and New Zealand, for example, are designed to attract skilled immigrants in the early phase of their productive lives, precisely because of the concern that their net contribution to the economy is likely to be negative if immigration takes place past a certain age. See DeVoretz and Ozsomer (1998) and DeVoretz (2001) for calculations on the net fiscal contribution of immigrants in Canada. Although immigration policies in the advanced countries have many dimensions, over the last couple of decades considerable attention has been focused on policy changes aimed at increasing the net fiscal contribution of immigrants. In addressing this issue, the 1996 Immigration Reform and Immigrant Responsibility Act in the United States has severely restricted immigrant access to means-tested social programs up until they become US citizens. In Western European countries, the conditions under which dependents of immigrants can reunite with the household head on a permanent basis have been tightened, with the effect of excluding those who are likely to become a heavy burden for the public sector. The instruments used include minimum-income and housing requirements that must be met by the sponsor. We do not model these instruments in the present study, as it would require much greater focus on the structural characteristics of immigrant households and potentially distract the reader from the main point of the paper.

the number of immigrants allowed to work in the economy reflects the influence that employers have in relation to native workers in shaping immigration policy.

We will not address this important domestic political-economy issue in the present study, as it has already received considerable attention. We will simply assume that the stock of immigrants, M, allowed to hold a valid work permit at any point in time is exogenously given, having been determined behind the scenes in a bargaining process involving various stakeholders in the host country.[7] We will focus, instead, on another key aspect of immigration policy that has not been treated in the theoretical literature on skilled-worker migration: The problem of deciding whether to admit immigrants on a permanent or temporary basis and, in the latter case, setting the optimal duration of the work permit.

With respect to the duration of stay, employers have a strong preference for having the same foreign worker over a relatively long period of time. High turnover is especially undesirable in the skilled occupations where the productivity of an employee can grow significantly with experience and on-the-job training, much of it being specific to the firm. We try to capture this in our analysis below by assuming that H, the marginal productivity of a skilled foreign worker, is an increasing function of the amount of time, t, spent on the job abroad, as well as her level of training, ε, at the time of arrival. A migrant's marginal productivity in host-country employment is thus given by $H(\varepsilon, t)$, where we assume $H_\varepsilon > 0$, $H_t > 0$, $H_{\varepsilon\varepsilon} < 0$, $H_{tt} < 0$, $\lim_{\varepsilon \to \infty} H_\varepsilon = 0$. One would also expect that $H_{\varepsilon t} \geq 0$.

Let the wage paid to foreign workers be a constant, w, which is lower than the marginal productivity of labor.[8] The average amount of rent, measured as a flow, enjoyed

[7] The numbers of immigrants admitted to the advanced countries are typically subject to numerical quotas for various types of workers, as in the case of the H1-B visa or the European "Blue Card" program, although in other cases the numbers merely represent loose targets, as in the case of Canadian immigration policy or that of Switzerland during its post-war boom.

[8] In the case of skilled HI-B workers in the USA, Martin, Chen and Madamba (2000) report evidence that foreign workers are paid less than the natives with comparable skills. In some economies, the underpayment

by an employer of a migrant worker is then

$$\frac{1}{\tau}\int_0^{\tau}\left[H(\varepsilon,t)-w\right]dt, \tag{1}$$

where τ represents the maximum duration of the work permit provided by the authorities.[9] If the permit is temporary, it is not renewable, requiring the migrant to return to the source country, S, on the date of expiration. Alternatively, if F offers permanent residence to a migrant worker, we assume that the latter does not return to S.

With respect to the fiscal impact of immigration, let us suppose that all income, whether from labor or capital, is taxed at the rate θ. The average flow of tax revenue from the output produced per migrant worker is then simply

$$\frac{1}{\tau}\int_0^{\tau}\theta H(\varepsilon,t)dt. \tag{2}$$

Concerning the cost of providing public services to an immigrant per unit of time, we shall assume that it amounts to a flow c if the migrant comes alone and $(1+a)c$ if s/he is accompanied by family members. The probability, π, that a migrant comes accompanied by family members, is clearly an increasing function of the expected duration of stay, τ. The cost of providing a migrant and any accompanying dependents with public services, measured as a flow, is therefore given by $c[1+a\pi(\tau)]$, where $\pi(\tau)\in[0,1]$ and a is likely to exceed unity.[10] It seems most realistic to assume that the second derivative of $\pi(\tau)$,

of migrant workers is institutionalized. For foreign contract workers in Taiwan, the wage set by the authorities is roughly one third lower than that paid to native workers. See Stein (2003).

[9] As hiring low-cost foreign labor generates a rent for an employer, there is an excess demand for migrant workers. For simplicity, we assume that employers are invited to participate in the program after being chosen at random by the authorities. The wage they are permitted to pay foreign workers is assumed to be strictly regulated and set below that received by native workers.

[10] In a dynamic setting, immigrant children (and particularly those of skilled immigrants in ageing societies) may have a positive net impact on public-sector finances. Chojnicki et al. (2011, p.344) find that the fiscal impact of immigration has been positive for the US economy, in spite of the fact that immigrants have been on average less educated than natives. This was mainly due to their younger age and higher fertility rates relative to natives, which resulted in a higher ratio of tax payers to beneficiaries of the welfare state. This potentially

$\pi_{\tau\tau} > 0$ for low values of τ, but becomes negative at some point as τ gets closer to T, where T is the length of the migrant's planning horizon. We shall therefore posit that the function $\pi(\tau)$ is initially increasing in a convex manner up to a certain (inflection) point after which it becomes concave.[11] We shall also assume that $\lim_{\tau \to 0} \pi(\tau) = 0$ and $\lim_{\tau \to T} \pi(\tau) = 1$.

Let us suppose that employers' rents and the net fiscal impact of hosting M migrant workers are the two key arguments in the objective function of the immigration authorities.[12] In this context, the problem for F is to choose τ that maximizes its objective function, W, which has two components: The flow of average annual (after-tax) rents enjoyed by the employers and the average annual net fiscal impact of hosting M migrant workers:

$$W = M\left[\frac{\lambda}{\tau}\int_0^\tau (1-\theta)[H(\varepsilon,t)-w]dt + \frac{\theta}{\tau}\int_0^\tau H(\varepsilon,t)dt - c\left[1+a\pi(\tau)\right]\right], \qquad (3)$$

where $\lambda \in (0,1)$ is the weight attached by the government to the employers' rents, captured by the first term in the large brackets, while the net fiscal impact is represented

positive impact of hosting dependents of immigrants is relevant for the case of permanent migration, examined in Section 5, but less so when the duration of the household head's stay in the host country is limited.

[11]This reflects the observation that for low values of τ, it is not economical for a migrant to bring the family along to the host country, as the associated migration costs impose a heavy burden without necessarily generating the offsetting benefits. For a low τ it makes more sense to leave the family in the source country, where the cost of consumption is typically lower and where the family can enjoy the continuity of residence along with a net increase in its standard of living due to higher earnings generated abroad by the household head. The vast majority of temporary migrants do in fact leave their family behind when the duration of the contract abroad is for just a year or two. For more extended stays, separation can become increasingly difficult to cope with and the advantage of avoiding migration costs and benefiting from the lower cost of family consumption at home can become small relative to the benefits of family unity. As the duration of stay abroad increases to the range of roughly 2-6 years, we would therefore expect π to rise quickly with τ and family migration to become the dominant mode. Further increases in τ can be expected to raise π further, but at a diminishing rate. The exact shape of the $\pi(\tau)$ function under various conditions in the host and source countries is an empirical question on which very little systematic data is available. Since the parameter values of the function are not crucial for the theoretical analysis of this paper, we leave this issue on the agenda for future research.

[12]On can easily add integration costs of immigration as a separate argument. For simplicity, we prefer to consider such costs as being reflected in the values of c and a.

by the difference between the last two terms. A necessary condition for the maximization of W with respect to τ is that

$$\frac{\partial W}{\partial \tau} \equiv W_\tau = \frac{M\left[\lambda(1-\theta)+\theta\right]}{\tau}\left[H(\varepsilon,\tau) - \frac{1}{\tau}\int_0^\tau H(\varepsilon,t)dt\right] - Mca\pi_\tau = 0, \tag{4}$$

where $H(\varepsilon,\tau)$ is the marginal productivity of a migrant worker at the moment just before she returns to the source country. Since we assumed that $H_t > 0$, $H(\varepsilon,\tau)$ is larger than the average productivity of a migrant worker, $\frac{1}{\tau}\int_0^\tau H(\varepsilon,t)dt$. This guarantees that the expression in the brackets of eq. (4) is positive. The last term captures the increase in the fiscal burden associated with the higher propensity for migrants to arrive accompanied by family members as τ is allowed to increase. In general, there can be zero, one, two or three internal values of τ that satisfy (4), given our assumptions on functions $\pi(\tau)$ and $H(\varepsilon,t)$. Note that $\tau = 0$ is never an optimum. Let us denote the vector of values of τ which satisfy eq. (4) by τ^0.

The second derivative of W with respect to τ is given by

$$\frac{\partial W_\tau}{\partial \tau} \equiv W_{\tau\tau} = \frac{M(\lambda(1-\theta)+\theta)}{\tau^2}\left[\tau H_\tau(\varepsilon,\tau) - 2H(\varepsilon,\tau) + \frac{2}{\tau}\int_0^\tau H(\varepsilon,t)dt\right] - Mca\pi_{\tau\tau}. \tag{5}$$

The first term in (5) is clearly negative (see Appendix A.1 for proof), while the second term can be either positive or negative, depending on whether τ lies on the convex or concave part of $\pi(\tau)$. Evaluating (5) at τ^0 we obtain the second-order condition

$$W_{\tau\tau}|_{\tau=\tau^0} = \frac{M(\lambda(1-\theta)+\theta)}{\tau^0}H_\tau(\varepsilon,\tau^0) - Mca\left(\frac{2\pi_\tau(\tau^0)}{\tau^0} + \pi_{\tau\tau}(\tau^0)\right) \gtrless 0,$$

Thus, the extrema τ^0 can be either maxima or minima (local or global). For a more detailed analysis of all possible outcomes see Appendix A.2. In the case of two extrema (one of which is necessarily a maximum and the other a minimum) we would also need to take into account the possibility of a corner solution $\tau = T$.[13] We examine the corner outcome in Section 5 on permanent migration, but for the moment wish to analyze a unique interior optimum such that $W_{\tau\tau}(\tau^0) < 0$. An analytical solution with specific functional forms is presented in Appendix B.[14]

2.2 Source Country

Suppose that the objective of the source country, S, is to maximize the welfare of its residents, while allowing them to have the freedom of international labor mobility. There is obviously a range of instruments available. The one we wish to focus on in the context of a model of skilled-worker migration is the level of public education and training, ε, provided to each member of the labor force. We shall assume that only the public educational system exists as liquidity-constrained households are unable to offer their children private education. Moreover, all students are assumed to be of identical ability.[15]

Education is costly, with government expenditure per individual assumed to be $x\varepsilon$, where x is the constant cost of providing more ε. The benefit of education for the economy manifests itself in a higher level of output, with the marginal productivity of a worker in source-country employment given by $H^*(\varepsilon)$ with $H^*_\varepsilon > 0$, $H^*_{\varepsilon\varepsilon} < 0$, and $\lim_{\varepsilon \to \infty} H^*_\varepsilon = 0$.[16]

[13] If τ^0 is a unique extremum and $W_{\tau\tau}(\tau^0) < 0$, then τ^0 is a global maximum. If there are three extrema, the first and the third are necessarily maxima, so that $W(T)$ cannot lie above the value of W evaluated at the third extremum. Thus, a corner solution $\tau = T$ may only occur when (a) W is monotonically increasing everywhere on $[0, T]$; (b) there are two extrema; (c) when W is monotonically increasing and has an inflection point, i.e., $W_\tau(\tau^0) = 0$ and $W_{\tau\tau}$ switches sign at τ^0. These cases are illustrated in the figure of the Appendix A.2: case (a) corresponds to Panel A on the left, case (b) to Panel B on the right, case (c) to Panel A on the right.

[14] Appendix B is available online at https://edit.ethz.ch/cer/resec/people/vinograa/Appendices_BCD.pdf

[15] The problem of international migration of skilled workers with heterogeneous abilities was first examined by Djajić (1989). We do not address this issue in the present study. Everyone in our model gets the same amount of education provided by the authorities and ends up with the same amount of skill when the training is completed.

[16] Note that we are assuming that local workers do not become more productive with experience in the source-country labor market. This is to sharpen our focus on the technological differences between countries

As some of the students will migrate at the time of graduation, the full benefits of the educational program are not captured by S. Some of the benefits spill over to F. This externality will obviously affect the optimal level of training provided to citizens. To define the problem in more concrete terms, let us assume that the objective of S is to maximize its steady-state GDP, net of educational expenditures. Suppose that L^* individuals are born at each instant, with their working lives being from the age of 0, when they graduate, to the age of T. The steady-state outflow of emigrants, M/τ, is set by the immigration policy of the host country, where M is the stock of migrants and τ is the duration of their stay abroad. Focusing here on the case of temporary migration, we may express the objective function of S as

$$W^* = (L^* - \frac{M}{\tau})TH^*(\varepsilon) + \frac{M}{\tau}(T - \tau)\phi H(\varepsilon, \tau) - xL^*\varepsilon, \tag{6}$$

where $\phi \leq 1$ is the proportion of a migrant's productivity in F, just before return, that is transferrable to the labor market of S. The first term in (6) corresponds to the productivity of the non-migrant population, the second term reflects the contribution of all the returnees and the last term corresponds to the public cost of education. One can assume that the returnees bring back valuable skills acquired abroad,[17] so that $\phi H(\varepsilon, \tau) > H^*(\varepsilon)$ or, at the other extreme, that the skills accumulated in F are largely firm specific and that having been away for τ units of time actually makes returnees less

and the possible benefits that a source country may enjoy due to return migration from a more advanced host country. None of the principal findings of the paper would change if we assumed that a worker's productivity is an increasing function of experience in the domestic labor market.

[17] Domingues Dos Santos and Postel-Vinay (2003) explicitly look at the effect of knowledge diffusion through return-migration. In their simple model they show that temporary migrants can boost the home country's productivity level by bringing a superior technology from the host country. In the long run this may lead to lower emigration and more return migration. Their analysis, however, is focused only on the sending (i.e., developing) economy, while our model considers the interaction between the policies of both the source and host countries. Dustmann et al. (2011) build a model in which individuals possess multiple skills and show that differences in the rates of return to these skills between the host and the source country may induce migrants to return home. By contrast, in our model, there is only one type of skill. See also a recent overview of this literature in Docquier and Rapoport (forthcoming).

productive in comparison with similarly educated non-emigrants [i.e., $\phi H(\varepsilon, \tau) < H^*(\varepsilon)$]. We shall ignore this second possibility on the grounds that it is much less likely to be empirically relevant than the first.

The source country will set ε to maximize W^*, so that

$$\frac{\partial W^*}{\partial \varepsilon} \equiv W^*_\varepsilon = (L^* - \frac{M}{\tau})TH^*_\varepsilon(\varepsilon) + \frac{M}{\tau}(T - \tau)\phi H_\varepsilon(\varepsilon, \tau) - xL^* = 0. \tag{7}$$

Given that H^*_ε and H_ε are both positive and monotonically declining in ε, with $\lim_{\varepsilon \to \infty} H_\varepsilon = 0$ and $\lim_{\varepsilon \to \infty} H^*_\varepsilon = 0$, the extremum of W^* is unique. Let us denote it by ε^0. The second-order derivative of W^* is

$$\frac{\partial W^*_\varepsilon}{\partial \varepsilon} \equiv W^*_{\varepsilon\varepsilon} = (L^* - \frac{M}{\tau})TH^*_{\varepsilon\varepsilon}(\varepsilon)dt + \frac{M}{\tau}(T - \tau)\phi H_{\varepsilon\varepsilon}(\varepsilon, \tau)dt < 0, \quad \forall \varepsilon, \tag{8}$$

ensuring that ε^0 is the global maximum. Rewriting (7) as

$$W^*_\varepsilon = L^*(TH^*_\varepsilon - x) + \frac{M}{\tau}\left[(T - \tau)\phi H_\varepsilon(\varepsilon, \tau) - TH^*_\varepsilon(\varepsilon)\right] = 0,$$

we see that if there is no migration (i.e., $M = 0$), the optimal level of training is such that, x, the marginal cost of an extra unit of education, is equal to $TH^*_\varepsilon(\varepsilon)$, which is the increase in the undiscounted lifetime productivity of a non-migrant.[18] With migration, either a higher or a lower level of training is optimal, depending on whether

$$D \equiv (T - \tau)\phi H_\varepsilon(\varepsilon, \tau) - TH^*_\varepsilon(\varepsilon) \tag{9}$$

is positive or negative, respectively. The second term in (9) corresponds to the increase

[18] Discounting the future benefits of public eduction would slightly complicate the notation. In terms of its impact on our findings, in an autarky equilibrium it would result in a lower ε, while in the case of temporary migration, with the benefits of education of those who migrate being deferred still further out in time, the effect on ε is even stronger. For formal treatment, see Appendix C, available online at https://edit.ethz.ch/cer/resec/people/vinograa/Appendices_BCD.pdf.

in the lifetime productivity of a non-migrant due to an increase in training by one unit. The first term captures a returnee's contribution to source-country output due to the same extra unit of training provided before emigration. If an additional unit of training results in an increase in the productivity of a returnee relative to that of a non-migrant in excess of $T/(T-\tau)$, then $D > 0$. In that case S benefits more by offering extra training to a worker who migrates temporarily than it does by offering it to another who remains at home. In consequence, it pays to provide more public education to citizens in a regime of temporary emigration than it does under autarky. Alternatively, if the skills accumulated in F are not easily transferrable to S (which might be due to a difference in the levels of development of the two countries) and/or $(T-\tau)/T$ is not sufficiently large, $D < 0$. It is then optimal to provide less training in the context of an open economy than it is under autarky. We shall consider both possibilities in the analysis below.[19]

3 Nash Equilibrium with Temporary Migration

Eqs. (4) and (7) are the reaction functions of F and S, respectively. The partial derivative of (4) with respect to ε is given by

$$\frac{\partial W_\tau}{\partial \varepsilon} \equiv W_{\tau\varepsilon} = M \left\{ \frac{(\lambda(1-\theta)+\theta)}{\tau} \left[H_\varepsilon(\varepsilon,\tau) - \frac{1}{\tau}\int_0^\tau H_\varepsilon(\varepsilon,t)dt \right] \right\} > 0. \qquad (10)$$

The sign of $W_{\tau\varepsilon}$ is positive because we assumed that $H_{\varepsilon t} \geq 0$, so that H_ε evaluated at $t=\tau$ is greater than the average of H_ε for $t \in [0,\tau]$. Since $W_{\tau\tau} < 0$ in the neighborhood of an internal solution for τ, the slope of the host country's reaction function, RR, is

[19] In a related paper, Wong and Yip (1999) consider an overlapping generations model of skilled migration, education, and endogenous growth. Emigration of skilled workers in their model lowers the growth rate of the economy, which in turn calls for greater expenditure on education by the authorities whose objective is to maintain the growth rate. The difference in the policy response to emigration of skilled workers in our model stems from the difference in the assumed policy objective.

positive (i.e., $d\tau/d\varepsilon|_{W_\tau=0} = -W_{\tau\varepsilon}/W_{\tau\tau} > 0$).

Differentiating the source-country reaction function (7) with respect to τ we obtain

$$\frac{\partial W_\varepsilon^*}{\partial \tau} \equiv W_{\varepsilon\tau}^* = \frac{MD}{\tau^2}\left(\xi_{D\tau} - 1\right), \tag{11}$$

where D is defined in (9) and the elasticity of D with respect to τ, $\xi_{D\tau} \equiv \frac{\partial D}{\partial \tau}\frac{\tau}{D} \gtrless 0$. The slope of the reaction function R^*R^* of country S is given by $d\tau/d\varepsilon|_{W_\varepsilon^*=0} = -W_{\varepsilon\varepsilon}^*/W_{\varepsilon\tau}^*$. Since $W_{\varepsilon\varepsilon}^* < 0$, the sign of the slope is the same as that of $W_{\varepsilon\tau}^*$ in eq. (11). It is therefore important to examine more closely the expression for $W_{\varepsilon\tau}^*$, which effectively determines whether it is optimal for S to increase or decrease spending on the training of its citizens in response to an increase in the value of τ chosen by country F. On the basis of (11), we observe that the slope of R^*R^* is positive in two cases. First, when $D > 0$ and $\xi_{D\tau} > 1$. A positive D means that the marginal effect of an extra unit of training on the productivity of a returnee exceeds the effect on the lifetime productivity of a non-migrant, i.e., there is a positive gap between these two marginal effects. The benefit of providing more ε is then larger for S, the greater the flow of migrants (and therefore returnees). An increase in τ reduces this flow in the same proportion because the stock of migrants, M, is held constant by F. This obviously calls for a reduction in ε. However, if $\xi_{D\tau} > 1$, the positive gap between the productivity of a returnee and a non-migrant expands *more* than in proportion to τ.[20] It then pays for S to raise ε in response to an increase in τ in spite of the associated reduction in the flow of returnees. R^*R^* is therefore positively sloped.

The second case in which the slope of R^*R^* is positive occurs when $D < 0$ and $\xi_{D\tau} < 1$. When $D < 0$, the benefit of providing more education to its citizens is larger for S, the *smaller* the flow of migrants. If, in addition, $\xi_{D\tau} < 1$, the reduction in the

[20] In general, an increase in τ has two effects on the gap. On the one hand, it reduces the time that a returnee spends back home ($T - \tau$ falls) and thus reduces her lifetime contribution to the GDP of S. On the other hand, it raises a migrant's marginal return to training ($H_{\varepsilon\tau} > 0$).

outflow of skilled workers due to an increase in τ has a more significant impact than any associated improvement in D. It is then beneficial, once again, for S to raise ε in response to a higher τ. In all other cases it is optimal for S to reduce the provision of public education in reaction to an increase in τ and hence R^*R^* is negatively sloped.

Figures 1 and 2 illustrate the determination of τ and ε in the Nash equilibrium. Figure 1 is drawn for the case $W^*_{\varepsilon\tau} < 0$ (negatively sloped R^*R^*) and Figure 2 for the case $W^*_{\varepsilon\tau} > 0$ (positively sloped R^*R^*). The host country's reaction function RR is positively sloped in both figures. Stability of the equilibrium requires that

$$\Delta \equiv W_{\tau\tau}W^*_{\varepsilon\varepsilon} - W_{\tau\varepsilon}W^*_{\varepsilon\tau} > 0,$$

which implies that if R^*R^* is positively sloped, it must be steeper than RR, as illustrated in Figure 2. We shall assume this to be the case.

FIGURES 1 AND 2 – POSITIONED HERE, SIDE BY SIDE

3.1 Comparative Statics

To examine the implications of changes in the key exogenous variables on the Nash-equilibrium values of the two policy instruments, we differentiate totally the reaction functions (4) and (7) to obtain

$$\begin{bmatrix} W_{\tau\tau} & W_{\tau\varepsilon} \\ W^*_{\varepsilon\tau} & W^*_{\varepsilon\varepsilon} \end{bmatrix} \begin{bmatrix} d\tau \\ d\varepsilon \end{bmatrix} = \begin{bmatrix} -W_{\tau\theta}d\theta - W_{\tau c}dc - W_{\tau\lambda}d\lambda \\ -W^*_{\varepsilon x}dx - W^*_{\varepsilon\phi}d\phi - W^*_{\varepsilon M}dM \end{bmatrix},$$

which enables us to solve for the effects of changes in the exogenous variables $\theta, c, \lambda, \phi, x,$

and M on the equilibrium values of τ and ε. The results are presented in the following subsections.

3.2 Increase in the Tax Rate in Country F

An increase in the tax rate, θ, of the host country has the following implications:

$$\Delta \frac{d\tau}{d\theta} = -W_{\tau\theta} W^{*}_{\varepsilon\varepsilon} > 0, \tag{12}$$

where $W_{\tau\theta} = M\frac{(1-\lambda)}{\tau}\left\{H(\varepsilon,\tau) - \frac{1}{\tau}\int_0^{\tau} H(\varepsilon,t)dt\right\} > 0$. It follows that a higher θ increases the Nash-equilibrium value of τ. Host countries with higher tax rates on earnings (including employer rents) can therefore be expected to allow skilled immigrants to stay longer. As we have assumed that the stock of migrants, M, is held constant, this comes at the expense of a smaller inflow of foreign workers.

The effect of a higher tax rate on the Nash equilibrium amount of training provided by country S is ambiguous and depends on the sign of $W^{*}_{\varepsilon\tau}$.

$$\Delta \frac{d\varepsilon}{d\theta} = W_{\tau\theta} W^{*}_{\varepsilon\tau} \gtreqless 0. \tag{13}$$

If $W^{*}_{\varepsilon\tau} < 0$, an increase in the tax rate lowers the amount of training, as that is the optimal response of S to a rise in τ. In terms of Figure 1, an increase in θ shifts the RR schedule up and to the left (shown by the dashed line $R'R'$), causing it to intersect the unaffected $R^{*}R^{*}$ locus at a lower value of ε. Alternatively, if $W^{*}_{\varepsilon\tau} > 0$, we have the case depicted in Figure 2, with an upward shift of RR giving rise to an increase in ε. This reflects the fact that when $W^{*}_{\varepsilon\tau} > 0$, an increase in each migrant's duration of stay abroad (along with a proportional reduction in the flow of returnees) actually raises the source-country benefit of training relative to the cost, making an increase in ε optimal.

3.3 Higher Cost of Public Services Absorbed by Immigrants

Consider next the implications of an increase in c, the cost of public services provided to immigrants:

$$\Delta \frac{d\tau}{dc} = -W_{\tau c} W^*_{\varepsilon\varepsilon} < 0, \tag{14}$$

$$\Delta \frac{d\varepsilon}{dc} = W_{\tau c} W^*_{\varepsilon\tau} \gtrless 0, \tag{15}$$

where $W_{\tau c} = -aM\pi_\tau < 0$. With an increase in c, the Nash-equilibrium duration of stay decreases. This stems from the assumption that if immigrants stay for a shorter period of time, they are less likely to bring with them their families that absorb costly public services. Thus, the more the public sector spends per unit of services provided to immigrants, the lower the value of τ. Host countries with highly developed welfare systems, particularly when it comes to services provided to dependent members of an immigrant household, can thus be expected to favor relatively shorter durations of stay.

The amount of training provided by the source country to its citizens either increases or decreases, depending on whether $W^*_{\varepsilon\tau}$ is positive or negative. The intuition here is the same as that in the previous subsection. The source country increases or cuts ε in response to a reduction in τ, depending on whether $W^*_{\varepsilon\tau}$ is negative or positive.

In the context of a somewhat richer model where the cost of providing public services to immigrants is a function of their education and skills, on might think of c as being a decreasing (possibly convex) function of ε. This modification of the model would not affect the qualitative results of our paper, but it would make the slope of RR steeper as the expression for $W_{\tau\varepsilon}$ would have an additional positive term, $-c'(\varepsilon)[1 + a\pi(\tau)] > 0$, where $c'(\varepsilon) < 0$.

3.4 Increase in the Weight of Employers' Rents

If the rents of host-country employers are assigned a larger weight, λ, in the objective function of country F, we have the following implications for the Nash-equilibrium values of τ and ε.

$$\Delta \frac{d\tau}{d\lambda} = -W_{\tau\lambda} W^*_{\varepsilon\varepsilon} > 0, \tag{16}$$

$$\Delta \frac{d\varepsilon}{d\lambda} = W_{\tau\lambda} W^*_{\varepsilon\tau} \gtreqless 0, \tag{17}$$

where $W_{\tau\lambda} = M\frac{(1-\theta)}{\tau}\left\{H(\varepsilon,\tau) - \frac{1}{\tau}\int_0^\tau H(\varepsilon,t)dt\right\} > 0$. A rise in λ therefore increases the Nash-equilibrium duration of stay while having an effect on ε that depends, once again, on the sign of $W^*_{\varepsilon\tau}$. This is precisely the same result that we had for an increase in θ and the same intuition follows.

3.5 Higher Transferability of Skills Acquired Abroad

An increase in ϕ has the following effects:

$$\Delta \frac{d\tau}{d\phi} = W_{\tau\varepsilon} W^*_{\varepsilon\phi} > 0, \tag{18}$$

$$\Delta \frac{d\varepsilon}{d\phi} = -W_{\tau\tau} W^*_{\varepsilon\phi} > 0, \tag{19}$$

where $W^*_{\varepsilon\phi} = \frac{M}{\tau}(T-\tau)H_\varepsilon(\varepsilon,\tau) > 0$. Greater source-country valuation of skills acquired by migrants in F increases the Nash-equilibrium amount of training and the duration of stay. If immigrants are effectively more productive at the point of return, it is then optimal for S to increase the amount of training it provides to all its citizens and for F to hold on to its skilled immigrants longer. This analysis suggests that over time, as source countries develop greater capacity to utilize the skills brought back by the returnees, the Nash-equilibrium values of ε and τ will tend to increase.

3.6 Increase in the Cost of Training

An increase in x is found to lower the Nash-equilibrium values of both ε and τ:

$$\Delta \frac{d\tau}{dx} = W_{\tau\varepsilon} W^*_{\varepsilon x} < 0, \tag{20}$$

$$\Delta \frac{d\varepsilon}{dx} = -W_{\tau\tau} W^*_{\varepsilon x} < 0, \tag{21}$$

where $W^*_{\varepsilon x} = -L^* < 0$. If there is an increase in the cost of training in country S, it no longer pays to provide as much of it as when the cost was lower. The optimal response of the host country is to cut the duration of stay of its skilled immigrants. In terms of Figures 1 and 2, an increase in x shifts the R^*R^* schedule to the left to intersect the unaffected RR locus at lower values of both ε and τ.

3.7 Increase in the Stock of Immigrants

Consider next a shift in immigration policy of country F that results in a larger desired stock of migrants, M, employed in the economy at any point in time. We have

$$\Delta \frac{d\tau}{dM} = W_{\tau\varepsilon} W^*_{\varepsilon M} \gtrless 0, \tag{22}$$

$$\Delta \frac{d\varepsilon}{dM} = -W_{\tau\tau} W^*_{\varepsilon M} \gtrless 0, \tag{23}$$

where $W^*_{\varepsilon M} = \frac{D}{\tau} \gtrless 0 \Leftrightarrow D \gtrless 0$, with D defined in (9). Since $W_{\tau\varepsilon} > 0$ and $W_{\tau\tau} < 0$, the Nash equilibrium values of τ and ε move in the same direction. They both decline if it is optimal for S to cut ε when its borders open up to temporary migration (i.e., $D < 0$) and increase when temporary emigration triggers an increase in ε (i.e., $D > 0$). The optimal response of country F is to shorten τ when training is reduced and to increase it when immigrants arrive with more skills.

4 Maximization of Joint Welfare

In this section we consider the case where country F chooses the duration of stay and country S chooses the amount of training to maximize joint welfare. The value of τ must then be set such that

$$\gamma W_\tau + (1-\gamma)W_\tau^* = 0. \tag{24}$$

The parameter $\gamma \in (0,1)$ is the relative weight attached to the welfare of F and may be interpreted to reflect its bargaining power.

Differentiating the welfare function of country S with respect to τ yields

$$W_\tau^* = \frac{M}{\tau^2}\left[TH^*(\varepsilon) - \phi(T-\tau)H(\varepsilon,\tau)\right] + \frac{M(T-\tau)\phi H(\varepsilon,\tau)}{\tau^2}\left[\eta_{H\tau} - \frac{\tau}{T-\tau}\right], \tag{25}$$

where $\eta_{H\tau} \equiv \frac{\partial H}{\partial \tau}\frac{\tau}{H}$. We can think of an increase in τ as having two effects on the welfare of S, represented by the two terms in eq. (25). First, for a given stock of migrants, an increase in τ implies a proportional reduction in the flow. More skilled workers therefore remain at home out of any generation of graduates, each contributing $TH^*(\varepsilon)$ to GDP of S. There is, however, a correspondingly smaller return flow of migrants, which implies a GDP loss amounting to $\phi(T-\tau)H(\varepsilon,\tau)$ units of output per returnee. If $TH^*(\varepsilon)$ is greater (smaller) than $\phi(T-\tau)H(\varepsilon,\tau)$, S experiences brain drain (gain) as a result of temporary emigration. A reduction in the flow of emigrants, due to an increase in τ, then benefits (harms) S, contributing to W_τ^* being positive (negative).

Second, with an increase in τ, each migrant stays abroad longer, accumulates skills, and returns to S with a higher productivity, albeit for a shorter period of time. This effect is captured by the second term in (25). If the elasticity of $H(.,.)$ with respect to τ, $\eta_{H\tau} > \tau/(T-\tau)$, an increase in τ contributes positively to source-country welfare through this channel. Such an outcome is likely to emerge in a migration regime where F allows migrants to stay for only a short period of time. For relatively high values of

τ, we would expect this second term in (25) to be negative.

In summary, taking into account both effects in (25), W_τ^* can be either positive or negative. The sign is unambiguously positive if S experiences a brain drain and migrants stay abroad for a relatively short period of time. Since $W_\tau = 0$ in the Nash equilibrium, $W_\tau^* > 0$ implies that joint welfare maximization calls for a relatively longer duration of stay for migrants in country F. Alternatively, if $W_\tau^* < 0$, joint welfare maximization results in a lower value of τ when compared with Nash.

Similarly, if country S chooses ε in order to maximize joint welfare of S and F, then

$$\gamma W_\varepsilon + (1-\gamma)W_\varepsilon^* = 0. \tag{26}$$

Differentiating the welfare function of country F with respect to ε, we find that

$$W_\varepsilon = M\frac{(\lambda(1-\theta)+\theta)}{\tau}\int_0^\tau H_\varepsilon(\varepsilon,t)dt > 0. \tag{27}$$

Since $W_\varepsilon^* = 0$ in the Nash equilibrium, joint welfare maximization requires a higher value of ε than the one that emerges in a non-cooperative setting.

In summary, maximization of joint welfare results in more training of workers by S and a longer or shorter duration of stay of skilled immigrants in F (depending on the sign of W_τ^*), when compared with the Nash-equilibrium values of these policy instruments. Note, in addition, that an increase in the bargaining power of F relative to that of S, as measured by γ, results in a higher ε and a shorter τ when $W_\tau^* > 0$ and a longer τ when $W_\tau^* < 0$. Moreover, maximization of joint welfare does not necessarily give rise to an increase in the individual level of welfare of both countries. Consider for example the case where W_τ^* is zero or close to zero. The duration of stay is then approximately the same with joint welfare maximization as it is at Nash, while the amount of training is higher. This means that the welfare of S is necessarily lower with joint welfare maximization than it is in the Nash equilibrium, while the welfare of F is unambiguously higher. In this case S has no incentive to cooperate and some side payment is needed in order to induce it to do so. A similar transfer mechanism might be necessary in order to induce S to cooperate in a situation where it is optimal for F to set $\tau = T$. This is the case of permanent immigration which we examine next.

5 Permanent Migration

Under certain conditions it is optimal for F to set $\tau = T$, i.e., invite skilled migrants to settle permanently. This corner solution may arise when (a) $\partial W/\partial \tau = 0$ has a unique root but is positive for all other values of τ, i.e., the objective function of F has an inflection point but is positively sloped everywhere else (see, e.g., Panel A on the right in Appendix A.2), or (b) $\partial W/\partial \tau = 0$ has two roots, the second of which is a (local) minimum (see Panel B on the right or Panel C on the left), or (c) the objective function W is positively sloped for all $\tau \in [0, T]$ (Panel A on the left). Case (c) requires no further discussion but in the other two cases it is possible that $W(\tau^0) < W(T)$. Evaluating the host country's objective (3) at τ^0 and T, we get

$$W(\tau^0) = M\left\{\frac{\lambda(1-\theta)+\theta}{\tau^0}\int_0^{\tau^0} H(\varepsilon,t)dt - \lambda w - c\left[1 + a\pi(\tau^0)\right]\right\}, \tag{28}$$

$$W(T) = M\left\{\frac{\lambda(1-\theta)+\theta}{T}\int_0^{T} H(\varepsilon,t)dt - \lambda w - c\left[1 + a\right]\right\}, \tag{29}$$

where we used the fact that $\lim\limits_{\tau \to T} \pi(\tau) = 1$. Subtracting (28) from (29), we find that the corner solution occurs when

$$[\lambda(1-\theta)+\theta]\left[\frac{1}{T}\int_0^{T} H(\varepsilon,t)dt - \frac{1}{\tau^0}\int_0^{\tau^0} H(\varepsilon,t)dt\right] - ca[1-\pi(\tau^0)] > 0.$$

That is, when the benefits of F stemming from the gain in a migrant's productivity (associated with the extension of the permit from τ^0 to T) more than compensate for the additional cost of public services provided to the immigrant household.

If migration is permanent, F simply retains a stock M of permanent immigrants, with a steady-state inflow of M/T skilled migrants filling the jobs of the retiring ones. The structure of the problem is then much simpler than in the case of temporary migration as τ is set at its maximum value of T. For S, the problem in this setting is to maximize

$$W^* = (L^*T - M)H^*(\varepsilon) - xL^*\varepsilon, \tag{30}$$

with respect to ε. This yields

$$\frac{\partial W^*}{\partial \varepsilon} = \left(\frac{L^*T - M}{L^*T} \right) TH^*_{\varepsilon}(\varepsilon)dt - x = 0, \tag{31}$$

which implies that the marginal cost of training must be equated to the product of the increase in the lifetime productivity of a non-migrant due to the extra unit of training and the proportion of graduates that remain at home. Comparing (31) with (7), we conclude that the optimal level of ε with permanent migration is unambiguously lower than that with temporary migration. Moreover, as the marginal productivity of training is assumed to be diminishing, it follows that the larger the stock of skilled migrants recruited on a permanent basis by F, the lower the optimal level of ε provided by S.

6 Conclusions

The vast literature on migration of skilled workers and the brain drain does not provide an analysis of the optimal interaction between immigration policy of the host country and the provision of public education in the source country. The present study attempts to fill this gap by developing a simple two-country model of skilled-worker migration where the host country chooses the optimal duration of stay of skilled migrants and the source country sets the level of training provided to its citizens.

In our analysis of the Nash equilibrium with temporary migration, we find that host countries that have relatively higher tax rates on incomes, that attribute a larger weight to employers' rents in their objective function, and that provide lower levels of public services to individuals, have a greater incentive to allow their skilled immigrants to work in the economy for a relatively longer period of time, including permanently. When a temporary immigration policy is chosen by the host country, the optimal level of training provided by the source country depends on the rate at which immigrants accumulate skills while working abroad and the valuation of those skills after return. Should the skills acquired abroad become more valuable in the labor market at home, it is optimal for the source country to provide a higher level of training to the workers. More training is also called for in response to a reduction in its cost. Finally, if the host country chooses to increase its *stock* of immigrants, this will lower (increase) the level of training provided by the source country if migration reduces (increases) its benefits from such training. This depends, in turn, on the rate at which migrants accumulate skills in the foreign country, the transferability of such skills to the labor market of the source country and the duration of each migrant's stay abroad. We also examine the implications of both countries acting to maximize joint welfare. The level of education provided to citizens of the source country is then greater, while the maximum duration of stay of migrant workers in the host country may be longer or shorter when compared with the Nash-equilibrium values of these instruments.

Our model can be extended to include the analysis of several host countries/regions that compete for skilled workers from a single source country/region. This problem would be more challenging and more interesting to consider in a setting where source-country workers differ in terms of their skills and host countries differ in terms of their technology. Moreover, in contrast with our simple model with infinitely elastic supply of migrants, host countries would have to make an effort to meet their immigration quotas. This implies that the stock of migrants becomes a key endogenous variable in their objective functions. To attract foreign workers, they would need to make compromises with respect to other objectives. We would expect this to be reflected in more favorable conditions being offered to migrants: conditions with respect to the duration of stay (i.e., longer τ), compensation (w), and even tax treatment, as we already observe in numerous advanced countries [See SOPEMI (2005, pp. 132-133)]. The optimal response of the source country is likely to be a cut in public expenditure on education below the level obtained under Nash equilibrium with a single host country.

There are a number of other directions in which the present model may be extended. In some cases this would complicate the analysis considerably, requiring simplifications in other dimensions. For example, our model has only one sector employing skilled labor with the authorities providing education to the entire labor force. A richer framework would consist of a two-sector economy, with one sector requiring skilled labor and the other unskilled labor. The size of the two sectors and the pattern of international trade in goods would then depend on the immigration and educational policies of the host and source countries, respectively. Second, as in Djajić (1989), one may look at the implications of emigration of skilled workers when individuals have heterogeneous abilities. In such a world, the workers with the highest abilities will likely be offered the strongest incentives to migrate, which in most modeling scenarios will accentuate the brain-drain effect for any given stock of migrants admitted abroad. These and other possible extensions of our model would contribute significantly to our understanding of the interaction between the optimal immigration and education policies of the host and source countries in a world where international mobility of skilled labor is becoming increasingly important.

References

[1] Asahi Shimbun, 2010, "Three Foreigners Pass Tough Nursing Exam," available at http://www.asahi.com/english/TKY201003260298.html

[2] Beine, M., F. Docquier, and H. Rapoport, 2001, “Brain Drain and Economic Growth: Theory and Evidence,” *Journal of Development Economics*, vol. 64: 275-289.

[3] Beine, M., F. Docquier, and H. Rapoport, 2007, "Measuring International Skilled Migration: A New Database Controlling for Age of Entry," *World Bank Economic Review*, vol. 21 (2): 249-254.

[4] Beine, M., F. Docquier, and H. Rapoport, 2008, “Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers,” *Economic Journal*, vol. 118: 631-652.

[5] Bhagwati, J.N., and K. Hamada, 1974, “The Brain Drain, International Integration of Markets for Professionals and Unemployment,” *Journal of Development Economics*, vol. 1: 19-42.

[6] Carrington, W.J., and E. Detragiache, 1998, “How Big is the Brain Drain,” IMF Working Paper 201.

[7] Chojnicki X., F. Docquier, and L. Ragot, 2011, "Should the US Have Locked Heaven’s Door?" *Journal of Population Economics*, vol. 24: 317-359.

[8] Council Directive 2009/50/EC available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:155:0017:0029:EN:PDF

[9] DeVoretz, D., 2001, "Canadian Immigration: Economic Winners and Losers," in S. Djajić, ed. *International Migration: Trends, Policies and Economic Impact* (London: Routledge).

[10] DeVoretz, D., and Y. Ozsomer, 1998, "Immigrants and Public Finance Transfers: Vancouver, Toronto and Montreal," Working Paper Series 99-06, Vancouver: Simon Fraser University.

[11] Djajić, S., 1986, "International Migration, Remittances and Welfare in a Dependent Economy," *Journal of Development Economics*, vol. 21: 226-234.

[12] Djajić, S., 1989, "Skills and the Pattern of Migration: The Role of Qualitative and Quantitative Restrictions on International Labor Mobility," *International Economic Review*, vol. 40: 795-809.

[13] Docquier, F., O. Faye, and P. Pestieau, 2008, "Is Migration a Good Substitute for Education Subsidies?"*Journal of Development Economics*, vol. 86: 263-276.

[14] Docquier, F., and A. Marfouk, 2006, International Migration by Educational Attainment (1990-2000)," in C. Ozden and M. Schiff, eds., *Internatinoal Migration, Remittances and the Brain Drain* (Palgrave-Macmillan).

[15] Docquier, F., and H. Rapoport, 2008, "Skilled Migration: The Perspective of Developing Countries," World Bank Policy Research Paper No. 3382.

[16] Docquier, F., and H. Rapoport, forthcoming, "Globalization, Brain Drain and Development,"*Journal of Economic Literature.*

[17] Domingues Dos Santos, M., and F. Postel-Vinay, 2003, "Migration as a Source of Growth: The Prospects of a Developing Country," *Journal of Population Economics*, vol. 16: 161-175.

[18] Dustmann, C., 2001, "Why Go Back? Return Motives of Migrant Workers," in S. Djajić, ed. *International Migration: Trends, Policies and Economic Impact* (London: Routledge).

[19] Dustmann, C., I. Fadlon, and Y. Weiss, 2011, "Return Migration, Human Capital Accumulation and the Brain Drain," *Journal of Development Economics*, vol. 95: 58-67.

[20] Grubel, H., and A. Scott, 1966, "The International Flow of Human Capital," *American Economic Review*, vol. 56: 268-274.

[21] Justman, M., and J. Thisse, 1997, "Implications of the Mobility of Skilled Labor for Local Public Funding of Higher Education," *Economics Letters*, vol. 55: 409-412.

[22] Javorcik, B.S., C. Özden, M. Spatareanu, and C. Neagu, 2011, "Migrant Networks and Foreign Direct Investment," *Journal of Development Economics*, vol. 94: 231-241.

[23] Kugler, M., and H. Rapoport, 2007, "International Labor and Capital Flows: Complements or Substitutes?" *Economics Letters*, vol. 94: 155-162.

[24] Lopez, R., and M. Schiff, 1998, "Migration and Skill Composition of the Labor Force: The Impact of Trade Liberalization in LDCs," *Canadian Journal of Economics*, vol. 31:318-336.

[25] Lucas, R.E.B., 2005, *International Migration Regimes and Economic Development*, Edward Elgar Publishing, Northampton, Massachusetts, USA.

[26] Meyr, K., and G. Peri, 2009, "Brain Drain and Brain Return: Theory and Application to Eastern-Western Europe," *The B.E. Journal of Economic Analysis & Policy*, vol. 9(1), Article 49.

[27] McCulloch, R., and J.T. Yellen, 1977, "Factor Mobility, Regional Development and the Distribution of Income," *Journal of Political Economy*, vol. 85: 79-96.

[28] Mountford, A., 1997, "Can a Brain Drain be Good for Growth in the Source Economy?" *Journal of Development Economics*, vol. 53: 287-303.

[29] Mountford, A., and H. Rapoport, 2011, "The Brain Drain and the World Distribution of Income," *Journal of Development Economics*, vol. 95: 4-17.

[30] Rauch, J.E., and A. Casella, 2003, Overcoming Informational Barriers to International Resource Allocation: Prices and Ties," *Economic Journal*, vol. 113: 21-42.

[31] Salt, J., 1997, "International Movements of the Highly Skilled," Occasional Paper no. 3, International Migration Unit, OECD.

[32] SOPEMI, 2005, *Trends in International Migration 2004*, OECD, Paris.

[33] Stark, O., C. Helmenstein, and A. Prskawetz, 1997, "A Brain Gain with a Brain Drain," *Economics Letters*, vol. 55: 227-234.

[34] Stein, N., 2003, "No Way Out," Fortune Magazine, January 30.

[35] Vidal, J.-P., 1998, " The Effect of Emigration on Human Capital Formation," *Journal of Population Economics*, vol. 11: 589-600.

[36] Wong, K.-y., 1997, "Endogenous Growth and International Migration," in B. Jensen and K.-y. Wong eds., *Dynamics, Growth and International Trade* (Ann Arbor: U. of Michigan Press).

[37] Wong, K.-y., and C.K. Yip, 1999, "Education, Economic Growth and Brain Drain," *Journal of Economic Dynamics and Control*, vol. 23: 699-726.

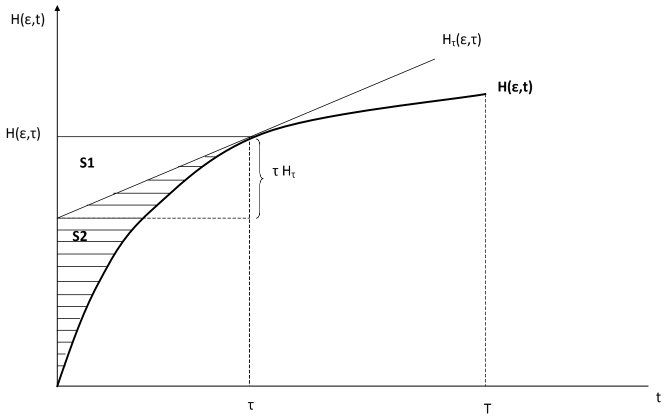
A Appendix to Section 2.1

A.1 Second Derivative of the Host Country's Objective

To see that the first term in (5) is negative, multiply the term in the square brackets by τ to get

$$\tau^2 H_\tau(\varepsilon,\tau) - 2\left[\tau H(\varepsilon,\tau) - \int_0^\tau H(\varepsilon,t)dt\right].$$

Note that the expression in the square brackets above is equal to the sum of the area marked by S_1 and the shaded area S_2 in the figure below.



Then write

$$\tau^2 H_\tau(\varepsilon,\tau) - 2\left[\tau H(\varepsilon,\tau) - \int_0^\tau H(\varepsilon,t)dt\right] =$$

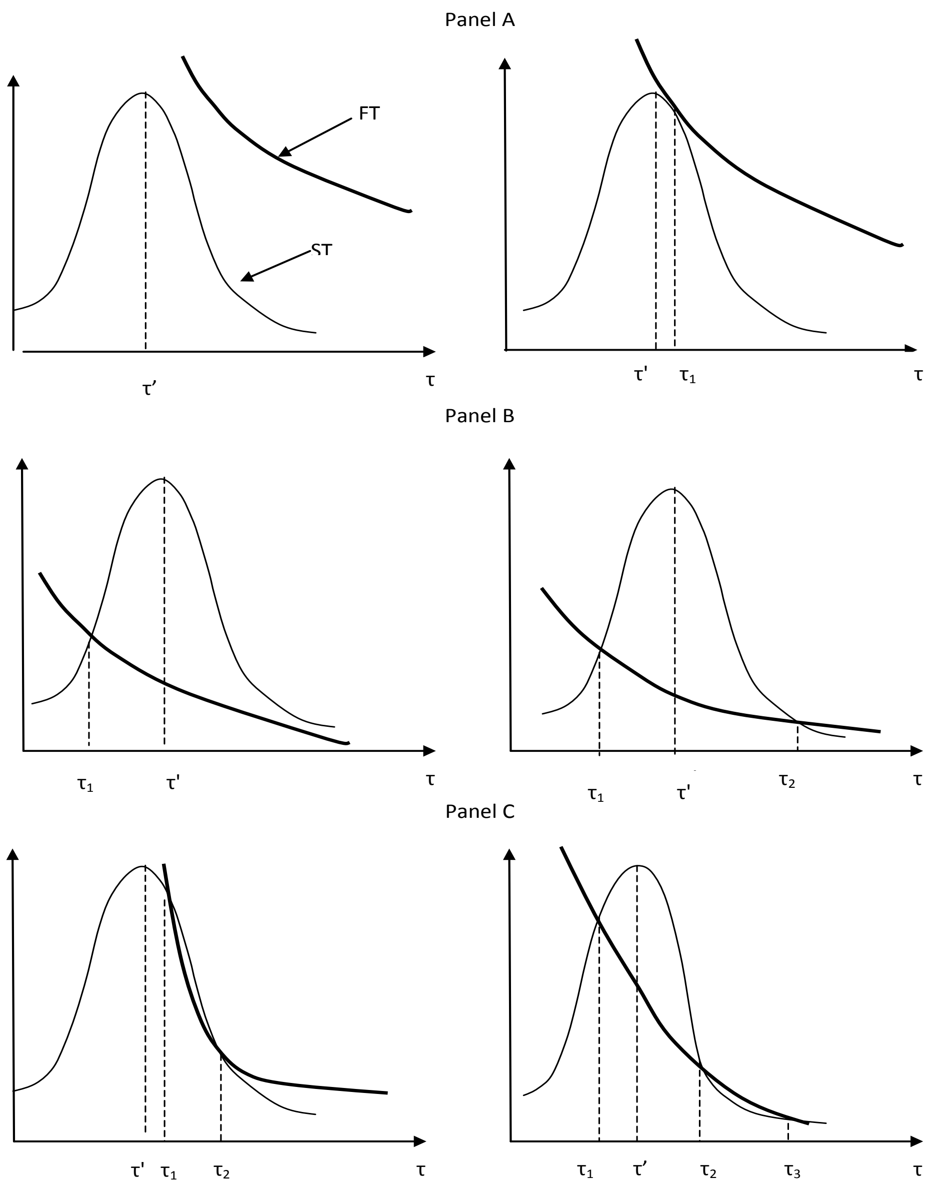
$$= \tau^2 H_\tau(\varepsilon,\tau) - 2(S_1+S_2) < \tau^2 H_\tau(\varepsilon,\tau) - 2S_1 = 0,$$

where the last equality follows from the fact that $S_1 = \tau^2 H_\tau(\varepsilon,\tau)/2$. Thus, the term in the brackets in (5) is unambiguously negative.

A.2 Optimal duration of the work permit

The first term (FT) in (4), $\frac{M[\lambda(1-\theta)+\theta]}{\tau}\left[H(\varepsilon,\tau)-\frac{1}{\tau}\int_0^\tau H(\varepsilon,t)dt\right]$, is positive and monotonically decreasing in τ, since, it's derivative with respect to τ, i.e., the first term in (5), is negative (proof in Appendix A.1). The second term (ST) in (4), $Mca\pi_\tau$, has a bell shape, with the maximum at the inflection point of $\pi(\tau)$, at $\tau=\tau'$ (see the figure below). The case with no interior solution corresponds to $W_\tau>0$, $\forall\tau$, so that the downward-sloping bold curve (labeled FT in the left half of Panel A) lies everywhere above the bell-shaped curve (labeled ST). It is then optimal for the host country to offer skilled migrants permanent residence. This corner solution is examined in Section 5.

The case of one optimum occurs if the downward-sloping FT curve just touches the ST curve, as shown on the right side of Panel A. This extremum cannot be a maximum, however, but rather an inflection point of $W(\tau)$, since the second derivative, $W_{\tau\tau}$, changes sign after passing through this point. A unique extremum may also occur if the FT curve crosses the ST curve from above and then lies everywhere below the decreasing portion of ST (see left side of Panel B, where the equilibrium is shown to occur to the left of the inflection point at $\tau=\tau'$). In this case, we have a global maximum. An extremum may also occur to the right of the inflexion point, on the downward-sloping portion of ST). Another possible case of two extrema is illustrated in Panel C on the left. Finally, three extrema may also occur, as shown in Panel C on the right. Among all these possible solutions we are interested only in maxima, that is, those which occur when FT crosses ST from above. In case of multiple maxima, as for example those at τ_1 and τ_3 in Panel C on the right, we cannot distinguish a local maximum from the global one without assuming specific functional forms.


Panel A
FT
ST
τ'
τ
τ'
τ1
τ
Panel B
τ1
τ'
τ
τ1
τ'
τ2
τ
Panel C
τ'
τ1
τ2
τ
τ1
τ'
τ2
τ3
τ

Figures

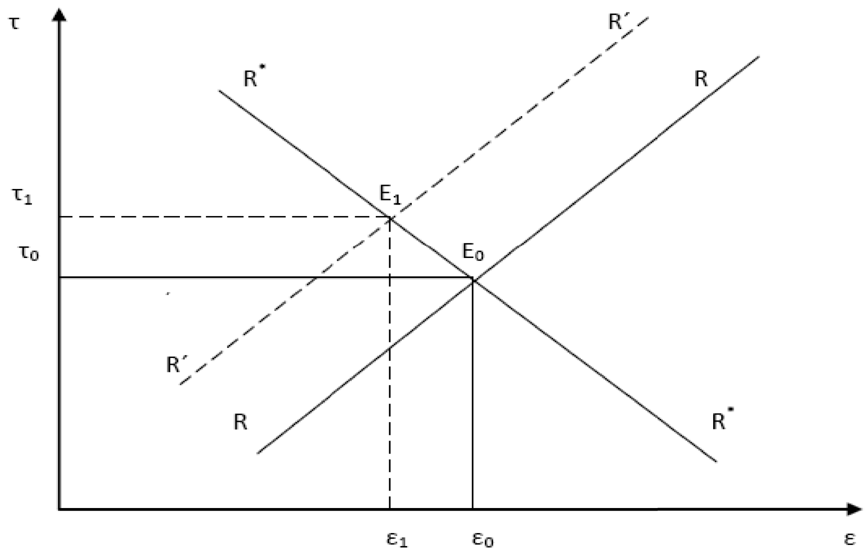


Figure 1: Solid lines depict the Nash equilibrium when the source country's reaction function R^*R^* is negatively sloped ($W^*_{\varepsilon\tau} < 0$). A higher tax rate on earnings in the host country, a lower cost of public services provided to immigrants or a higher weight attached to employers' rents in F, result in an upward shift of the host country's reaction function to $R'R'$, and hence a longer duration of the work permit and a lower level of public training.

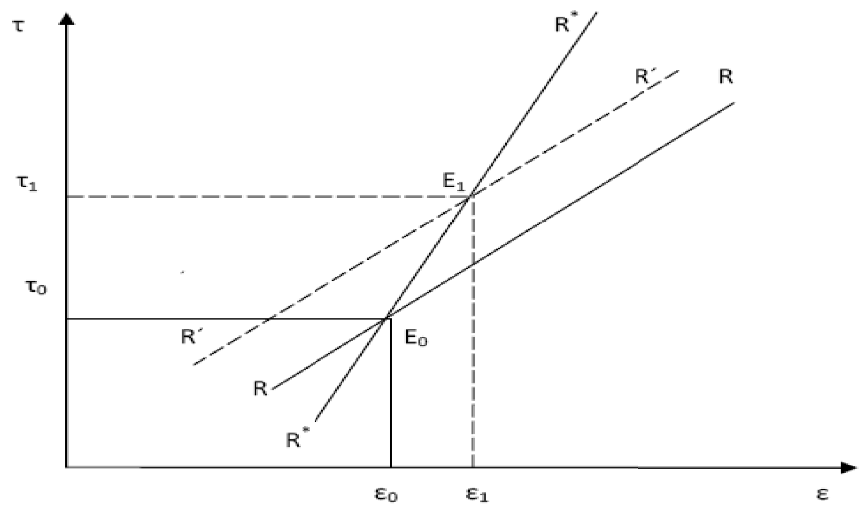


Figure 2: Solid lines depict the Nash equilibrium when the source country's reaction function R^*R^* is positively sloped ($W^*_{\varepsilon\tau} > 0$). A higher tax rate on earnings in F, a lower cost of hosting immigrants or a higher weight attached to employers' rents in F result in a longer duration of the work permit and more expenditure on public training provided by country S.