# Labor Migration and the Case for Flat Tax

**Tomer Blumkin**[*] **Efraim Sadka**[♠] **Yotam Shem-Tov**[♦]

**March, 2012**

**Abstract**

In this paper we employ a tax-competition model to demonstrate that in the presence of migration the re-distributive advantage of a non-linear income tax system over a linear (flat) one is significantly mitigated relative to the autarky (no-migration) equilibrium. When migration threats are sufficiently strong, a coordinated shift from a non-linear (*prima-facie* superior) system to a flat (inferior) regime is not too welfare-costly, even when the extent of re-distribution is significant. Therefore, such a shift may be warranted on administrative grounds. We also show, as expected, that migration reduces the extent of redistribution.

**JEL Classification:** D6, H2, H5

Key Words: flat tax, re-distribution, migration, tax-competition.

[*] Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel, CesIfo, IZA. E-mail: tomerblu@bgumail.bgu.ac.il.

[♠] The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, CesIfo, IZA. E-mail: sadka@post.tau.ac.il

[♦] The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, E-mail: yotam121@yahoo.com

## 1. Introduction

Linear (flat) tax systems are considered simpler, hence cheaper to administer than non-linear ones [often raised arguments refer to enhanced compliance and lower extent of avoidance associated with flat tax regimes; see Hall and Rabushka (1985), for an elaborate discussion of the merits of flat tax]. Reluctance to 'flattening' the tax system (via the consolidation of tax brackets and/or income sources), notwithstanding the administrative advantages associated with a flat tax, is often attributed to the latter's limited re-distributive capacity. Unlike a linear system, which in its most simple form, accords a universal demo-grant (basic income) across the board, a non-linear tax may employ means-testing to enhance the target-efficiency of the re-distributive system.

Prior to the 1990's hardly any countries enacted flat tax systems (a rare exception was Hong Kong). Since 1994, when the Baltic republics of Estonia and Lithuania first introduced a flat tax regime, many countries followed suit. In 2001 Russia introduced a flat personal income tax (PIT) of 13%.[1] By 2005, already nine countries in Central and Eastern Europe had a personal income flat tax in place, up and running (OECD, 2005).

In his seminal work Mirrlees (1971) raised the possibility that a flat tax may be the optimal choice of the government, by presenting simulations which showed that the optimal tax schedule for the US is approximately linear (notably, this result is derived without taking into account the additional administrative advantages associated with a flat system). This result has received fairly limited attention by the subsequent literature, and

[1] Notably, a year after the reform has been implemented, tax revenues from PIT increased by 46%. Ivanova et al. (2005) found that the reform had a significant behavioral effect on tax compliance. Gorodnichenko et al. (2007) found a strong influence of the reform on tax evasion.

has been disputed on the grounds of the parametric assumptions underlying the simulations, that were allegedly driving the result [see Tuomala (1990)]. Mirrlees (1971) examined a closed economy, where workers were utterly immobile. Four decades later, in the backdrop of a globalized world economy in which workers have become ever-more mobile, addressing the effect of labor migration on the extent of re-distribution and the desirable properties of the tax-and-transfer system has never been more timely.

The voluminous literature on tax competition has primarily focused on capital taxation [see Wilson (1999) for a comprehensive survey]. The key message conveyed by the literature suggests that competition over mobile capital would lead to inefficiently low taxes and under-provision of public goods, in contrast with the Tiebout paradigm [see Wilson (1986) and Zodrow and Mieszkowski (1986)]. This prediction is empirically supported by a documented shift from capital to labor taxation: over the years 1965-1995 the share of wage taxes in total tax revenues increased from 45% to 65% in the OECD.[2]

Another strand in the literature, more directly related to our paper, examines the optimal labor income tax system in the presence of mobile labor. A key feature of this literature is that competition over mobile labor limits the re-distributive power of the state [see Wilson (1980), (1992), Mirrlees (1982), Widasin (1994), Hindriks (1999) and Osmundsen (1999), amongst others)]. Most of the literature on tax competition and the effect of mobility on redistributive policy focused on linear tax-transfer schemes. Recently, several papers have revisited the issue in the context of non-linear taxation. Some papers cast the problem in a partial-equilibrium setting, examining the effect of migration on the properties of the optimal non-linear tax schedule of the state, taking the other states' tax schedules as exogenous outside options [see Wilson (2006), Krause (2007) and Simula and Trannoy (2010)]. Hamilton and Pestieau (2005) consider a general

[2] With the exception of Turkey and the UK.

equilibrium setting where some of the workers are (perfectly) mobile. They focus on the case of small-open economies, where each country ignores the effect of its re-distributive policy on international migration; hence, governments are not strategic competitors. Other papers [Piaser (2007), Brett and Weymark (2008) and Morelli et al. (2010)] consider a general equilibrium setting and explicitly model the strategic interaction across tax authorities. Piaser (2007) considers a setting with two identical countries and two skill levels, and demonstrates that when migration costs are sufficiently small, the income tax schedule entails no distortions. This result stands in contrast to the case of autarky (no migration), where a standard result in the literature suggests that low-skill individuals would be subject to a strictly positive marginal tax rate [Balcer and Sadka (1982) and Stiglitz (1982)]. Brett and Weymark (2008) consider a setting with two governments and a finite number of types. They illustrate a 'race to the bottom' argument in the case of tax competition and perfect mobility of the workforce, by showing that there do not exist equilibria in which either the most highly-skilled pay taxes or the lowest skilled receive transfers. Morelli et al. (2010) consider an extension of Piaser (2007) with three types of workers. They focus on the constitutional choice, within a federation comprised of two states, between a unified (centralized) tax system, where an identical tax system for both states is set by the central authority (hence there are no incentives to migrate between the two states); and an independent (decentralized) tax system, where the tax schedule is independently determined by each state, taking into account that citizens can migrate from one state to the other. They show that as migration costs rise, it becomes increasingly likely that the decisive middle class (the plausible scenario in the constitutional choice phase) will prefer to have a unified system.

In this paper, we re-visit Mirrlees (1971) by examining the case for a flat tax in the presence of migration threats. Employing the analytical framework used by Piaser (2007), we consider a tax competition game between two identical countries populated with individuals with two skill-levels. We compare between a non-linear tax regime and a flat tax system and demonstrate that in the backdrop of a high-skill migration threat (due to a reduction of the migration costs faced by high-skill individuals), the re-distributive advantage of a non-linear system over a linear (flat) one is significantly mitigated. In the presence of migration, and in sharp contrast to the autarky case, a coordinated shift to a flat system (with its entailed administrative advantages), still allowing for fiscal competition between countries (by maintaining the countries' sovereignty over the welfare state generosity), is not too welfare-reducing; and when administrative costs are taken into account, such a shift may prove to be mutually beneficial for both countries. We also examine the stability of the linear-tax equilibrium. Starting from equilibrium in the tax competition game between the two countries where both countries are restricted to linear schedules, we show that the gain associated with a unilateral (uncoordinated) deviation to a non-linear tax system by one of the two countries is fairly small, even when the extent of re-distribution is significant. Thus, taking into account the administrative gains associated with a flat system (relative to a non-linear tax regime), even when both countries may choose a general non-linear tax regime, an equilibrium where both do set a flat system in place is likely to form. We also confirm the race-to the-bottom hypothesis that suggests that migration reduces the extent of redistribution.

The structure of the remainder of the paper will be as follows. In section 2 we present the model. In section 3 we introduce the government problem and solve the tax competition game under the non-linear tax regime. In section 4 we present the tax competition game under the linear tax regime. Section 5 compares the two regimes. Section 6 examines the stability of the linear-tax equilibrium. Section 7 briefly concludes.

## 2. The Model

We consider a global economy which is comprised of two identical countries ($i$=1, 2). Each country produces a single consumption good employing labor inputs with different skill levels. We follow Mirrlees (1971) by assuming that the production technology exhibits constant returns to scale and perfect substitutability across skill levels.

Individuals differ in three attributes: (i) innate productive ability (skill-level), (ii) mobility costs (between the two countries) and (iii) mobility opportunity – only high-skill workers are able to migrate. For simplicity we assume that there are only two skill levels, where we denote by $w_1$ and $w_2 > w_1$, the productive ability (and the competitive wage rate) of the low-skill individual and high-skill individual, respectively. We follow Mirrlees (1971) by assuming that skill levels are private information unobserved by the government. We normalize the world population to 2 and assume that the measures of the low-skill population and the high-skill population are given, respectively, by $1 < \alpha_1 < 2$ and $0 < \alpha_2 < 1$, where $\alpha_1 + \alpha_2 = 2$. This assumption plausibly reflects the observed (right) skewed wage distributions.

Turning next to mobility costs, we assume that in the absence of any differences across the two countries (in terms of the fiscal policy implemented by the local government) the world population of each skill-group is equally divided between the two countries. Without being excessively unrealistic we assume that only high-skill individuals can migrate; that is, migration is prohibitively costly for all low-skill individuals.[3] The mobility cost, in consumption terms, incurred by a high-skill resident of

[3] For supporting empirical evidence see Docquier and Marfouk (2005) who show that in 2000, high skilled individuals were six times more likely to emigrate than low skilled ones. The departure of high-skill individuals to tax-havens has become a major concern amongst governments [OECD, (2002) and (2008)].

country $i$ who migrates to the other country, is denoted by $m$ and, in order to render our analysis tractable, is assumed to be distributed uniformly over the support $[0, \delta / 2]$. When the parameter $\delta$ assumes extreme values (zero or infinity), we obtain, respectively, the limiting cases of costless migration and no migration (autarky).

Individuals share the same preferences. Following Diamond (1998), we simplify by assuming that preferences are represented by some quasi-linear utility function of the form:

(1) $$U(c, l, d) = c - h(l) - d \cdot m,$$

where $c$ denotes consumption (gross of migration costs), $l$ denotes labor, $d$ is an indicator function assuming the value of one if the individual migrates and zero otherwise, and $h(\cdot)$ is strictly increasing and strictly convex.[4]

For later purposes, as is common in the optimal tax literature, we reformulate the utility (gross of migration costs) and represent it as a function of gross income ($y$), net income ($c$) and the individual's skill-level ($w$):

(2) $$V(w, c, y) \equiv c - h(y / w).$$

Hence, utility (net of migration costs) is given by:

(2') $$U(w, c, y, d) \equiv V(w, c, y) - d \cdot m.$$

---

In the numerical simulations we also examined the case where individuals of all skill-levels are faced with the same migration costs. Our qualitative results remain robust to this specification.

[4] For technical reasons, we make two additional assumptions. First, we assume that the term $h''/h'$ is (weakly) decreasing with respect to $l$. This assumption is satisfied, for instance, when $h$ is iso-elastic, which is the functional form used in our simulations and is commonly used in the literature [see Diamond (1998), Salanie (2003) and Simula and Trannoy (2010), amongst others]. In addition, we assume that $h''' \geq 0$. When $h$ is iso-elastic, the assumption implies that the (constant) elasticity of labor supply is bounded above by unity, which is consistent with empirical evidence [see, e.g., Salanie (2003)].

As we assume that the majority of the population (in each country) is of low skill, it is plausible to assume that each government will resort to some re-distributive policy towards the low-skill individuals. Thus, we assume, applying median-voter considerations, that the government of each country will maximize a *Rawlsian* social welfare function; that is, the utility of a representative low-skill resident.

## 3. The Government Problem

We turn next to formulate the government problem. For concreteness we will focus on country *i*=1, that takes as given the fiscal policy (tax and transfer system) implemented by country *i*=2. We will then solve for the symmetric *Nash* equilibrium of the fiscal-competition game formed between the two countries. We first introduce some useful notation. Denote by $\alpha_{ij}$ the measure of individuals of skill-level *j* in country *i*. Denote by $V_{ij}$ the utility level (gross of migration costs) derived by an individual of skill level *j* in country *i*. Finally denote by $c_{ij}$ and $y_{ij}$, correspondingly, the net income and gross income chosen by an individual of skill level *j* in country *i*.

By virtue of our quasi-linear specification, a high-skill individual who incurs mobility cost $m$ will migrate from country *i*=2 if, and only if, the following condition holds:

(3) $$V_{12} - m \geq V_{22}$$

Denote by $m^* \equiv V_{12} - V_{22}$, the cost of migration incurred by the high-skill individual who is just indifferent between staying in country 2 or migrating to country 1. Thus, any individual incurring a cost of migration lower than or equal to the above threshold will migrate to country 1. Recalling our assumption that migration cost is distributed uniformly over the support $[0, \delta/2]$ in both countries, it follows, by symmetry, that the term $\frac{\alpha_2 \cdot m^*}{\delta}$ represents the extent of migration of high-skill individuals between the two countries. If the term is positive there is migration from country 2 to country 1, and vice-versa.

Clearly, a more generous policy of the government in country *i*=1 towards high-skill individuals will attract more high-skill migration, *ceteris paribus*, and vice versa. In a symmetric equilibrium no migration will take place ($m^* = 0$), hence, $\alpha_{1j} = \alpha_j / 2 = \alpha_{2j}$.

The *Rawlsian* government in country *i*=1 is seeking to maximize the utility derived by a representative low-skill individual; namely:

(4) $$W = V(w_1, c_{11}, y_{11}),$$

subject to the following two self-selection/incentive compatibility constraints (for the low-skill individual and the high-skill individual, respectively), ensuring that each type of individual is as well-off with his bundle as he would be with mimicking the other type:

(5) $$V(w_1, c_{11}, y_{11}) \geq V(w_1, c_{12}, y_{12}),$$

(6) $$V(w_2, c_{12}, y_{12}) \geq V(w_2, c_{11}, y_{11});$$

a resource constraint:

(7) $$\sum_j \alpha_{1j} \cdot (y_{1j} - c_{1j}) \geq R,$$

where $R$ denotes the (pre-determined) level of government revenue needs;[5] and, a migration condition, which (endogenously) determines the number of high-skill individuals in country 1:

$$(8) \qquad \alpha_{12} = \alpha_2 \cdot \left[ 1/2 + \frac{1}{\delta} \cdot \left[ V(w_2, c_{12}, y_{12}) - V_{22} \right] \right].^{6}$$

Note, that unlike the standard formulation of the optimal tax problem, the number of high-skill individuals is endogenously determined, rather than being a fixed parameter. The standard case of no migration is obtained for the special limiting case where change: $\delta = \infty$. In this case, by virtue of (8), it follows that the number of individuals of each skill level is given by $\alpha_j / 2$. Each government takes the tax policy of the other country as given, i.e., country 1 takes $V_{22}$, the utility derived by the high-skill individuals in country 2, as given when choosing its tax policy. We will look for a symmetric *Nash* equilibrium for the fiscal-competition game between the two countries. Note that symmetry implies that in equilibrium the same tax schedule will be implemented by both countries.

## 3.1 Characterization of the Optimal Policy

[5] Note that when $R$>0, the equilibrium for the fiscal competition game between the two countries (to be characterized below) exists only for values of $\delta$ sufficiently bounded away from zero. To avoid this (purely technical) complication and to enhance the clarity of our presentation without changing the qualitative nature of our results, we will henceforth focus on the case where the fiscal-system is purely re-distributive ($R$=0). Note further, that setting $R$>0 does in fact strengthen our key results, by making the case for a flat system stronger.

[6] The formulation of the condition in equation (8) implicitly assumes an interior solution; namely, only a fraction of the high-skill population migrates in equilibrium. Notice, that in the symmetric equilibrium for the tax competition game between the two (identical) countries (to be characterized in what follows), no migration will actually take place. Thus, the necessary first-order (stability) conditions for each country (stating that no country will gain by deviating from the symmetric equilibrium profile) will indeed refer to an interior allocation. Notice further that as low-skill individuals cannot migrate, by assumption, it follows that $\alpha_{11} = \alpha_1 / 2$.

We turn next to characterize the solution for the government problem. It is straightforward to prove that the revenue constraint has to bind in the optimum.[7] Turning next to the two self-selection constraints, one can show that our formulation differs from the standard optimal tax setting with no migration, in that it may well be the case that in the optimal solution both self-selection constraints will not bind. Thus, in addition to the standard efficiency at the top property, we may well obtain no distortion at the bottom. We first state a result due to Piaser (2007), demonstrating that the patterns of binding self-selection constraints crucially hinge on the level of migration costs (all proofs and formal derivations are relegated to the Appendix).

**Proposition 1 :** There exists some critical level of migration costs, above which the high-skill self-selection constraint is binding, and below which both self-selection constraints are not binding. In the former case, only the marginal tax rate at the top is zero; in the latter case, the marginal tax rate at both the top and the bottom is zero.

**Proof:** See Appendix A.

According to proposition 1, when migration costs are sufficiently large (but still bounded away from infinity), the standard result in the literature applies; namely the incentive compatibility constraint associated with the high-skill individuals binds. However, when migration costs are small enough (but still bounded away from zero), both self-selection constraints do not bind. To see the intuition for this result, recall that an egalitarian government seeking to redistribute wealth from the high-skill towards the low-skill residents is essentially faced with two challenges. The first one is the standard one on the

[7] To see this, note that when the revenue constraint does not bind, the government may increase slightly the net income of both skill-levels by the same amount, thereby the utility of the low-skill individual, without violating the revenue constraint (by continuity considerations) or the two self-selection constraints (by construction).

intensive margin (which applies in the case of an autarky, as well) and derives from the mimicking threat of high skill individuals. The second one on the extensive margin (which applies only when tax competition takes place) derives from the migration threat of high-skill residents. With large enough migration costs, the impact of the extensive margin consideration (the potential threat of a massive migration of the high-skill) is relatively small; hence, the standard result (as in the case of autarky) applies. When migration costs are small enough the migration threat kicks in, in-earnest. Although the government can increase the tax burden shifted on the high-skill residents without inducing the latter to mimic, the reduction in the tax base due to the ensued migration is large enough to offset the gain from increasing the tax rate. The (lump-sum) tax levied on the high-skill residents in this case is essentially set (optimally) at the *Laffer* level; namely, the tax is set so as to maximize the total revenues raised from the high-skill population (hence the total transfers granted to the low-skill population).[8]

## 3.2 Characterization of the Equilibrium

We turn to solve the government problem. By virtue of symmetry, it suffices to focus on country 1. Formulating the *Lagrangean* yields

$$
\begin{aligned}
(9)\qquad & L \equiv V(w_1, c_{11}, y_{11}) + \lambda\left[V(w_2, c_{12}, y_{12}) - V(w_2, c_{11}, y_{11})\right] \\
& +\mu\left[\alpha_{12}\cdot(y_{1j} - c_{1j}) + \alpha_{11}\cdot(y_{11} - c_{11})\right],
\end{aligned}
$$

where $\lambda \geq 0$ (=0, when the incentive constraint is not binding, for small enough migration costs, as shown in proposition 1 above) and $\mu > 0$ denote, correspondingly, the *Lagrange*

[8] The lump-sum tax naturally introduces no distortions at the intensive margin, but being country-specific, does affect the decision on the extensive margin (whether or not to migrate). In this sense, the allocation attained in equilibrium is not first-best efficient.

multipliers associated with the high-skill individual self-selection constraint and the government revenue constraint, $\alpha_{12}$ is given by the condition in (8) and $\alpha_{11} = \alpha_1 / 2$, by virtue of the assumption that low-skill individuals cannot migrate.

The first-order conditions are given by:

(10) $1 - \lambda - \mu \cdot \alpha_{11} = 0$,

(11) $\partial V_{11} / \partial y_{11} - \lambda \cdot \partial V_{12} / \partial y_{11} + \mu \cdot \alpha_{11} = 0$,

(12) $\lambda + \mu[-\alpha_{12} + \alpha_2 / \delta \cdot (y_{12} - c_{12})] = 0$,

(13) $\lambda \cdot \partial V_{12} / \partial y_{12} + \mu[\alpha_{12} + \alpha_2 / \delta \cdot \partial V_{12} / \partial y_{12} \cdot (y_{12} - c_{12})] = 0$.

The optimal policy is given by a solution to the system of 7 equations [the four first-order conditions given in (10)-(13) and the three constraints in (6)-(8)]. When the self-selection constraint in not binding, $\lambda = 0$ (the constraint is dropped) and the optimal policy is obtained as a solution to a system of 6 equations.

We let $\hat{c_{11}}(V_{22}), \hat{y_{11}}(V_{22}), \hat{c_{12}}(V_{22})$ and $\hat{y_{12}}(V_{22})$ denote the optimal solution for the government problem in country 1 as a function of the utility derived by the high-skill individuals in country 2, $V_{22}$. A symmetric equilibrium for the game between the two countries is given by the implicit solution to the following equation:

(14) $V_{12} \equiv V[w_2, \hat{c_{12}}(V_{22}), \hat{y_{12}}(V_{22})] = V_{22}$.

In the symmetric equilibrium, by construction, the tax schedules implemented by both countries are identical and, therefore, no migration takes place. The tax schedule offered in equilibrium by country *i*=1 (and country *i*=2, by symmetry) is given by the 4-tuple:

$\hat{c_{11}}(\overline{V}_{22}), \hat{y_{11}}(\overline{V}_{22}), \hat{c_{12}}(\overline{V}_{22})$ and $\hat{y_{12}}(\overline{V}_{22})$ , where $\overline{V}_{22}$ is the implicit solution to the condition given in (14).

Employing the first-order conditions in (10)-(13), one can prove the following proposition:

**Proposition 2:** There exists a unique symmetric *Nash* equilibrium for the tax competition game between the two countries.

**Proof:** See Appendix B.

### 3.3 The Effect of Migration on the Optimal Tax Schedule

In this section we turn to investigate the effect of migration on the properties of the optimal tax schedules in equilibrium. For this purpose we conduct comparative static analysis with respect to the parameter $\delta$ which measures the intensity of migration. The lower the parameter is the lower are the mobility costs incurred by migrants; hence, the stronger are the migration pressures. The following proposition summarizes the comparative statics results:

**Proposition 3:** In the unique symmetric *Nash* equilibrium for the game between the two-countries, as mobility costs decrease: (i) the net transfers received by low-skill individuals as well as the net taxes paid by high-skill individuals decrease, (ii) the utility level of the low-skill individuals decreases, whereas that of the high-skill individuals increases.

Furthermore, when migration costs are sufficiently large, the marginal tax rate levied on low-skill individuals increases with respect to $\delta$.[9]

**Proof:** See Appendix D.

The implications of proposition 3 are straightforward. Further integration of the world economy, reflected in a reduction in mobility costs, triggers an enhanced fiscal competition between the two countries over mobile skilled-labor, which limits the scope of re-distributive policy. In the case where mobility costs are sufficiently large (in which the incentive constraint of the high-skill individuals is binding) the reduction in mobility costs would induce a 'flattening' of the tax schedule; that is, reduced differences in the marginal tax rates across income levels. The intuition for this result is as follows. As mobility costs decrease, governments offer less generous welfare policies to avoid emigration of high-skill individuals to the other country.[10] This implies that there is a lower incentive for the high-skill individuals to 'mimic' their low-skill counterparts in order to be eligible for welfare benefits, hence a lower need to impose a relatively high marginal tax rate at the bottom to deter such mimicking. Notice that in the case where mobility costs are small enough, as both incentive constraints do not bind; hence, there are no distortions either at the top or at the bottom, the marginal tax rate levied on both types of individuals is constantly zero, for any level of migration costs.

[9] Proposition 3 summarizes the comparative-static properties of the (unique) symmetric equilibrium. Unlike the symmetric equilibrium that exists for any level of migration costs, one can show (see Appendix C for details) that for sufficiently small levels of migration costs, no a-symmetric equilibrium exists.

[10] Notice that by virtue of the substitutability between the skill levels in the production function, migration of the high-skill does not affect the productivity of the low-skill (the standard brain-drain argument in the migration literature) but rather gives rise to a sort of 'fiscal brain-drain' effect through the erosion of the tax base of the government, thereby limiting the extent of re-distribution attained by the progressive tax-and-transfer system.

## 4. A Linear Tax Schedule

In this section we re-consider the tax competition game between the two countries assuming, now, that tax systems are restricted to be linear. This restriction implies, in particular, that transfers are accorded on a universal basis rather than being means-tested, as in the non-linear system. We first examine the effects of migration on the equilibrium and then turn (in the coming section) to compare the optimal linear tax schedule with the optimal non-linear system (characterized in the previous section).

We denote by *t* and *T*, respectively, the (constant) marginal tax rate and the (universal) demo-grant set by the government in country *i*=1 (taking as given the linear tax system in country *i*=2). We further denote by $y_{1j}(t,T)$ and $\alpha_{1j}(t,T)$, respectively, the gross income level chosen by a *j*-type individual and the number of *j*-type individuals residing in country *i*=1.[11] Maintaining our notation from the previous sections, the government is faced with the following revenue constraint (where we simplify by omitting the tax arguments to abbreviate notation):

(15) $$t \cdot (\alpha_{11} y_{11} + \alpha_{12} y_{12}) - \left(\alpha_{11} + \alpha_{12}\right) T \geq 0 .$$

Denoting by $V(w_{1j},t,T)$, the maximal utility derived by a *j*-type individual (in country *i*=1) faced with the linear tax system, (*t*, *T*), the government is seeking to maximize the well-being of the low-skill individual, $V(w_{11},t,T)$, subject to the revenue constraint in (15).

[11] $y_{1j}$ is given by the implicit solution to the individual first-order condition: $w_j \cdot (1-t) = h'(y_{1j} / w_j)$. It is straightforward to verify, by full differentiation of the first order condition with respect to the tax rate, *t*, employing the properties of the utility function ( $h'' > 0, h''' \geq 0$ ), that: $\partial y_{1j} / \partial t < 0$ and $\partial^2 y_{1j} / \partial t^2 \leq 0$. We will make use of these properties in the formal arguments in appendices E and F.

It is straightforward to verify that the revenue constraint in (15) is binding. Otherwise, the government could slightly increase the lump-sum transfer, thereby increasing the utility of the low-skill individuals. Attracting high-skill migrants (due to the higher demo-grant offered) will further expand the tax base and allow for enhanced redistribution. As in the previous sections we look for a symmetric Nash equilibrium for the game between the two countries.

Re-formulating the *Lagrangean* of the government problem (in country $i$=1) yields:

$$(16) \qquad L \equiv V(w_1, t, T) + \lambda\left[t\cdot(\alpha_{11}y_{11} + \alpha_{12}y_{12}) - (\alpha_{11} + \alpha_{12})T\right],$$

where $\lambda$ denotes the multiplier associated with the government budget constraint. Employing the migration condition in (8), modified to the case of the linear tax regime, the first-order conditions with respect to the two tax parameters ($t$ and $T$) are given by:

$$(17) \qquad \frac{\partial V_{11}}{\partial t} + \lambda\left(\frac{\alpha_2}{\delta}\frac{\partial V_{12}}{\partial t}ty_{12} + \frac{\partial y_{11}}{\partial t}t\alpha_{11} + \alpha_{11}y_{11} + \frac{\partial y_{12}}{\partial t}t\alpha_{12} + \alpha_{12}y_{12} - \frac{\alpha_2}{\delta}\frac{\partial V_{12}}{\partial t}T\right) = 0,$$

$$(18) \qquad 1 + \lambda\left[\frac{\alpha_2 t y_{12}}{\delta} - \frac{\alpha_2 T}{\delta} - (\alpha_{11} + \alpha_{12})\right] = 0.$$

Employing the first-order conditions in (17) and (18), one can show that there exists a unique symmetric Nash equilibrium for the game between the two countries. Formally,

**Proposition 4:** When both countries are restricted to linear tax schedules, a unique symmetric Nash equilibrium for the tax-competition game between the two countries exists.

**Proof:** see Appendix E.

The following proposition characterizes the effect of migration on the properties of the optimal linear tax schedules in equilibrium.

**Proposition 5:** In the unique symmetric Nash equilibrium for the game between the two-countries, as mobility costs decrease: (i) the lump-sum transfer decreases, (ii) the tax rate decreases and (iii) the utility level of the low-skill individuals decreases, whereas that of the high-skill individuals increases.

**Proof:** see Appendix F.

From proposition 5 it follows that a linear tax schedule has similar characteristics to a non-linear tax system. Under both tax systems, a decrease in the costs of migration (reflecting an enhanced threat of high-skill migration) implies that the government has to offer a less generous welfare system to its low-skill residents.

## 5. Comparing the Non- Linear and Linear Tax Schedules

Linear (flat) tax systems are commonly perceived to be much simpler and hence cheaper to administer (enhanced compliance, lower extent of avoidance etc.) than non-linear ones [see Hall and Rabushka (1985), for an elaborate discussion of the merits of flat systems]. Much of the criticism against a reform aiming at 'flattening' the tax system (say, through the consolidation of tax brackets and/or income sources) despite its well-known entailed administrative gains, dwells on its perceived limited re-distributive capacity. A linear system accords a universal demo-grant across the board and, thus, fails to employ screening devices (notably, means-testing) to enhance the target-efficiency of the tax-transfer system. In this section we demonstrate that in the backdrop of a high-skill migration threat (due to a reduction of the migration costs faced by high-skill individuals), the re-distributive advantage of a non-linear system is significantly mitigated. Thus, in the presence of migration and in sharp contrast to the autarky case, a coordinated shift to a flat system (with its entailed administrative advantages), still allowing for fiscal competition between countries (by maintaining the countries' sovereignty over the welfare state generosity), may well prove to be mutually beneficial for both countries.[12]

One obvious case in which the advantage of the non-linear system (relative to a flat one) utterly disappears is the limiting case of costless migration. With costless migration ($\delta = 0$) the redistributive system ultimately unravels under both tax regimes and the equilibrium of the tax competition game converges to the *laissez-faire* (redistributive-free) allocation under both tax systems (a standard *Bertrand*-type competition argument). By continuity considerations, with sufficiently small migration costs, the welfare gain associated with a shift from a flat to a non-linear system is small enough to render the former system preferred due to its administrative advantages. However, costless (or almost costless) migration is clearly an unrealistic paradigm for drawing concrete policy conclusions. We thus turn next to demonstrate that even in the far more plausible case where migration costs are sufficiently bounded away from zero, so that the extent of re-distribution attained (in equilibrium) is substantial, the welfare difference between the two tax regimes is fairly small. Being unable to provide a closed-form solution, we resort to numerical simulations, based on a calibrated version of the model.

We make the following parametric assumptions for the numerical analysis. We follow Simula and Trannoy (2010) in assuming that the utility function takes the

[12] Coordination can take different forms such as binding international agreements as part of a treaty (such as the EU) or via a federal system in which the restriction can be imposed on the states by the federal authority.

following (iso-elastic) functional form: $V(w_i, c_i, y_i) = c_i - \frac{(y_i / w_i)^{1+1/e}}{1+1/e}$, where $e$ is measuring the labor supply (taxable income) elasticity. We calibrate the wage rates (of both types of individuals) and the proportion of high-skill workers, using data from the US Bureau of Labor Statistics (BLS). By virtue of the assumption that the low-skill individuals form the majority of the population ($\alpha_2 < 1$) it follows that the median wage rate is equal to the wage rate earned by a low-skill individual, $w_{med} = w_1$. Using the median hourly wage rate of $w_{med} = 17.9$ [National Compensation Survey (2009)], the elasticity of taxable income *e*=0.4 [Gruber and Saez (2002)] and assuming a constant tax rate of 40 percent [Saez (2002)], one can solve for the gross income level earned by a high-skill individual.[13] The mean income, $y_{mean} = 22.36$ [Current Population Survey (2009)] is given by $y_{mean} = \frac{\alpha_2}{2} y_2 + \left(\frac{2-\alpha_2}{2}\right) y_1$. In order to obtain the gross income earned by a high-skill individual, $y_2$, we define high-skill (respectively, low-skill) workers as those individuals who earn above (below) the mean income and, accordingly, set the proportion of high-skill workers in the population at $\alpha_2 / 2$ =0.3518 [Current Population Survey (2009)], in line with our assumption that the income distribution is right-skewed. We then solve for the wage rate of the high-skill individual (employing the same parametric assumptions used above with respect to the elasticity of taxable income and the marginal tax rate in place) to obtain $w_2 = 45.645$. We turn next to discuss the results.

Figure 1 below demonstrates the difference between the welfare levels associated with equilibrium of the tax competition game under the two tax regimes (non-linear versus linear) as a function of the cost of migration. The difference between the two

[13] Our estimated parameters are robust to the tax rate being used.

regimes is measured on the vertical axis in compensating-variation terms (as a fraction of the *laissez-faire* output). As a guide to interpreting the figure, notice that in the limiting case of costless migration the welfare difference between the two tax regimes is equal to zero , as argued above.

Figure 1 indicates that there is a non-monotonic relationship between the cost of migration and the difference in welfare levels between the two tax regimes. Starting from the costless-migration equilibrium ( $\delta = 0$ ), as the cost of migration increases, the difference between the welfare levels associated with the two tax regimes increases (that is, the gain associated with shifting from a flat system to a non-linear one is rising). This pattern is maintained over some range up to some critical (sufficiently high) level of migration costs (incidentally, the level at which the incentive constraint of the high-skill individual starts to bind; see the characterization in proposition 1). The increasing pattern reverses itself, for values of $\delta$ higher than the critical level (the case of autarky is captured by sufficiently high levels of migration costs).

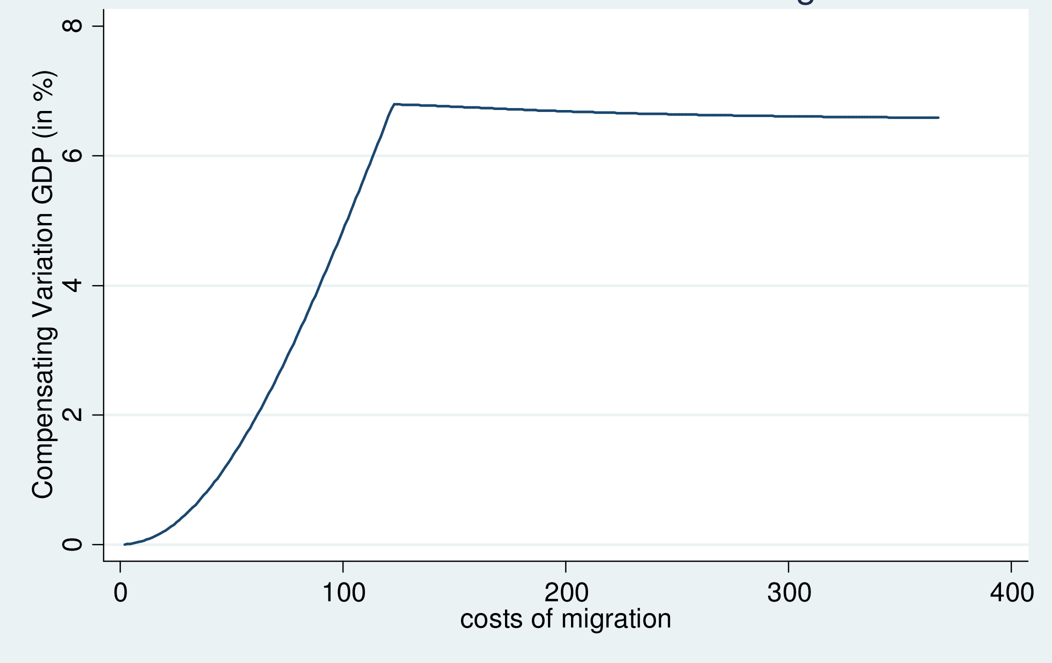



Figure 1: Comparison between the Non-linear and Linear Tax Regimes

The see the intuition underlying the non-monotonic relationship exhibited by figure 1, consider first the case where migration costs are sufficiently small in which none of the incentive constraints binds. Starting from the costless-migration case (in which no re-distribution takes place), as migration costs increase, the extent or re-distribution expands (due to the mitigated threat of high-skill migration) under both tax regimes. However, whereas re-distribution is being carried out through a (distortion-free) system of differential lump-sum transfers and taxes under the non-linear regime, attaining enhanced re-distribution via a linear system implies an increase in the (flat) marginal tax rate; hence, in the magnitude of the labor-leisure distortion entailed. Thus, as migration costs increase, the re-distributive advantage of the (efficient) non-linear system over the (distortive) linear regime becomes more manifest. Now consider the case where migration costs are sufficiently large, in which the incentive constraint of the high skill individual binds. Similar to the case of low migration costs, as migration costs increase, the extent of re-distribution expands under both tax regimes. However, whereas efficiency at the top (zero marginal tax rate levied on the high-skill individual) is maintained, in order to ensure no mimicking by the high-skill individual (effective means-testing), as migration costs increase, the government has to raise the marginal tax rate imposed on the low-skill individuals. The entailed distortion at the bottom limits the gains from enhanced re-distribution under the non-linear system; hence, the re-distributive advantage of the non-linear system relative to the linear regime. When the distortion entailed at the bottom is large enough, the patterns reverse and the welfare difference between the two tax regimes decreases as migration costs rise. When the distortion at the bottom is small in magnitude (for instance, in the case where low/high-skill wage ratio is sufficiently high) the welfare difference between the two regimes will rise monotonically over the entire range of migration costs, but at a decreasing rate over the range in which the incentive constraint is binding (see Figure G1 in Appendix G). Our numerical analysis (see Appendix G) shows that the patterns exhibited by figure 1 remain robust to changes in the other parameters of the model: the taxable-income elasticity (Figure G2) and the proportion of high-skill workers (Figure G3).

Figure 2 below depicts the relationship between the welfare-difference (on the vertical axis) measured, as in the previous figure, in compensating-variation terms (as a fraction of the *laissez-faire* output); and, the extent of re-distribution (on the horizontal axis), measured as the increase, in percentage terms, of the net income (consumption) of the low-skill individuals under a linear tax-regime relative to the *laissez-faire* benchmark.

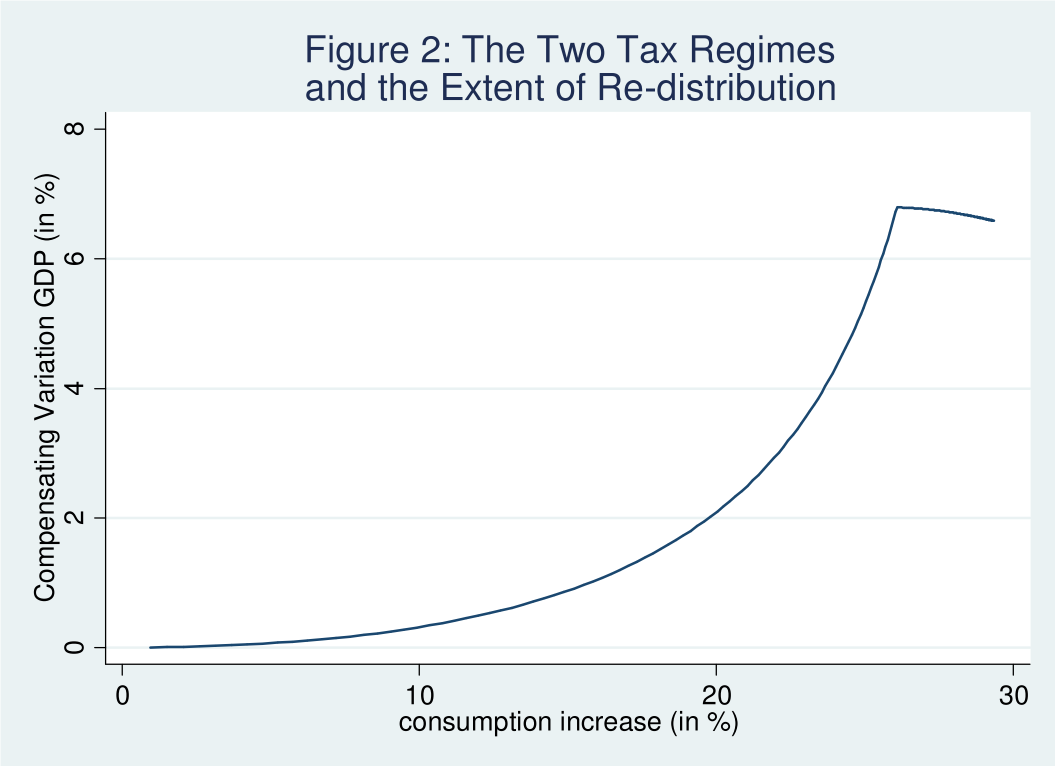


Figure 2: The Two Tax Regimes and the Extent of Re-distribution

Notice that the patterns are similar to those exhibited by figure 1. The figure indicates that for sufficiently small costs of migration, yet large enough to support a substantial amount of re-distribution, the welfare difference between the two-tax regimes is fairly small. For instance, when the linear tax regime attains a 14.8 percent increase in the level of consumption derived by the low-skill individuals relative to the *laissez-faire* equilibrium, the transfer (in consumption terms) that would be required to fully compensate the low-skill individuals for a shift from a non-linear to a linear regime, measured as a fraction of the total output in the *laissez-faire* equilibrium, is less than 1 percent. . When the linear regime attains an increase of 19.3 percent in the low-skill level of consumption relative to the level attained under the *laissez-faire* equilibrium, the welfare difference in compensating-variation terms (as a fraction of the *laissez-faire* output) is less than 2 percent.

To gain some perspective on the significant restraining impact tax competition bears on the extent of re-distribution in equilibrium (thereby on the welfare dominance of the non-linear tax regime), notice that when the linear regime attains an increase of 14.8 percent in low-skill consumption level (relative to the *laissez-faire* equilibrium) in which case the welfare difference between the two tax regimes (in compensating-variation terms) amounts to less than 1 percent of the *laissez-fair* output, the flat tax rate is given by 17.7 percent. In contrast, in the closed economy case (with no migration) the linear regime attains an increase of 30 percent in low-skill consumption (relative to the *laissez-fair* benchmark), the welfare difference between the two regimes (in compensating-variation terms) is 6.6 percent of the *laissez-fair* output, and the flat tax rate is over 50 percent.

Figure 2 illustrates that under plausible parametric assumptions the welfare gain associated with a shift from a linear system to a non-linear regime is fairly small, even when migration costs are sufficiently bounded away from zero to support a substantial amount of re-distribution in equilibrium. Notice that our result is in fact stronger than that inferred from the figure for several reasons. First, we assume a *Rawlsain* objective, which exhibits the strongest taste for re-distribution. Invoking a more moderate re-distributive objective is likely to enhance the restraining effect of tax competition on the entailed extent of re-distribution; thus, further narrowing the welfare difference between the two tax regimes. Second, allowing for migration of low-skill workers is likely to reduce the extent of re-distribution attained in equilibrium under both tax regimes, with similar implications to those driven by setting a less egalitarian objective. Third, in practice, linear systems often allow for an exemption level (an income threshold below which the individual does not pay any taxes), which enhances the extent of re-distribution attained. Finally, a non-linear system which is by construction means-tested (unlike the universal linear regime) is often mired by compliance issues (notably, misreporting by agents that claim eligibility for transfers) and by low take-up rates, both of which reduce the effective extent of re-distribution attained under the non-linear regime. Taking into account all the above considerations is likely to tilt the balance in favor of a linear system, in the presence of a sufficiently strong migration threat.

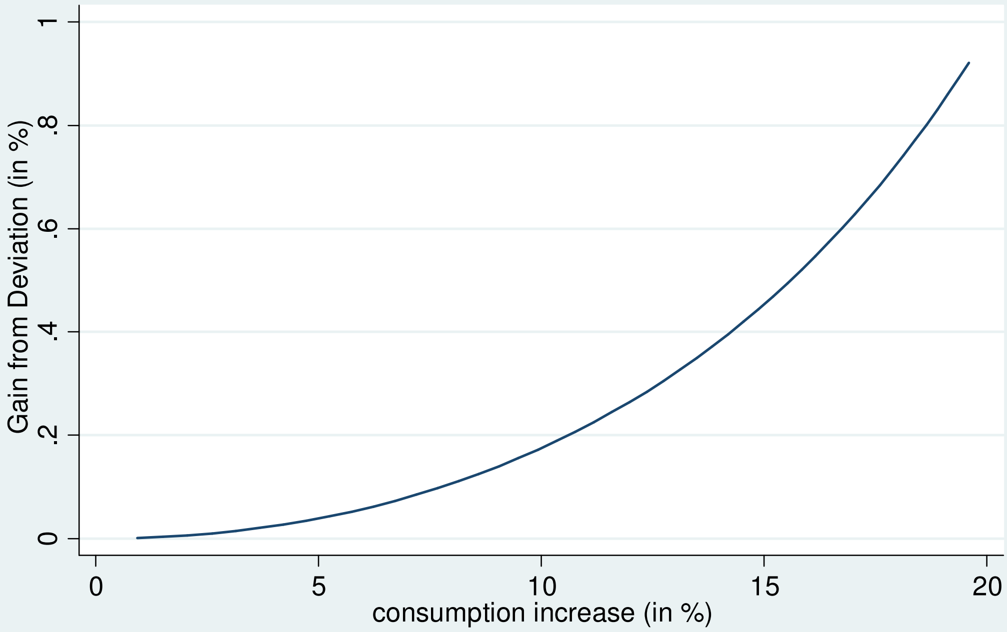
## 6. Stability of the Linear-Tax Equilibrium

In the previous section we demonstrated that under plausible parametric assumptions, the gain associated with a coordinated shift (by both countries) from a flat tax regime to a non-linear one may be fairly small when migration threat is strong. In this section we consider the case of an uncoordinated (unilateral) shift by one of the two countries. Starting from equilibrium in the case where both countries are restricted to linear-tax regimes, we examine the gain associated with a unilateral shift by one of the two countries to the optimal non-linear schedule. We maintain the same parametric assumption used in the previous calibrated simulations.

Figure 3 below depicts the relationship between the gain from a (unilateral) deviation from the linear-tax equilibrium to a non-linear schedule (on the vertical axis) measured, as in the previous figures, in compensating-variation terms (as a fraction of the *laissez-faire* output); and, the extent of re-distribution (on the horizontal axis), measured as the increase, in percentage terms, of the net income (consumption) of the low-skill individuals under a linear tax-regime relative to the *laissez-faire* benchmark.

Figure 3: Stability of the Linear Equilibrium

When the Difference in CV is Less than 2%




As can be seen from the figure, even when the extent of re-distribution is fairly significant (20 percent increase in the low-skill consumption level under the linear regime relative to the *laissez-faire* benchmark) the gain from deviation amounts to less than 1 percent of *laissez-faire* output. This modest gain is likely to be more than offset by the additional administrative costs associated with a shift from a flat regime to a non-linear one. Thus, taking into account the administrative costs, the figure essentially illustrates the stability of the linear equilibrium. Put differently, even in the case where both countries are free to choose a general non-linear tax schedule (no coordination is imposed), an equilibrium where both countries optimally choose to set a flat system in place is likely to occur. As in the previous section, the crucial factor at play is the threat of high-skill migration that significantly reduces the gain from deviation.

## 7. Conclusions

Linear (flat) tax systems are commonly perceived to be much simpler and hence cheaper to administer than non-linear ones. Much of the criticism against a reform aiming at 'flattening' the tax system (say, by consolidating tax brackets and/or income sources), despite its well-known entailed administrative gains, dwells on its perceived limited re-distributive capacity. A linear system accords a universal demo-grant across the board and, thus, fails to employ screening devices (notably, means-testing) to enhance the target-efficiency of the tax-transfer system. In this paper we employ a tax competition model to demonstrate that in the backdrop of a high-skill migration threat the re-distributive advantage of a non-linear system is significantly mitigated. In the presence of migration and in sharp contrast to the autarky case, a coordinated shift to a flat system (with its entailed administrative advantages), may be warranted. Furthermore, we demonstrate that an equilibrium where both countries, free to choose a non-linear tax system, set a flat system in place, is likely to form when migration threat is sufficiently strong.

## Appendix A: Proof of Proposition 1

We turn first to show that when migration costs are sufficiently large, the high-skill self-selection constraint is binding. Formally, there exists some threshold level of migration costs, $\delta_{critical}$, such that for all $\delta > \delta_{critical}$, the high-skill self-selection constraint is binding. Let $\delta_{critical} = \alpha_1 \left( y^*_{12} - y^*_{11} - h\left( y^*_{12} / w_2 \right) + h\left( y^*_{11} / w_2 \right) \right)$, where $y^*_{1j}$ denotes the *laissez-faire* income level derived by an individual with skill-level *j*=1,2, given by the implicit solution to: $h'(y^*_{1j} / w_j) = w_j$ . Notice that the threshold level is well defined (positive) by construction. Suppose by negation that for some $\delta > \delta_{critical}$ the high-skill self-selection constraint does not bind. As we re-distribute towards the low-skill individuals, it obviously cannot be the case that the low-skill self-selection constraint is binding. Consider then the case where both self-selection constraints do not bind. In this case the government problem is given by the following *Lagrangean*:

(A1) $$L \equiv V(w_1, c_{11}, y_{11}) + \mu \left[ \sum_j \alpha_{1j} \cdot (y_{1j} - c_{1j}) \right].$$

Formulating the first-order conditions yields:

(A2) $$1 - \mu \cdot \alpha_{11} = 0,$$

(A3) $$\frac{\partial V_{11}}{\partial y_{11}} + \mu \cdot \alpha_{11} = 0,$$

(A4) $$\mu[-\alpha_{12} + \frac{\alpha_2}{\delta} \cdot (y_{12} - c_{12})] = 0,$$

(A5) $$\mu[\alpha_{12} + \frac{\alpha_2}{\delta} \cdot \frac{\partial V_{12}}{\partial y_{12}} \cdot (y_{12} - c_{12})] = 0.$$

Notice that as both self-selection constraints do not bind (by our presumption) individuals of both skills set their income at the *laissez-faire* levels. Employing the symmetry property (as we characterize a symmetric equilibrium) and re-arranging the condition in (A4) yields:

(A6) $$(y_{12} - c_{12}) = \frac{\delta}{2}.$$

Substituting into the government revenue constraint yields:

(A7) $$c_{11} = \delta\left(\frac{\alpha_2}{2\alpha_1}\right) + y_{11}.$$

The high-skill self-selection constraint is, by assumption, satisfied as a strict inequality, and given by:

(A8) $$c_{12} - c_{11} - h\left(y_{12} / w_2\right) + h\left(y_{11} / w_2\right) > 0.$$

Substituting for $c_{11}$ and $c_{12}$ from (A6) and (A7) into (A8) yields:

(A9) $$\begin{aligned} & y_{12} - \frac{\delta}{2} - \delta\left(\frac{\alpha_2}{2\alpha_1}\right) - y_{11} - h\left(y_{12} / w_2\right) + h\left(y_{11} / w_2\right) > 0 \\ & \Leftrightarrow \delta < \alpha_1\left(y^*_{12} - y^*_{11} - h\left(y^*_{12} / w_2\right) + h\left(y^*_{11} / w_2\right)\right) = \delta_{critical}, \end{aligned}$$

where the last inequality follows from our previous observation that $y_{1j} = y^*_{1j}$. We thus obtain a contradiction to the presumption that $\delta_{critical} < \delta$.

We turn next to show that for $0 < \delta < \delta_{critical}$ both self-selection constraints are not binding. By the single-crossing property it cannot be the case that both constraints are binding. Moreover, as we re-distribute towards the low-skill individuals, it cannot be the case that only the low-skill incentive constraint is binding. We thus assume by negation, that only the high-skill self-selection constraint is binding.

The government problem is given by the following *Lagrangean*:

(A10) $L \equiv V(w_1, c_{11}, y_{11}) + \lambda\left[V(w_2, c_{12}, y_{12}) - V(w_2, c_{11}, y_{11})\right] + \mu\left[\sum_j \alpha_{1j} \cdot (y_{1j} - c_{1j})\right],$

Formulating the first-order conditions yields:

(A11) $1 - \lambda - \mu \cdot \dfrac{\alpha_1}{2} = 0,$

(A12) $\dfrac{\partial V_{11}}{\partial y_{11}} - \lambda \dfrac{\partial V_{12}}{\partial y_{11}} + \mu \cdot \dfrac{\alpha_1}{2} = 0,$

(A13) $\lambda + \mu[-\alpha_{12} + \dfrac{\alpha_2}{\delta} \cdot (y_{12} - c_{12})] = 0,$

(A14) $\lambda \cdot \dfrac{\partial V_{12}}{\partial y_{12}} + \mu[\alpha_{12} + \dfrac{\alpha_2}{\delta} \cdot \dfrac{\partial V_{12}}{\partial y_{12}} \cdot (y_{12} - c_{12})] = 0.$

Re-arranging (A13) yields:

(A15) $(y_{12} - c_{12}) = \left(\dfrac{\alpha_2}{2} - \dfrac{\lambda}{\mu}\right) \cdot \dfrac{\delta}{\alpha_2} < \dfrac{\delta}{2}.$

By virtue of the government revenue constraint, it follows that:

(A16) $(y_{12} - c_{12}) = -\dfrac{\alpha_1}{\alpha_2}(y_{11} - c_{11}).$

Substituting from (A16) into (A15) yields:

(A17) $c_{11} < y_{11} + \dfrac{\delta \alpha_2}{2\alpha_1}.$

By virtue of (A16) and (A17), it follows that,

(A18) $c_{11} - c_{12} < y_{11} - y_{12} + \dfrac{\delta}{\alpha_1}$ .

Substituting for the term $c_{11} - c_{12}$ from (A18) into the low-skill (non-binding) self-selection constraint and re-arranging yields:

(A19) $\delta > \alpha_1 \cdot \left[ y_{12} - h\left( y_{12} / w_1 \right) - y_{11} + h\left( y_{11} / w_1 \right) \right].$

Now let $F(w) \equiv h\left( y_{11} / w \right) - h\left( y_{12} / w \right)$. As $y_{12} - y_{11} > 0$ and by virtue of the convexity of $h$, it follows that $\frac{\partial F}{\partial w} < 0$. Thus,

(A20) $\delta > \alpha_1 \cdot \left[ y_{12} - h\left( y_{12} / w_2 \right) - y_{11} + h\left( y_{11} / w_2 \right) \right].$

By virtue of the efficiency-at-the-top property, $y_{12} = y_{12}^*$. Moreover, as $y_{11} < y_{11}^* < y_{12}^*$ it follows that $y_{11} - h\left( y_{11} / w_2 \right) < y_{11}^* - h\left( y_{11}^* / w_2 \right)$. It therefore follows that:

(A21) $\delta > \alpha_1 \cdot \left[ y_{12}^* - h\left( y_{12}^* / w_2 \right) - y_{11}^* + h\left( y_{11}^* / w_2 \right) \right].$

Thus, we obtain a contradiction to our presumption that $\delta < \delta_{critical}$.

## Appendix B: Proof of Proposition 2

We consider first the case where migration costs are sufficiently large, so that, by virtue of proposition 1, the self-selection constraint associated with the high-skill individuals is binding. The *Largangean* for the government problem is given by:

(B1) $$L \equiv V(w_1, c_{11}, y_{11}) + \lambda\left[V(w_2, c_{12}, y_{12}) - V(w_2, c_{11}, y_{11})\right] + \mu\left[\sum_j \alpha_{1j} \cdot (y_{1j} - c_{1j})\right].$$

Formulating the first-order conditions yields:

(B2) $$1 - \lambda - \mu \cdot \frac{\alpha_1}{2} = 0,$$

(B3) $$\frac{\partial V_{11}}{\partial y_{11}} - \lambda \frac{\partial V_{12}}{\partial y_{11}} + \mu \cdot \frac{\alpha_1}{2} = 0,$$

(B4) $$\lambda + \mu[-\alpha_{12} + \frac{\alpha_2}{\delta} \cdot (y_{12} - c_{12})] = 0,$$

(B5) $$\lambda \cdot \frac{\partial V_{12}}{\partial y_{12}} + \mu[\alpha_{12} + \frac{\alpha_2}{\delta} \cdot \frac{\partial V_{12}}{\partial y_{12}} \cdot (y_{12} - c_{12})] = 0.$$

Substituting from (B2) into (B3) yields:

(B6) $$H \equiv \frac{\partial V_{11}}{\partial y_{11}} - \lambda\left(\frac{\partial V_{12}}{\partial y_{11}} + 1\right) + 1 = 0.$$

Employing (B2), (B4) and the government revenue constraint, following some algebraic manipulations and re-arranging, yields:

(B7) $$c_{11} = y_{11} + \frac{\delta}{\alpha_1} \cdot \left(1 - \frac{\alpha_1}{2(1-\lambda)}\right),$$

(B8) $$c_{12} = y_{12} + \frac{\delta}{\alpha_2} \cdot \left(\frac{\alpha_1}{2(1-\lambda)} - 1\right).$$

Substituting (B7) and (B8) into the high-skill incentive constraint and re-arranging, yields:

(B9) $$J \equiv \left(\frac{2\delta}{\alpha_1\alpha_2}\right)\left(\frac{\alpha_1}{2(1-\lambda)}-1\right)+y_{12}-h(y_{12}/w_2)-y_{11}+h(y_{11}/w_2)=0.$$

By virtue of the efficiency at the top, $y_{12}=y_{12}^*$, the level of income chosen under a *laissez-faire* regime. Thus, equilibrium is given by the solution to the system of two equations (B6) and (B9) solved for the two unknowns: $y_{11}$ and $\lambda$. From (B**6**) it follows that:

(B10) $$\frac{\partial y_{11}}{\partial\lambda}\Big|_{H=const}=-\frac{\frac{\partial H}{\partial\lambda}}{\frac{\partial H}{\partial y_{11}}}=\frac{\frac{\partial V_{12}}{\partial y_{11}}+1}{\frac{\partial^2 V_{11}}{\partial y_{11}^2}-\lambda\frac{\partial^2 V_{12}}{\partial y_{11}^2}}.$$

Recalling that $\frac{\partial V}{\partial y}=-h'(y/w)\cdot(1/w)$ and $y_{11}<y_{11}^*<y_{12}^*$, where $y_{1j}^*$ denotes the level of income chosen by a '*j*-type' individual under *laissez-faire*, it follows by the convexity of *h* that $0<\frac{\partial V_{11}}{\partial y_{11}}+1<\frac{\partial V_{12}}{\partial y_{11}}+1$. Thus, by virtue of (B6), it follows that $\lambda<1$. Note that $\frac{\partial^2 V_{11}}{\partial y_{11}^2}-\lambda\frac{\partial^2 V_{12}}{\partial y_{11}^2}=-\left(h''\left(\frac{y_{11}}{w_1}\right)\cdot(1/w_1^2)-\lambda h''\left(\frac{y_{11}}{w_2}\right)\cdot(1/w_2^2)\right)<0$ , by virtue of the assumption that $h'''\geq 0$. Thus, by virtue of (B10) it follows that $\frac{\partial y_{11}}{\partial\lambda}\Big|_{H=const}<0$.

Denote the solutions to (B6) and (B9) by $f_1(\lambda)$ and $f_2(\lambda,\delta)$, respectively. Note that as the expression in (B9) is strictly concave in $y_{11}$, there are potentially two implicit solutions, but as the optimal tax schedule implies that the marginal tax rate levied on the low-skill individual is strictly positive, the only feasible solution is the smaller value of the two candidate solutions.

Now let $f(\lambda,\delta) \equiv f_1(\lambda) - f_2(\lambda,\delta)$. We next show that $f(0,\delta) > 0$ and $f[\alpha_2/2,\delta] < 0$.

We turn first to establish that $f(0,\delta) > 0$. To see this note first that by virtue of (B6), setting $\lambda = 0$ implies $\frac{\partial V_{11}}{\partial y_{11}} = -1$; hence, $f_1(0) = y_{11}^*$. It suffices to show that $f_2(0,\delta) < y_{11}^*$. To see this, note that by virtue of (B9), setting $\lambda = 0$ implies that $-\frac{\delta}{\alpha_1} + y_{12} - y_{11} - h(y_{12}/w_2) + h(y_{11}/w_2) = 0$.

Thus, $\delta = \alpha_1 \cdot \left( y_{12} - h(y_{12}/w_2) - y_{11} + h(y_{11}/w_2) \right)$.

Recalling that $\delta_{critical} < \delta$ (where the critical delta is defined in the proof of proposition 1) implies:

$$\alpha_1 \left( y_{12}^* - h(y_{12}^*/w_2) - y_{11}^* + h(y_{11}^*/w_2) \right) < \alpha_1 \left( y_{12}^* - h(y_{12}^*/w_2) - y_{11} + h(y_{11}/w_2) \right)$$
$$\Leftrightarrow y_{11} - h(y_{11}/w_2) < y_{11}^* - h(y_{11}^*/w_2) \Leftrightarrow y_{11} < y_{11}^*$$

We turn next to establish that $f[\alpha_2/2,\delta] < 0$. To see this note first that by virtue of (B6), setting $\lambda = \alpha_2/2$ and recalling that $f_1(0) = y_{11}^*$ implies, by virtue of the fact that $\frac{\partial y_{11}}{\partial \lambda}\Big|_{H=const} < 0$, that $f_1[\alpha_2/2] < y_{11}^*$. Substituting $\lambda = \alpha_2/2$ into (B9) implies that $y_{12}^* - h(y_{12}^*/w_2) - y_{11} + h(y_{11}/w_2) = 0 \Leftrightarrow f_2(\alpha_2/2,\delta) = y_{12}^*$. Thus, $f[\alpha_2/2,\delta] < 0$.

By the continuity of $f$ it follows, by virtue of the intermediate value theorem, that there exists some $\lambda' \in (0,\alpha_2/2)$, such that $f(\lambda') = 0$. Moreover, it follows that $f_1(\lambda') < y_1^*$. Thus, we have proved existence and the solution is well defined (satisfies the condition that the marginal tax rate on the low-skill individual is positive at the optimum).

We turn next to prove uniqueness. Differentiating the expression in (B9) with respect to $\lambda$ and $y_{11}$ yields:

$$\frac{\partial J}{\partial \lambda} = \left(\frac{2}{\alpha_1 \alpha_2}\right)\left(\frac{2\alpha_1}{4(1-\lambda)^2}\right) > 0,$$
$$\frac{\partial J}{\partial y_{11}} = -\left[1 - h'(y_{11}/w_2)\cdot\frac{1}{w_2}\right] < 0,$$

where the last inequality follows from $y_{11} < y_{11}^* < y_{12}^*$ and the convexity of *h*.

Thus,

(B11) $$\frac{\partial y_{11}}{\partial \lambda}\Big|_{J=const} = -\frac{\frac{\partial J}{\partial \lambda}}{\frac{\partial J}{\partial y_{11}}} > 0.$$

Uniqueness follows, as $f_1$ is decreasing and $f_2$ is increasing. Hence, both schedules intersect only once.

We consider next the case where migration costs are small enough, so that both self-selection constraints do not bind. Formulating the *Lagrangean* for this case yields:

(B12) $$L \equiv V(w_1, c_{11}, y_{11}) + \mu\left[\alpha_{12}\cdot(y_{1j} - c_{1j}) + \alpha_{11}\cdot(y_{11} - c_{11})\right],$$

where $\mu$ denotes the *Lagrange* multiplier associated with the government budget constraint. The first-order conditions are given by:

(B13) $$1 - \mu\cdot\alpha_{11} = 0,$$

(B14) $$\frac{\partial V_{11}}{\partial y_{11}} + \mu\cdot\alpha_{11} = 0,$$

(B15) $\mu[-\alpha_{12} + \frac{\alpha_2}{\delta} \cdot (y_{12} - c_{12})] = 0$,

(B16) $\mu[\alpha_{12} + \frac{\alpha_2}{\delta} \cdot \frac{\partial V_{12}}{\partial y_{12}} \cdot (y_{12} - c_{12})] = 0$.

By substituting from (B13) into (B14) and re-arranging, one obtains:

(B17) $\partial V_{11} / \partial y_{11} = 1$.

By substituting from (B15) into (B16) and re-arranging, one obtains:

(B18) $\partial V_{12} / \partial y_{12} = 1$.

Thus, the gross income levels chosen by both the high-skill and the low-skill individuals, denoted by $y^*_{1j}$; $j$=1,2, are the efficient *laissez-faire* ones, given by the implicit solution to: $\partial V_{1j} / \partial y_{1j} = 1$.

Substituting for $\alpha_{12} = \alpha_2 / 2$ into (B15), by virtue of the construction of a symmetric equilibrium, and re-arranging, yields:

(B19) $c_{12} = y^*_{12} - \frac{\delta}{2}$,

By substituting for the term $y^*_{12} - c_{12}$ from (B19) into the government (binding) revenue constraint in (7) and re-arranging, one obtains:

(B20) $c_{11} = y^*_{11} + \frac{\delta\alpha_2}{2\alpha_1}$.

The symmetric equilibrium is uniquely defined by the 8-tuple: $y^*_{ij}$ and $c_{ij}$, $i, j = 1,2$, with $y^*_{1j} = y^*_{2j}$ and $c_{1j} = c_{2j}$, where $y^*_{ij}$ denotes the *laissez-faire* gross income level derived by an individual with skill-level $j$=1,2 in country $i$=1,2; and $c_{ij}$ denotes the net income level derived by an individual with skill-level $j$=1,2 in country $i$=1,2, given by the expressions on right-hand side of (B19) and (B20).

Notice that when $\delta=0;$ namely, in the case of costless migration, the equilibrium naturally converges to the *laissez-faire* allocation, given by: $c_{1j} = y^*_{1j}, \; j = 1,2$.

## Appendix C: Non-existence of Asymmetric Equilibria

In what follows we prove that when migration costs are sufficiently small; hence, the incentive constraint of the high-skill individuals is not binding, there exists no a-symmetric Nash equilibrium for the tax competition game between the two countries.

Formulating the first-order conditions for the program solved by country *i*, *i*=1,2, yields:

$$(C1) \quad 1+\mu_i[-\alpha_{i1}+\alpha_1/\delta\cdot(y_{i1}-c_{i1})]=0,$$

$$(C2) \quad \partial V_{i1}/\partial y_{i1}+\mu_i\cdot\alpha_{i1}=0,$$

$$(C3) \quad \mu_i\cdot[-\alpha_{i2}+\alpha_2/\delta\cdot(y_{i2}-c_{i2})]=0,$$

$$\text{(C4)} \quad \mu[\alpha_{i2}+\alpha_2/\delta\cdot\partial V_{i2}/\partial y_{i2}\cdot(y_{i2}-c_{i2})]=0.$$

Substituting (C3) into (C4) and re-arranging, yields:

$$\text{(C5)} \quad y_{i2}=y_2^*,$$

where $y_2^*$ denotes the *laissez-faire* gross level of income chosen by type-2 individual.

By virtue of (C5) it follows that the difference between the utility levels derived by a high-skill individual residing in countries *i* and *j*, respectively, is given by:

$$\text{(C6)} \quad V(w_2,c_{i2},y_{i2})-V(w_2,c_{j2},y_{j2})=c_{i2}-c_{j2}.$$

Substituting from (C6) into the migration condition given in (8), in the main text, yields:

$$\text{(C7)} \quad \alpha_{i2}=\alpha_2\cdot\left[1/2+1/\delta\cdot\left(c_{i2}-c_{j2}\right)\right].$$

In an asymmetric equilibrium $\alpha_{i2} \neq \alpha_{j2}$; hence, with no loss of generality, we henceforth assume that, $\alpha_{12} > \alpha_{22}$. By virtue of (C7) it follows that $c_{12} > c_{22}$.

A necessary condition for a *Nash* equilibrium to exist is that either one of the two countries cannot attain a fiscal surplus by slightly modifying the net income (consumption) level derived by a typical high-skill individual (leaving all other tax parameters unchanged). Notice that by slightly increasing or decreasing the net income level derived by the high-skill individuals, none of the two incentive constraints is violated, as both are satisfied as strict inequalities when migration costs are small enough. Differentiating the revenue constraint (in countries 1 and 2, respectively) with respect to the corresponding net income (consumption) level derived by the high-skill individual yields the following two conditions that necessarily hold in equilibrium:

(C8) $$\frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{12}\right) - \alpha_{12} = 0 \Leftrightarrow \frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{12}\right) = \alpha_{12},$$

(C9) $$\frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{22}\right) - \alpha_{22} = 0 \Leftrightarrow \frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{22}\right) = \alpha_{22}.$$

By subtracting (C9) from (C8) and re-arranging, it follows that:

(C10) $$\frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{12}\right) - \frac{\alpha_2}{\delta}\cdot\left(y_2^* - c_{22}\right) = \alpha_{12} - \alpha_{22} \to \frac{\alpha_2}{\delta}\left(c_{22} - c_{12}\right) = \alpha_{12} - \alpha_{22}$$

From (C7) it also follows that:

(C11) $$\alpha_{12} - \alpha_{22} = 1/\delta\cdot\left(c_{12} - c_{22}\right) - 1/\delta\cdot\left(c_{22} - c_{12}\right) = 2/\delta\cdot\left(c_{12} - c_{22}\right)$$

By substituting from (C11) into (C10) it follows that:

(C12) $$\frac{\alpha_2}{\delta}\left(c_{22} - c_{12}\right) = 2/\delta\cdot\left(c_{12} - c_{22}\right) \Leftrightarrow \left(2 + \alpha_2\right)\left(c_{12} - c_{22}\right) = 0$$

Thus, we obtain a contradiction to our presumption that $c_{12} > c_{22}$. This completes the proof.

## Appendix D: Proof of Proposition 3

We consider first the case where migration costs are sufficiently large; hence, the incentive constraint of the high-skill individual is binding. We first prove that in equilibrium, the marginal tax rate levied on the low-skill individuals increases with respect to $\delta$. As shown in Appendix B, the unique equilibrium for the game between the two countries is given by the (unique feasible) solution to the following system of two equations:

(D1) $$H\left(y_{11},\lambda\right)\equiv\frac{\partial V_{11}}{\partial y_{11}}-\lambda\left(\frac{\partial V_{12}}{\partial y_{11}}+1\right)+1=0,$$

(D2) $$J\left(y_{11},\lambda,\delta\right)\equiv\left(\frac{2}{\alpha_1\alpha_2}\right)\left(\frac{\alpha_1}{2(1-\lambda)}-1\right)\delta+y^*_{12}-h(w_2,y^*_{12})-\left[y_{11}-h(w_2,y_{11})\right]=0,$$

where $y^*_{12}$ denotes the *laissez-faire* income level associated with the high-skill individual (efficiency at the top property).

Fully differentiating (D1) and (D2) with respect to $\delta$ yields:

(D3) $$\frac{\partial H}{\partial\lambda}\cdot\frac{\partial\lambda}{\partial\delta}+\frac{\partial H}{\partial y_{11}}\cdot\frac{\partial y_{11}}{\partial\delta}=0,$$

(D4) $$\frac{\partial J}{\partial\lambda}\cdot\frac{\partial\lambda}{\partial\delta}+\frac{\partial J}{\partial y_{11}}\cdot\frac{\partial y_{11}}{\partial\delta}+\frac{\partial J}{\partial\delta}=0.$$

Using Cramer's rule one obtains:

(D5) $$\frac{\partial y_{11}}{\partial\delta}=-\frac{(\partial H/\partial\lambda)\cdot(\partial J/\partial\delta)}{(\partial H/\partial\lambda)\cdot(\partial J/\partial y_{11})-(\partial H/\partial y_{11})\cdot(\partial J/\partial\lambda)}.$$

We turn next to sign the expression on the right-hand side of (D5). Note first, that by virtue of our earlier derivations (see Appendix B), $\frac{\partial H}{\partial \lambda} < 0$. Moreover, efficiency at the top implies that $y^*_{12} - h(w_2, y^*_{12}) > y_{11} - h(w_2, y_{11})$. Thus, by virtue of (D2), it follows that $\left(\frac{2}{\alpha_1 \alpha_2}\right)\left(\frac{\alpha_1}{2(1-\lambda)} - 1\right) < 0$, hence, $\frac{\partial J}{\partial \delta} < 0$. We conclude that the expression in the numerator on the right-hand side of (D5) is positive. Turning next to the expression on the denominator on the right-hand side of (D5), it follows, by virtue of our earlier derivations (see Appendix B), that $\frac{\partial J}{\partial \lambda} > 0$, $\frac{\partial J}{\partial y_{11}} < 0$ and $\frac{\partial H}{\partial y_{11}} < 0$. Thus, the expression in the denominator on the right-hand side of (D5) is positive. We conclude that $\frac{\partial y_{11}}{\partial \delta}$<0.

The (implicit) marginal tax rate levied on the low-skill individual is given by:

(D6) $$MRT_1 = 1 + \frac{\partial V_{11} / \partial y_{11}}{\partial V_{11} / \partial c_{11}} = 1 + \partial V_{11} / \partial y_{11}.$$

Differentiation with respect to $\delta$ yields:

(D7) $$\frac{\partial MRT_1}{\partial \delta} = (\partial^2 V_{11} / \partial y_{11}{}^2) \cdot (\partial y_{11} / \partial \delta) > 0,$$

where the last inequality follows from (D5) and the convexity of *h*.

Thus, indeed, as $\delta$ decreases, the marginal tax rate levied on the low-skill individuals decreases.

We turn next to prove that when $\delta$ decreases the net transfers received by the low-skill individuals decrease (correspondingly, by virtue of the balanced budget constraint of the government, the net taxes paid by the high-skill individuals decrease as well). Suppose by way of contradiction that as $\delta$ decreases, the term $(c_{11} - y_{11})$ weakly increases. Formally,

(D8) $$\frac{\partial(c_{11} - y_{11})}{\partial\delta} \leq 0.$$

To satisfy the government budget constraint [given in (7)] it necessary follows that,

(D9) $$\frac{\partial(y_{12} - c_{12})}{\partial\delta} \leq 0.$$

By virtue of the 'efficiency at the top' property, the gross income received by the high-skill individual does not change in response to the decrease in $\delta$. Hence,

(D10) $$\frac{\partial y_{12}}{\partial\delta} = 0.$$

Combining (D9) and (D10) implies that,

(D11) $$\frac{\partial[c_{12} - h(y_{12}/w_2)]}{\partial\delta} \geq 0.$$

By virtue of the binding self-selection constraint [given in (7)] it follows that,

(D12) $$\frac{\partial[c_{11} - h(y_{11}/w_2)]}{\partial\delta} \geq 0 \Leftrightarrow \frac{\partial c_{11}}{\partial\delta} \geq h'(y_{11}/w_2)\cdot(1/w_2)\cdot\frac{\partial y_{11}}{\partial\delta}$$
$$\Leftrightarrow \frac{\partial(c_{11} - y_{11})}{\partial\delta} \geq \frac{\partial y_{11}}{\partial\delta}\cdot[h'(y_{11}/w_2)\cdot(1/w_2) - 1] > 0,$$

where the last inequality follows from (D5) and the fact that $y_{11} < y_{11}^* < y_{12}^*$, where $y_{1j}^*$, $j$=1,2, denotes the *laissez-faire* level of income chosen by a $j$-type individual.

Comparing (D8) and (D12), one obtains the desired contradiction.

Finally, we turn to prove that the utility level of the high-skill individuals (respectively, that of the low-skill individuals) is decreasing (increasing) with respect to $\delta$. We first consider the high-skill individuals. Differentiation with respect to $\delta$ yields:

(D13) $$\frac{\partial V_{12}}{\partial \delta} = \partial c_{12} / \partial \delta - h'(y_{12} / w_2) \cdot (1/ w_2) \cdot \partial y_{12} / \partial \delta .$$

As shown above, $\partial y_{12} / \partial \delta = 0$ and $\partial (y_{12} - c_{12}) / \partial \delta > 0$; hence, $\frac{\partial V_{12}}{\partial \delta} < 0$.

We turn next to the low-skill individuals. With slight abuse of notation, denote by $V_{1j}(\delta)$ the utility derived by an individual of skill-level *j* in country *i*=1 (and hence in country *i*=2) in equilibrium as a function of $\delta$. Further denote by $c_{1j}(\delta)$ and $y_{1j}(\delta)$, the corresponding net income and gross income levels of an individual of skill level *j* in country *i*=1. Now suppose by way of contradiction, that as $\delta$ increases, the utility derived by a low-skill individual in equilibrium is weakly decreasing. Formally,

(D14) $$\frac{\partial V_{11}}{\partial \delta} = \partial c_{11} / \partial \delta - h'(y_{11} / w_1) \cdot (1/ w_1) \cdot \partial y_{11} / \partial \delta \leq 0 .$$

We turn to show that if the condition in (D14) holds, then choosing the bundles $(c_{1j}(\delta), y_{1j}(\delta)); j = 1,2$, is not a best-response for country *i*=1. To see this, fix some arbitrary $\delta$, and consider a deviation to an alternative tax schedule given by $(c_{1j}(\delta'), y_{1j}(\delta')); j = 1,2$, where $\delta - \delta' > 0$ and is arbitrarily small. Clearly, by construction, the self-selection constrains are satisfied under the new tax schedule. Moreover, the self-selection constraint of the high-skill individual [given in (7)] is satisfied as equality. We turn to show that such a deviation creates a fiscal surplus. Let $\Omega(\delta, \delta')$ denote the fiscal surplus of the government in country 1, when the migration costs are $\delta$, and the government in country 1 deviates to the alternative tax schedule,

$(c_{1j}(\delta'), y_{1j}(\delta')); j = 1,2$. Note that by construction, $\Omega(\delta,\delta) = 0$. Formally, we need to show that $\Omega(\delta,\delta') > 0$. Taking a first-order approximation, we need to show that:

(D15) $$\Omega(\delta,\delta') = \Omega(\delta,\delta) - (\delta - \delta') \cdot \left.\frac{\partial\Omega(\delta,\delta')}{\partial\delta'}\right|_{\delta'=\delta} = -(\delta - \delta') \cdot \left.\frac{\partial\Omega(\delta,\delta')}{\partial\delta'}\right|_{\delta'=\delta} > 0$$

Differentiating the budget constraint in (7), it suffices to show that:

(D16)

$$\left.\frac{\partial\Omega(\delta,\delta')}{\partial\delta'}\right|_{\delta'=\delta} = \left[\alpha_1 / 2 \cdot (\partial y_{11} / \partial\delta - \partial c_{11} / \partial\delta) + \alpha_2 / 2 \cdot (\partial y_{12} / \partial\delta - \partial c_{12} / \partial\delta)\right] + \left[\alpha_2 / \delta \cdot (\partial V_{12} / \partial\delta) \cdot (y_{12} - c_{12})\right] < 0.$$

Consider first the first term in brackets on the right-hand side of (D16). By virtue of the binding budget constraint (which holds for any $\delta$ in equilibrium) this term is equal to zero. Consider next the second term in brackets. By virtue of our earlier derivation $\frac{\partial V_{12}}{\partial\delta} < 0$, hence this term is negative. We have established, therefore, that some deviation from the best-response (by presumption) tax schedule results in a fiscal surplus. This surplus can be used to attain a Pareto improvement. We obtain the desired contradiction.

We turn next to the case where migration costs are low enough (hence, the incentive-constraint of the high-skill individuals does not bind). As shown in Appendix B [see conditions (B19) and (B20)], in this case, the net income (consumption) levels derived by the high-skill and low-skill individuals, respectively, are given by:

(D17) $$c_{12} = y^*_{12} - \frac{\delta}{2},$$

(D18) $$c_{11} = y^*_{11} + \frac{\delta\alpha_2}{2\alpha_1},$$

where the gross income levels chosen by both the high-skill and the low-skill individuals, denoted by $y^{*}_{1j}$; *j*=1,2, are the efficient *laissez-faire* ones, given by the implicit solution to: $\partial V_{1j} / \partial y_{1j} = 1$.

It directly follows from conditions (D17) and (D18) that the net transfers (net taxes) received (paid) by the low-skill (respectively, high skill) individuals increase with respect to mobility costs; and, correspondingly, the utility derived by the low-skill (respectively, high-skill) individuals is increasing (decreasing) with respect to mobility costs. This completes the proof.

## Appendix E: Proof of Proposition 4

In proving the proposition we will repeatedly use the following simple lemma:

**Lemma:** (i) $\partial y_{1j} / \partial t < 0$, (i) $\partial^2 y_{1j} / \partial t^2 \leq 0$.

**Proof:** Follows straightforward from differentiation of the individual first-order condition, $h'(y_{1j} / w) = w_j \cdot (1-t)$, with respect to *t*, employing the properties of the function *h* ($h''' \geq 0$ and $h'' > 0$).

Employing the first-order conditions in (17) and (18), the equilibrium for the game between the two countries is given by the solution to the following system of three equations (for the three unknowns, *t*, *T* and $\lambda$):

(E1) $$\frac{\partial V_{11}}{\partial t} + \lambda\left(\frac{\alpha_2}{\delta}\frac{\partial V_{12}}{\partial t} t y_{12} + \frac{\partial y_{11}}{\partial t} t\alpha_{11} + \alpha_{11} y_{11} + \frac{\partial y_{12}}{\partial t} t\alpha_{12} + \alpha_{12} y_{12} - \frac{\alpha_2}{\delta}\frac{\partial V_{12}}{\partial t} T\right) = 0,$$

(E2) $$1 + \lambda\left[\frac{\alpha_2}{\delta}\left(t y_{12} - T\right) - 1\right] = 0,$$

(E3) $$t \cdot (\alpha_{11} y_{11} + \alpha_{12} y_{12}) - \left(\alpha_{11} + \alpha_{12}\right) T = 0.$$

Let *A* denote the effect of an increase in *t* on the government budget constraint in a closed economy (that is, in the absence of migration). Formally,

(E4) $$A = \frac{\alpha_1}{2}\left(\frac{\partial y_{11}}{\partial t} t + y_{11}\right) + \frac{\alpha_2}{2}\left(\frac{\partial y_{12}}{\partial t} t + y_{12}\right).$$

Re-arranging the revenue constraint in (E3), employing symmetry, yields:

(E5) $$t y_{12} - T = \frac{\alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}{2}.$$

Substituting for *T* from (E5) into (E2) and re-arranging yields:

$$\text{(E6)} \quad \lambda = \frac{2\delta}{2\delta - \alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}.$$

Substituting from (E4)-(E6) into (E1) and re-arranging yields:

$$\text{(E7)} \quad F\left(t,\delta\right) \equiv \frac{\partial V_{11}}{\partial t} + \left(\frac{2\delta}{2\delta - \alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}\right)\left(\frac{\alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}{2\delta}\frac{\partial V_{12}}{\partial t} + A\right) = 0.$$

We turn first to prove existence.

Substituting $t$=0 into (E7), employing the individual optimization envelope condition and re-arranging, yields:

$$\text{(E8)} \quad F\left(0,\delta\right) = \frac{\alpha_2\left(y_{12}^* - y_{11}^*\right)}{2} > 0,$$

where $y_{1j}^*, j = 1,2$, denotes the *laissez-faire* level of income chosen by a *j*-type individual.

Substituting $t$=1 into (E7), noting that in this case, $y_{1j} = 0, j = 1,2$, it follows:

$$\text{(E9)} \quad F\left(1,\delta\right) = \frac{\alpha_1}{2} \cdot \left.\frac{\partial y_{11}}{\partial t}\right|_{t=1} + \frac{\alpha_2}{2} \cdot \left.\frac{\partial y_{12}}{\partial t}\right|_{t=1} < 0,$$

where the sign of the inequality follows from part (i) of the lemma.

Existence follows then by the continuity of *F* in *t*, employing the intermediate value theorem. The linear system is indeed progressive ($0<t<1$) as expected.

We turn next to prove uniqueness by showing that $\left.\frac{\partial F(t,\delta)}{\partial t}\right|_{F(t,\delta)=0} < 0$. Uniqueness will then follow by the continuity of *F* in *t*.

By differentiating (E6) with respect to $t$ and re-arranging, one obtains:

(E10) $$\frac{\partial \lambda}{\partial t} = \frac{2\delta \cdot \alpha_2 \cdot \alpha_1 \cdot \phi}{\left(2\delta - \alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)\right)^2} = \lambda^2 \cdot \frac{\alpha_2 \cdot \alpha_1 \cdot \phi}{2\delta},$$

where $\phi \equiv y_{12} + t \cdot \frac{\partial y_{12}}{\partial t} - \left(y_{11} + t \cdot \frac{\partial y_{11}}{\partial t}\right)$.

By manipulating (E7), employing (E4) and the envelope condition for the individual optimization problem and re-arranging, one obtains:

(E11) $$\phi = \frac{2}{\alpha_2} \cdot \left[(1/\lambda - 1) \cdot (y_{11} - y_{12}) - t \cdot \frac{\partial y_{11}}{\partial t}\right] > 0,$$

where the inequality follows, as by our earlier derivations, $\lambda > 1, y_{12} > y_{11}$ and $\partial y_{11} / \partial t < 0$.

By virtue of (E10), it follows $\frac{\partial \lambda}{\partial t}$>0.

By differentiating (E7) with respect to $t$, one obtains:

(E12) $$\left.\frac{\partial F(t,\delta)}{\partial t}\right|_{F(t,\delta)=0} = \frac{\partial^2 V_{11}}{\partial t^2} + \frac{\partial \lambda}{\partial t} \cdot \left(\frac{\alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}{2\delta} \cdot \frac{\partial V_{12}}{\partial t} + A\right) + \lambda \cdot \left(\frac{\alpha_2 \cdot \alpha_1 \cdot \phi}{2\delta} \cdot \frac{\partial V_{12}}{\partial t} + \frac{\alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}{2\delta} \cdot \frac{\partial^2 V_{12}}{\partial t^2} + \frac{\partial A}{\partial t}\right).$$

From (E7) it follows that:

(E13) $$\lambda \cdot \left(\frac{\alpha_2 \cdot \alpha_1 \cdot t \cdot \left(y_{12} - y_{11}\right)}{2\delta} \frac{\partial V_{12}}{\partial t} + A\right) = -\frac{\partial V_{11}}{\partial t}.$$

Substituting from (E13) and (E10) into (E12) and re-arranging, yields:

(E14)
$$\left.\frac{\partial F(t,\delta)}{\partial t}\right|_{F(t,\delta)=0} = \frac{\partial^2 V_{11}}{\partial t^2} + \lambda\cdot\left(\frac{\alpha_2\cdot\alpha_1\cdot\phi}{2\delta}\left(\frac{\partial V_{12}}{\partial t} - \frac{\partial V_{11}}{\partial t}\right) + \frac{\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12} - y_{11}\right)}{2\delta}\cdot\frac{\partial^2 V_{12}}{\partial t^2} + \frac{\partial A}{\partial t}\right).$$

Further re-arranging (E14), yields:

(E15)
$$\left.\frac{\partial F(t,\delta)}{\partial t}\right|_{F(t,\delta)=0} = \lambda\cdot\left(\frac{\alpha_2\cdot\alpha_1\left(y_{11} - y_{12}\right)\cdot}{2\delta}\left(\phi + \frac{\partial y_{11}}{\partial t}t\right) + \frac{\partial A}{\partial t} - \frac{\partial y_{11}}{\partial t}\cdot\frac{1}{\lambda}\right).$$

Notice that,

(E16)
$$\phi + \frac{\partial y_{11}}{\partial t}t = \frac{2}{\alpha_2}\cdot\left[(1/\lambda - 1)\cdot(y_{11} - y_{12})\right] - t\cdot\frac{\partial y_{11}}{\partial t}\left(\frac{2}{\alpha_2} - 1\right) > 0.$$

Hence, a sufficient condition for the expression on the right-hand side of (E15) to be negative is the following

(E17)
$$\frac{\partial A}{\partial t} - \frac{\partial y_{11}}{\partial t}\cdot\frac{1}{\lambda} < 0.$$

As $\lambda > 1$, it follows that:

(E18)
$$\frac{\partial A}{\partial t} - \frac{\partial y_{11}}{\partial t}\cdot\frac{1}{\lambda} < \frac{\partial A}{\partial t} - \frac{\partial y_{11}}{\partial t} = \frac{\alpha_1}{2}\frac{\partial^2 y_{11}}{\partial t^2}t + \frac{\alpha_2}{2}\frac{\partial^2 y_{12}}{\partial t^2}t + \frac{\alpha_2}{2}\left(\frac{\partial y_{12}}{\partial t} - \frac{\partial y_{11}}{\partial t}\right).$$

By part (ii) of the lemma it follows that $\partial^2 y_{1j} / \partial t^2 \leq 0$. Thus, a sufficient condition for the inequality in (E17) to hold is the following:

(E19)
$$\frac{\partial y_{12}}{\partial t} - \frac{\partial y_{11}}{\partial t} < 0.$$

A sufficient condition for the inequality in (E19) to be satisfied is the following:

(E20) $\frac{\partial^2 y}{\partial t \partial w} < 0$.

The individual first-order condition is given by:

(E21) $h'(y/w) = w \cdot (1-t)$.

Differentiation with respect to *w* then yields:

(E22) $h''(y/w) \cdot \left[ \frac{\frac{\partial y}{\partial w} \cdot w - y}{w^2} \right] = (1-t)$.

Fully differentiating the expression in (E22) with respect to *t*, employing (E21) and (E22) and re-arranging yields:

(E23) $\frac{\partial^2 y}{\partial t \partial w} = -\frac{w}{h''(y/w)} \cdot \left[ 1 + \frac{\partial \left( \frac{h'(l)}{h''(l)} \right)}{\partial l} \right] < 0,$

where the last inequality follows from our assumption that *h''/h'* is non-increasing in *l*. This completes the proof.

## Appendix F: Proof of Proposition 5

The unique equilibrium for the game between the two countries is given by the (unique) implicit solution to [see equation (E7) in appendix E]:

$$\text{(F1)}\quad F\left(t,\delta\right)\equiv\frac{\partial V_{11}}{\partial t}+\lambda\left(\frac{\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)}{2\delta}\frac{\partial V_{12}}{\partial t}+A\right)=0$$

From (F1) it follows that,

$$\text{(F2)}\quad \frac{\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)}{2\delta}\frac{\partial V_{12}}{\partial t}+A>0.$$

Differentiating the expression in (E6) with respect to $\delta$, yields:

$$\text{(F3)}\quad \frac{\partial\lambda}{\partial\delta}=\frac{2\cdot\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)}{\left(2\delta-\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)\right)^2}>0.$$

Differentiating the expression on the right-hand side of (F1) with respect to $\delta$, employing (F2) and (F3) and the fact that $\frac{\partial V_{12}}{\partial t}=-y_{12}$, yields:

$$\text{(F4)}\quad \frac{\partial F\left(t,\delta\right)}{\partial\delta}=\frac{\partial\lambda}{\partial\delta}\cdot\left(\frac{\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)}{2\delta}\frac{\partial V_{12}}{\partial t}+A\right)-\lambda\left(\frac{\alpha_2\cdot\alpha_1\cdot t\cdot\left(y_{12}-y_{11}\right)}{2\delta^2}\frac{\partial V_{12}}{\partial t}\right)>0.$$

By virtue of the fact that $\frac{\partial F\left(t,\delta\right)}{\partial t}<0$ (see Appendix E for details) one obtains:

$$\text{(F5)}\quad \left.\frac{\partial t}{\partial\delta}\right|_{F(t,\delta)=0}=-\frac{\dfrac{\partial F\left(t,\delta\right)}{\partial\delta}}{\dfrac{\partial F\left(t,\delta\right)}{\partial t}}>0.$$

By virtue of (F5) it follows that $\frac{\partial T}{\partial \delta} > 0$. To see this, suppose by negation, that $\frac{\partial T}{\partial \delta} \leq 0$ for some $\delta$. Let $t(\delta)$ and $T(\delta)$ denote, respectively, the tax rate and the demo-grant set in the symmetric equilibrium by both countries, when migration costs are given by $\delta$. One can show that choosing $t(\delta)$ and $T(\delta)$ is not a best response for country 1. To see this, consider a deviation to an alternative tax system given by the pair $t(\delta')$ and $T(\delta')$, where $\delta > \delta'$; that is, $t(\delta')$ and $T(\delta')$ denote, respectively, the tax rate and the demo-grant set in the symmetric equilibrium by both countries, when migration costs are given by $\delta'$. Notice that $t(\delta) > t(\delta')$, by virtue of (F5); and $T(\delta) \leq T(\delta')$, by our presumption. By deviating, country 1 will increase the utility of both individuals, and will also attract high-skill migrants from country 2, which will create a fiscal surplus [notice that in the absence of these additional migrants, the pair $t(\delta')$ and $T(\delta')$ satisfies the government revenue constraint as equality, by construction]. We thus obtain the desired contradiction.

We finally turn to show that the utility of the low-skill individuals is increasing with respect to $\delta$, whereas, the utility of their high-skill counterparts is decreasing with respect to $\delta$. We start by examining the utility of the low-skill individuals.

Consider the optimization program solved by a government seeking to maximize the utility of the typical low-skill individual under autarky. The government will maximize:

(F6) $\quad V(w_1, t, T) \equiv \max_y [y(1-t) + T - h(y / w_1)]$,

subject to the revenues constraint:

(F7) $\quad T = t \cdot [\alpha_1 / 2 \cdot y_1(t) + \alpha_2 / 2 \cdot y_2(t)] \equiv K(t)$,

where $y_j(t)$ is the implicit solution to $w_j \cdot (1-t) = h'(y / w_j)$.

Substituting for *T* from (F7) into (F6), yields:

(F8) $J(w_1,t) \equiv V[w_1,t,K(t)]$.

Let $t^* = \arg\max J(w_1,t)$. As the case of autarky (no migration) is obtained as the limiting case of the migration equilibrium, it follows that $t^* = \lim_{\delta\to\infty} t(\delta)$, where, as before, $t(\delta)$ denotes the tax rate set by both countries in the migration equilibrium, when migration costs are given by $\delta$. By the second-order conditions for the government optimization, it follows that $\partial J(w_1,t)/\partial t > 0$ for all $t < t^*$. It follows then by virtue of (F5) that the utility derived by the low-skill individual is indeed rising with respect to $\delta$ (notice that in a symmetric equilibrium the distribution of population across the two countries will be identical to that under autarky).

We turn next to examine the utility derived by the high-skill individuals. Consider the optimization program solved by a government seeking to maximize the utility of the typical high-skill individual under autarky. Suppose that the tax rate is restricted to be non-negative (the system is restricted to be progressive). The government will maximize:

(F9) $V(w_2,t,T) \equiv \max_y [y(1-t)+T-h(y/w_2)]$,

subject to the revenues constraint:

(F10) $T = t\cdot[\alpha_1/2\cdot y_1(t)+\alpha_2/2\cdot y_2(t)] \equiv K(t)$,

where $y_j(t)$ is the implicit solution to $w_j\cdot(1-t) = h'(y/w_j)$.

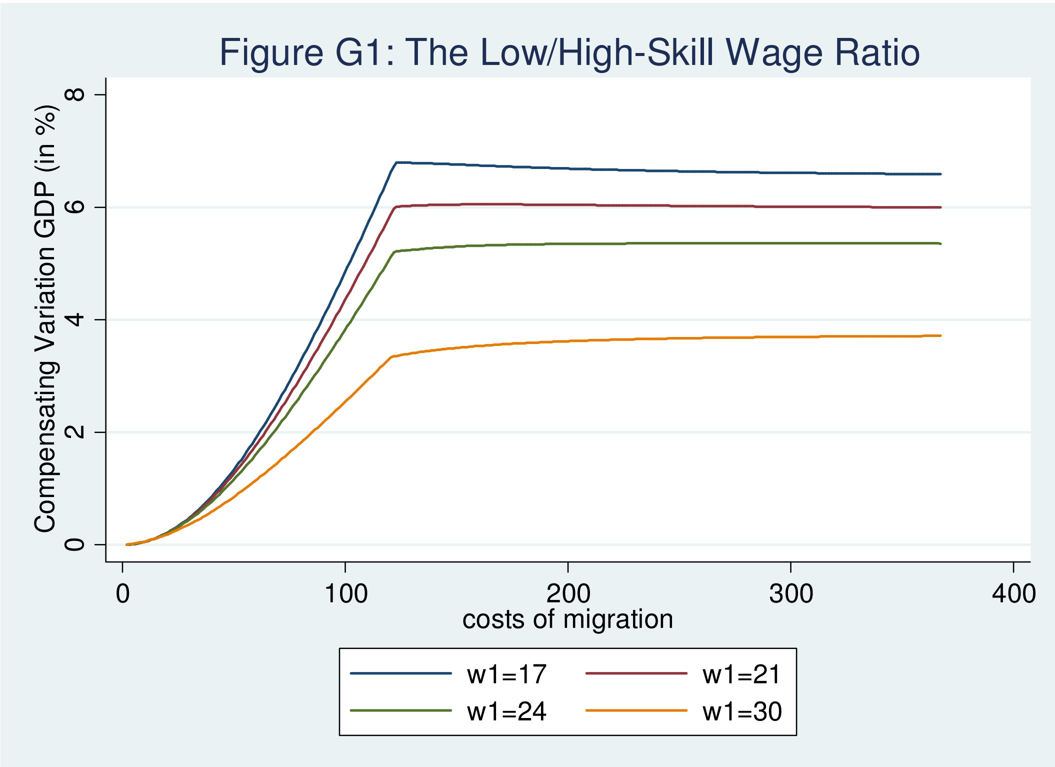
Substituting for *T* from (F10) into (F9), yields:

(F11) $H(w_2,t) \equiv V[w_2,t,K(t)]$.

Let $t^{**} = \arg\max H(w_2, t)$ . As we restrict ourselves to progressive tax-and-transfer systems, it follows that $t^{**} = 0$ , because the high-skill individuals have nothing to gain from re-distribution. As the case of no re-distribution is obtained as a limiting case of the migration equilibrium with costless migration (due to the *Bertrand* competition between the two countries) it follows that $t^{**} = \lim_{\delta \to 0} t(\delta)$. By the second-order conditions for the government optimization, it follows that $\partial H(w_2, t) / \partial t < 0$ for all $t > 0$. It follows then by virtue of (F5) that the utility derived by the high-skill individual is indeed decreasing with respect to $\delta$ (notice that in a symmetric equilibrium the distribution of population across the two countries will be identical to that under autarky). This completes the proof.

## Appendix G: Robustness Simulations

In this appendix we provide several simulations demonstrating the robustness of our key results to the change in the parameters specification. Our benchmark parameters are: $w_1 = 17.9$ (low-skill wage-rate), $w_2 = 45.645$ (high-skill wage-rate), $e$=0.4 (elasticity of taxable income) and $\alpha_2$=0.736 (number of high-skill workers).[14] In figure G1 we examine the effect of the low/high-skill wage ratio. In figure G2 we focus on the effect of the elasticity of taxable income. We conclude by examining the effect of the proportion of high-skill workers in the general population in figure G3.




[14] Recall that total world population is normalized to 2, so that the proportion of high-skill workers is given by 0.736/2=0.3518 (consistent with our assumption that the low-skill workers form the majority in the population).

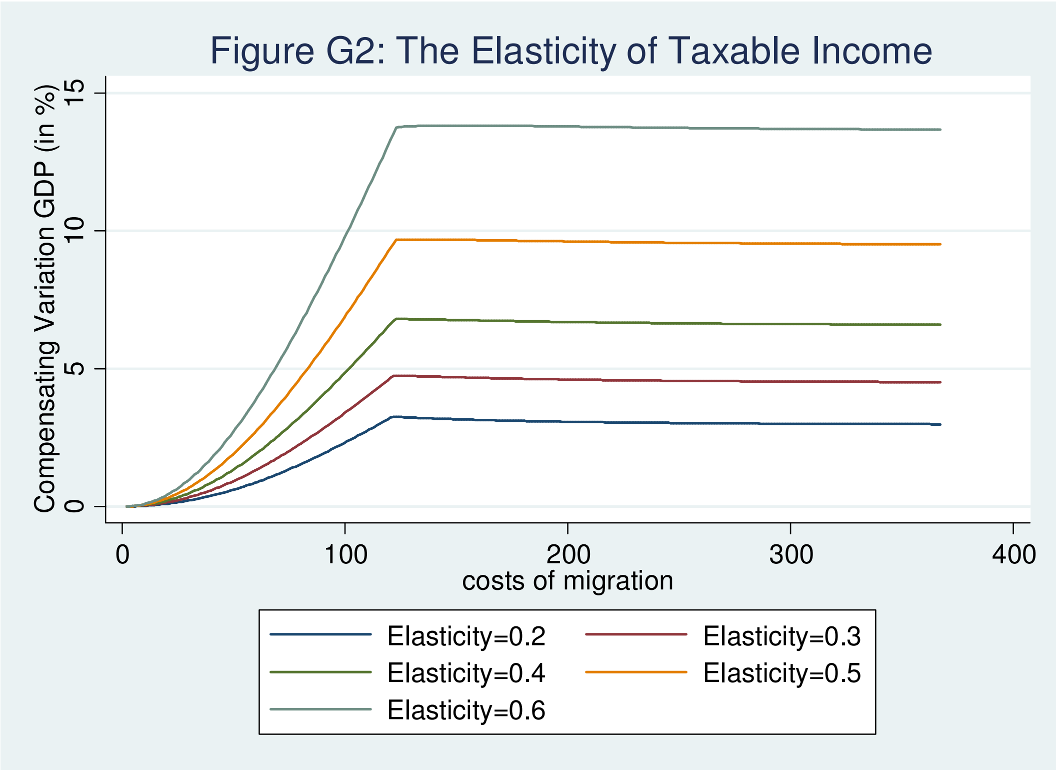


Figure G2: The Elasticity of Taxable Income
Compensating Variation GDP (in %)
0
5
10
15
0
100
200
300
400
costs of migration
Elasticity=0.2
Elasticity=0.3
Elasticity=0.4
Elasticity=0.5
Elasticity=0.6

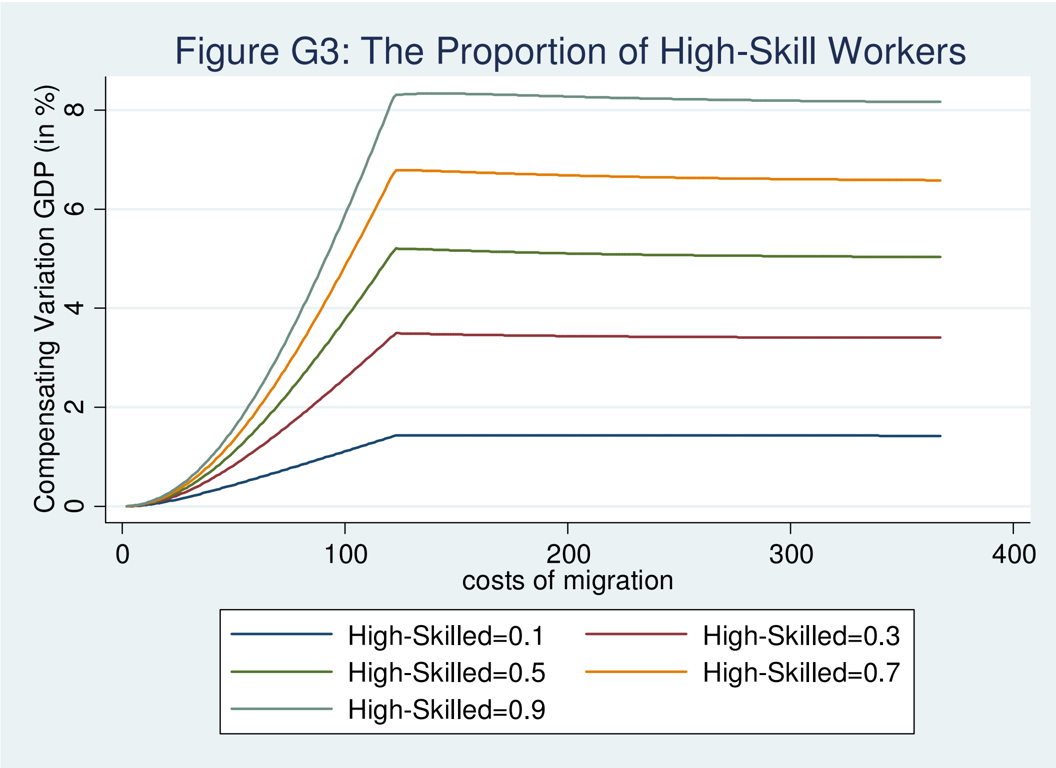


Figure G3: The Proportion of High-Skill Workers
Compensating Variation GDP (in %)
0
2
4
6
8
0
100
200
300
400
costs of migration
High-Skilled=0.1
High-Skilled=0.3
High-Skilled=0.5
High-Skilled=0.7
High-Skilled=0.9

## References

Balcer, Y. and E. Sadka (1982) "Horizontal Equity, Income Taxation and Self-selection with an Application to Income Tax Credits", Journal of Public Economics, 19, 291-309

Brett, C. and J. Weymark (2008) "Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility", Working Paper No. 08-W12, Department of Economics, Vanderbilt University.

Docquier, F. and A. Marfouk (2005) "International Migration by Education Attainment, 1990-2000", in International Migration, Remittances and the Brain-Drain, 151-200, World Bank.

Gruber, J. and E. Saez (2000) "The Elasticity of Taxable Income: Evidence and Implications", Journal of Public Economics, 84, 1-32.

Gorodnichenko, Y. et al. (2009) "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia", Journal of Political Economy, 117, 504-54.

Hamilton, J. and P. Pestieau (2005) "Optimal Income Taxation and the Ability Distribution: Implications for Migration Equilibria", International Tax and Public Finance, 12, 29-45.

Hall, R. and A. Rabushka (1985) "The Flat Tax", Hoover Institution Press, Stanford.

Hindriks, J. (1999) "The Consequences of Labor Mobility for Redistribution: Tax vs. Transfer Competition", Journal of Public Economics, 74, 215-234.

Ivanova, A. et al. (2005) "The Russian 'Flat Tax' Reform", Economic Policy, 43, 397-435.

Krause, A. (2007) “Education and Taxation Policies in the Presence of Countervailing Incentives”, Economica, forthcoming.

Mirrlees, J. (1971) “An exploration in the Theory of Optimum Income Taxation”, Review of Economic Studies, 38, 175-208.

------- (1982) “Migration and Optimal Income Taxes”, Journal of Public Economics, 18, 319–341.

Morelli, M. et al (2010) “Competitive Nonlinear Taxation and Constitutional Choice”, mimeo, Department of Economics, Ohio State University.

OECD, (2002) International Mobility of Highly Skilled, OECD, Paris.

OECD, (2005) OECD Economic Surveys: Slovakia, OECD, Paris.

OECD, (2008) OECD Economic Surveys: Netherlands, Volume I, OECD, Paris.

Osmundsen, P. (1999) “Taxing Internationally Mobile Individuals: A Case of Countervailing Incentives”, International Tax and Public Finance, 6, 149-164.

Piaser, G. (2007): “Labor Mobility and Income Tax Competition”, In: Gregoriou, G., Read, C. (Eds.), International Taxation Handbook: Policy, Practice, Standards and Regulation, CIMA Publishing, Oxford, 73-94.

Saez, E. (2002) “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses”, Quarterly Journal of Economics, 117, 1039-73.

Salanie, B. (2003) “The Economics of Taxation”, The MIT Press, Cambridge, Mass.

Simula, L. and A. Trannoy (2010) “Optimal Income Tax under the Threat of Migration by Top-Earners”, Journal of Public Economics, 94, 163-73.

Stiglitz, J. (1982) "Self-Selection and Pareto Efficient Taxation", Journal of Public Economics, 17, 213–240.

Tuomala, M. (1990) "Optimal Income Tax and Redistribution", Clarendon Press, Oxford.

Wildasin, D. (1994) "Income Redistribution and Migration", Canadian Journal of Economics, 27, 637-656.

Wilson, J. (1980) "The Effect of Potential Emigration on the Optimal Linear

Income Tax", Journal of Public Economics, 14, 339-353.

**--------** (1986) "A Theory of Interregional Tax Competition", Journal of Urban Economics, 19, 296-315.

**--------** (1992) "Optimal Income Taxation and International Personal Mobility", American Economic Review, 82, 191-196.

------- (1999) "Theories of Tax Competition", National Tax Journal, 52, 269-304.

----- (2006) "Income Taxation and Skilled Migration: The Analytical Issues", mimeo, Department of Economics, Michigan State University.

Zodrow, G. and P. Mieszkowski (1986) "Pigou, Tiebout, Property Taxation, and the

Underprovision of Local Public Goods", Journal of Urban Economics, 19, 356-70.