UPPSALA UNIVERSITET

Uppsala Center for Fiscal Studies
Department of Economics

Working Paper 2009:9

Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective

Laurent Simula and Alain Trannoy

Uppsala Center for Fiscal Studies
Department of Economics
Uppsala University
P.O. Box 513
SE-751 20 Uppsala
Sweden
Fax: +46 18 471 14 78

Working paper 2009:9
July 2009

Shall we keep highly skilled at home? The optimal income tax perspective

Laurent Simula and Alain Trannoy

Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective*

Laurent SIMULA
Uppsala University and Uppsala Center for Fiscal Studies†

Alain TRANNOY
Ecole des Hautes Etudes en Sciences Sociales‡

July 18, 2009

Abstract

We examine how allowing individuals to emigrate to pay lower taxes abroad changes the optimal non-linear income tax scheme in a Mirrleesian economy. An individual emigrates if his domestic utility is less than his utility abroad net of migration costs, utilities and costs both depending on productivity. Three average social criteria are distinguished – national, citizen and resident – according to the agents whose welfare matters. A curse of the middle-skilled occurs in the first-best and it may be optimal to let some highly skilled leave the country under the resident criterion. In the second-best, we provide an extension of Saez's formula for the optimal marginal tax rates. The middle-skilled can support the highest average tax rates and the marginal tax rates can be negative. Preventing emigration of the highly skilled is not necessarily optimal under the citizen and resident criteria.

Keywords: Optimal Income Tax, Emigration, Participation Constraints, Highly Skilled

JEL classification: H21; H31; D82; F22.

*We are grateful to Thomas Aronsson, David Bevan, Chuck Blackorby, Julia Cagé, Philippe Choné, David de la Croix, Jeremy Edwards, Jonathan Hamilton, Jean Hindriks, Mathias Hungerbühler, Laurence Jacquet, Karolina Kaiser, Etienne Lehman, Jean-Marie Lozachmeur, Hamish Low, François Maniquet, Michael S. Michael, James Mirrlees, Gareth Myles, Frank Page, Pierre Pestieau, Panu Poutvaara, Emanuela Sciubba, Gwenola Trotin, John Weymark and David Wildasin for their very helpful comments and suggestions. The usual caveats apply.

†Uppsala University, Department of Economics, P.O. Box 513, SE-75120 Uppsala, Sweden. E-mail: laurent.simula@nek.uu.se

‡EHESS, GREQAM and IDEP, Centre de la Vieille Charité, 2 rue de la Charité, 13236 Marseille cedex 02, France. E-mail: alain.trannoy@univmed.fr

1. INTRODUCTION

As noted by Mirrlees (1982), "high tax rates encourage emigration. The resulting loss of tax revenue is widely believed to be an important reason for keeping taxes down". Today, the governments of many developed countries, notably in continental Europe, fear the departure of their highly skilled to less redistributive countries and adjust their top income tax rates consequently. However, is it always socially optimal to prevent top-income earners from emigrating?

To address this issue, we study the optimal nonlinear income tax in a Mirrleesian economy the citizens of which have the possibility to vote with their feet and settle down in a less redistributive country in order to pay lower taxes. The tax policy abroad is given and an illustration is the laissez-faire or a flat tax with a low tax rate. The government wants to redistribute incomes from the more to the less productive individuals as in Mirrlees's (1971) model. Yet, in addition to the standard incentive constraints, it has to take account of participation constraints for the individuals it wants to keep at home. Because more productive individuals are likely to have more attractive outside options (e.g.,Hanson (2005) and Docquier and Marfouk (2006)), these participation constraints are type dependent. We borrow these constraints from recent articles in contract theory (cf. Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995), and Jullien (2000)) and introduce them in the optimal income tax problem.

The social objective is more complex to specify when individuals are allowed to vote with their feet because the set of agents whose welfare is to count can depend on the income tax itself. We distinguish three social criteria. Under the National criterion, the domestic government maximizes the average welfare of its citizens whilst ensuring that every citizen lives at home. Under the Citizen criterion, it maximizes the average welfare of its citizens, irrespective of their country of residence. Under the Resident criterion, it maximizes the average welfare of its residents.

The model is designed to cast light on the impact of the potential emigration of highly skilled, driven by a significant asymmetry in tax levels between home and abroad. In particular, we consider that foreigners do not migrate to the high tax country and are, thus, immobile. Our main findings can be summarized as follows.

When each individual's productivity is public information (first-best), it is socially optimal to prevent emigration of the highly-skilled individuals under the Citizen criterion, which coincides therefore with the National criterion at the optimum. By contrast, emigration of highly-skilled workers may be socially optimal under the Resident criterion. In every case, there is a curse of the middle-skilled workers at the optimum, instead of the curse of the highly skilled obtained in closed economy (Mirrlees, 1974). Indeed, it is no longer possible to demand as much work as without mobility from the highly skilled individuals, so the productive rent is extracted to the maximum from the most productive individuals among those insufficiently talented to threaten to emigrate. However, these middle-skilled workers cannot be taxed at will because they would otherwise threaten to emigrate. Consequently, the redistribution in favour of the low-skilled individuals has to be reduced.

When each individual's productivity is private information (second-best), two qualitative properties of the optimal marginal tax rates are lost: they can be non-positive at interior points and strictly negative at the top. Consequently, individual mobility does not only render the tax schedule less progressive, but can also make the tax function decreasing. In fact, the small tax reform perturbation around the optimal tax scheme used by Piketty (1997) and Saez (2001) has an additional participation effect on social welfare, which favours a decrease in the optimal marginal tax rates even for individuals below the productivity levels where the individuals threaten to emigrate. This new effect results in changes in Mirrlees's formula to ensure that the optimal average tax rates are compatible with the participation constraints of the individuals threatening to emigrate. In addition, the interaction between the type-dependent participation constraints and the incentive compatibility conditions can give rise to countervailing incentives, in which case less skilled individuals want to mimic more skilled individuals because the latter have more appealing outside options. Countervailing incentives cause an indirect social cost of the presence in A of the highly-skilled individuals. The Citizen and Resident criteria allow us to consider whether it is not too expensive in terms of social welfare to implement a tax scheme which prevents emigration of the highly skilled workers. When the indirect cost due to countervailing incentives prevails over the benefits of them staying in A, implementing a tax schedule inducing them to emigrate increases social welfare.

As far as we know, Osmundsen (1999) is the first to examine income taxation with type-dependent participation constraints. He studies how highly skilled individuals distribute their working time between two countries. Because he directly uses the model developed by Maggi and Rodriguez-Clare (1995), there is no individual trade-off between consumption and leisure. By contrast, our model takes this trade-off into account. In a recent article, Krause (2008) has examined income taxation and education policy when there exist conflicting incentives for individuals to understate and overstate their productivity. Highly-skilled individuals are better educated and can thus benefit from higher outside options when emigrating. Using quasilinear-in-leisure preferences and a two-type model, different possible regime are identified but no optimal tax scheme is characterized. Simula and Trannoy (2006) address the same issue as in the present paper when the income tax is linear. We show that it may be socially optimal to let highly skilled leave the home country under the resident criterion. In Simula and Trannoy (2009), we examine the impact of the threat of migration by highly skilled under a set of simplifying assumptions and derive simple formulae for the top marginal income tax rates. In particular, we only consider the National social objective and, thus, migration never actually occurs. Moreover, several articles have adopted the viewpoint of tax competition, restricting attention to personalised lump-sum taxes (Leite-Monteiro, 1997), considering a two-type population as in Stiglitz (1982) (Huber, 1999, Hamilton and Pestieau, 2005, Piaser, 2003) or a population with many types (Brett and Weymark, 2008, Morelli, Yang, and Ye, 2008).

Our article is organized as follows. Section 2 sets up the model. Section 3 examines the first-best optimal allocations. Section 4 studies the properties of the second-best optimal allocations. Section 5 concludes.

2. THE MODEL

The world consists of two countries, the home country A and the foreign country B. All individuals are initially living in country A. Country A's government implements a redistributive tax policy. Country B is committed to being a laissez-faire country, but we could more generally consider that it implements a (fixed) income tax, with a low constant marginal tax rate and no basic income. Governments provide no public goods. Both countries have the same production function with constant returns to scale. Hence, productivity levels, equal to pre-tax wage rates, are independent of the country in which an individual is working.

Individuals differ in productivities θ, which are private information. The cumulative distribution function of θ, denoted F, is common knowledge. It is defined on $\left[\underline{\theta}, \overline{\theta}\right] := \Theta \subseteq \mathbb{R}^+$, where it admits a continuous and strictly positive density f.

2.1. Individual Behaviour

All individuals have the same preferences over consumption x and labour ℓ. If $\overline{\ell}$ is the time endowment, these preferences are represented by a cardinal utility function $U : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} := \left\{(x, \ell) \in \mathbb{R}^+ \times [0, \overline{\ell})\right\}$.

Assumption 1. *U is a $\mathcal{C}^2$ strictly concave function such that $U_x > 0$, $U_\ell < 0$ and $U \to -\infty$ as $x \overset{>}{\to} 0$ or $\ell \overset{<}{\to} \overline{\ell}$.*

Assumption 2. *Leisure is a normal good.*

A θ-individual working ℓ units of time has gross income $z := \theta\ell$. We call

$$u\left(x, z; \theta\right) := U\left(x, z/\theta\right) \tag{1}$$

the personalized utility function and note that $u'_x = U'_x$, $u'_z = U'_\ell/\theta$, $u''_{xx} = U''_{xx}$, $u''_{xz} = U''_{x\ell}/\theta$, $u''_{zz} = U''_{\ell\ell}/\theta^2$. The marginal rate of substitution of gross income for consumption of a θ-individual at (x, z) is

$$s\left(x, z; \theta\right) := -\frac{u'_z\left(x, z; \theta\right)}{u'_x\left(x, z; \theta\right)}. \tag{2}$$

Each individual decides about the optimal amount of consumption and labour to maximize his utility subject to his budget constraint. Country A's government uses a tax function $T\left(\theta, \ell\right)$, with $T\left(\theta, \ell\right) = T\left(\theta\right)$ in the first-best and $T\left(\theta, \ell\right) = T\left(\theta\ell\right)$ in the second-best. The utility maximization programme in A implicitly defines the consumption and labour supply functions in A, denoted $x_A\left(\theta\right)$ and $\ell_A\left(\theta\right)$ respectively. The indirect utility in A is thus $V_A\left(\theta\right) := U\left(x_A\left(\theta\right), \ell_A\left(\theta\right)\right)$.

The utility maximization programme in B defines implicitly the consumption and labour supply functions in B, denoted $x_B\left(\theta\right)$ and $\ell_B\left(\theta\right)$ respectively. The indirect utility in B is thus $V_B\left(\theta\right) := U\left(x_B\left(\theta\right), \ell_B\left(\theta\right)\right)$, which is strictly increasing in θ.

2.2. Emigration and Participation Constraints

An individual leaving country A pays a strictly positive *migration cost* c. Given the cardinality of individual preferences, this cost can be expressed as a "time-equivalent" loss in utility, due to different material and psychic costs of moving: application fees, transportation of persons and household's goods, forgone earnings, costs of speaking a different language and adapting to another culture, costs of leaving one's family and friends, etc. "[These migration] costs probably vary among persons [but] the sign of the correlation between costs and wages is ambiguous" (Borjas, 1999, p. 12). We consider that they depend on productivity and that their distribution is known to A's government. Hence, A's government knows $c(\theta)$ when it knows θ, which is thus the sole parameter of heterogeneity within the population. In addition:

Assumption 3. *$c : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}^{++}$ is a $\mathcal{C}^2$ function satisfying $c'(\theta) < V_B'(\theta)$.*

The *reservation utility* is the maximum utility an individual staying in A can obtain abroad. It is thus equal to $V_B(\theta) - c(\theta)$. Assumption 3 amounts therefore to considering that the outside opportunities are increasing in productivity. This is in accordance with many empirical studies, which find that the propensity to migrate increases with productivity (Sahota, 1968, Schwartz, 1973, Gordon and McCormick, 1981, Nakosteen and Zimmer, 1980, Inoki and Surugan, 1981, Hanson, 2005, Docquier and Marfouk, 2005). For instance, within the EU, the migration rate of the skilled population is 8.1% versus 4.8 for the unskilled one (Docquier and Marfouk, 2005). It thus does not seem too egregious to consider that more skilled individuals should have more attractive outside options, i.e. higher reservation utilities. Moreover, Assumption 3 places *no* restriction on the *level* of the migration costs, expect that they are positive.

The *location rent* of a θ-individual is the excess of his indirect utility in A over his reservation utility, i.e.

$$R(\theta) = V_A(\theta) - V_B(\theta) + c(\theta). \tag{3}$$

An individual stays in A if and only if

$$R(\theta) \geq 0, \tag{4}$$

and therefore leaves A if and only if $R(\theta) < 0$.

A *citizen* is defined as an individual born in country A, so all individuals have country A's citizenship. Individuals are committed to working in the country where they live. Since the focus is on the mobility of highly skilled individuals, we consider that there is a partition of citizens between A and B, with the less skilled individuals being immobile and staying in A.

Assumption 4. *Country $A's$ resident population is a closed interval of types $\left[\underline{\theta}, \widehat{\theta}\right]$, with $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$.*[1]

[1] This assumption is a restriction on the class of feasible tax schedules. In the absence of this assumption, the resident population in A might consist of several disjoint intervals. It thus not seem possible to address such issues with optimal control theory (Seierstad and Sydsaeter, 1987).

We consider that A's government is not able to levy taxes in B, since the fiscal prerogative is closely linked to national sovereignty, and not willing to redistribute income to the individuals staying in B. Consequently, $T : \mathcal{T} \to \mathbb{R}$ with $\mathcal{T} = \left[\underline{\theta}, \widehat{\theta}\right] \times \left[0, \overline{\ell}\right]$. Since $T := z_A - x_A$, a tax policy is *budget balanced* if and only if it satisfies the tax revenue constraint

$$\int_{\underline{\theta}}^{\widehat{\theta}} (z_A - x_A)\, dF(\theta) \geq 0. \tag{TR}$$

In the rest of the paper, we denote by γ the Lagrange multiplier associated with the budget constraint (TR).

2.3. Social Criteria

A's government is a benevolent policy maker which intends to implement the tax policy corresponding to the best compromise between equity and efficiency. Its desire to redistribute income is captured through its aversion to income inequality $\rho \in \mathbb{R}^{+}$. A zero aversion corresponds to utilitarianism and an infinite one to the Rawlsian maximin.

The social objective is more difficult to specify than in closed economy. Indeed, it does not only depend on ρ which is captured through an isoelastic function defined by $\phi_\rho : \mathbb{R}^{++} \to \mathbb{R}$, $\phi_\rho(U) = U^{1-\rho}/(1-\rho)$ for $\rho \neq 1$ and $\phi_1(U) = \ln U$ for $\rho = 1$, but also on the answers to the following questions. First, should we maximize total or average social welfare? Classical utilitarianism has been criticized on the ground that it leads to the so-called repugnant conclusion and this is a significant shortcoming (For details, see Blackorby, Bossert, and Donaldson (2005)). Average utilitarianism does not suffer from this drawback. So, we consider that the government is interested in social welfare per capita, which allows us to compare allocations differing in population size. Second, who are the agents whose welfare is to count? At least three social criteria can be proposed, each of which corresponds to a specific answer.

Under the *National criterion*, $A's$ government cares about the welfare of all its citizens and wants each citizen to choose to stay in A. The social objective is

$$W_{A,\rho}^{N} := \int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A(\theta))\, dF(\theta), \text{ with } W_{A,\rho}^{N} = -\infty \text{ for } \widehat{\theta} < \overline{\theta}. \tag{5}$$

This objective corresponds to the mercantilist idea, formulated by Bodin (1578), that "the only source of welfare is mankind itself". Emigration should therefore be prevented to keep the state prosperous. This social criterion is considered to provide a building block for the solutions of the following more appealing Citizen and Resident criteria.

Under the *Citizen criterion*, A's government cares about the average social welfare of its citizens, whether they are in A or B. Under Assumption 4, the social objective is

$$W_{A,\rho}^{C}\left(\widehat{\theta}\right) := \int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A(\theta))\, dF(\theta) + \int_{\widehat{\theta}}^{\overline{\theta}} \phi_\rho(V_B(\theta) - c(\theta))\, dF(\theta). \tag{6}$$

This criterion rests on the idea that the fiscal system finds its legitimacy in its democratic adoption. Consequently, the welfare of every individual who has the right to vote should be taken into account, irrespective of his country of residence[2]. When this objective is chosen, the optimal tax function depends on the choice of $\widehat{\theta}$ and determines an allocation of A's citizens between A and B. Hence, A's resident population is endogenous while the set of agents the welfare of whom matters is exogenously fixed.

Under the *Resident criterion*, A's government cares about the average social welfare of its residents. Under Assumption 4, the social objective is

$$W_{A,\rho}^{R}\left(\widehat{\theta}\right) := \frac{1}{F\left(\widehat{\theta}\right)} \int_{\underline{\theta}}^{\widehat{\theta}} \phi_{\rho}\left(V_A\left(\theta\right)\right) dF\left(\theta\right). \tag{7}$$

This criterion is based on the idea that a public policy should take the welfare of all taxpayers into account. Consequently, the welfare of the citizens living in B does not count. When this objective is chosen, the tax function as well as the set of agents whose welfare is to count depend on the choice of $\widehat{\theta}$.[3] $W_{A,\rho}^{R}\left(\widehat{\theta}\right)$ is based on average utilitarianism, which is known to face the Mere Addition Paradox: the addition of individuals whose utility is less than the average utility in the initial population is regarded as suboptimal even if this change in population size affects no one else and does not involve social injustice. In the second-best framework, this paradox does not really matter herein because we are focusing on emigration of the highest skilled individuals initially living in A, whose utility is greater than the maximum utility in A.

3. FIRST-BEST OPTIMAL ALLOCATIONS

This section characterizes the first-best optimal allocations where each individual's productivity is public information. Consequently, A's government implements a tax policy depending on productivity, i.e. $T\left(\theta, \ell\right) = T\left(\theta\right)$. We restrict attention to the tax schedules which are continuous and differentiable almost everywhere.

The indirect utility if A were a closed economy, $V_A^{c\ell}\left(\theta\right)$, is used as a benchmark. When ρ is finite, it is decreasing in θ at the social optimum if and only if Assumption 2 holds (Mirrlees, 1974): there is therefore a *curse of the highly skilled workers*. When ρ is infinite, all individuals receive the same utility level. In this section, we assume $V_A^{c\ell}\left(\overline{\theta}\right) < V_B\left(\overline{\theta}\right) - c\left(\overline{\theta}\right)$ since otherwise the participation constraints would never be active. It is worth noting that, in the first-best setting, individuals for whom the participation constraints are active pay strictly positive taxes. Indeed, since $V_B\left(\theta\right) - c\left(\theta\right) < V_B\left(\theta\right)$ under Assumption 3, their budget constraints must be below

[2] In France, the 14th Article of the Declaration of the Rights of Man and of the Citizen, which has constitutional value, provides that: "*All citizens* have the right to vote, by themselves or through their representatives, for the need for the public contribution, to agree to it voluntarily, to allow implementation of it, and to determine its appropriation, the amount of assessment, its collection and its duration". For example, twelve senators represent the French citizens living abroad.

[3] In other words, a population problem consisting in "*different* number choices" (Parfit, 1984) is embedded in the optimal income tax problem.

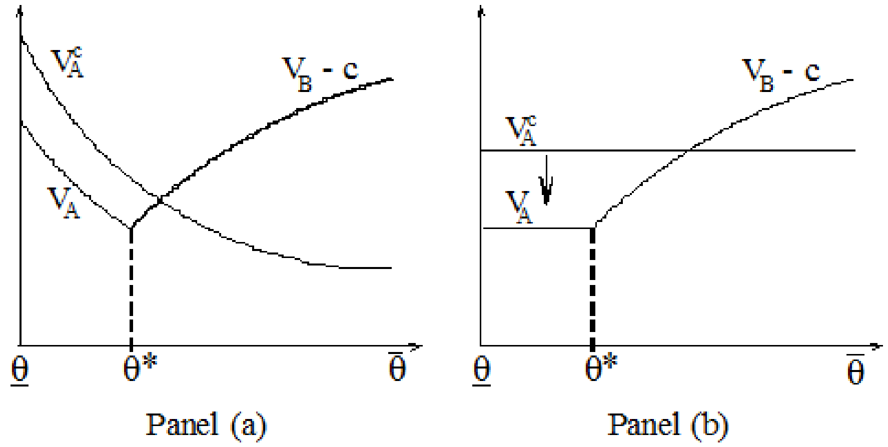

Figure 1: The curse of the middle-skilled workers

the 45°-line through the origin in the gross-income/consumption space.

3.1. National criterion

A's government chooses the tax paid by each individual or, equivalently, the consumption-labour bundle intended for each individual.

Problem 1 (National Criterion, First-Best)**.** *Find* $(x, \ell) \in \mathcal{X}$ *to maximize* $W^N_{A,\rho}$*, with* $\widehat{\theta} = \overline{\theta}$*, subject to (TR) and*

$$R(\theta) \geq 0 \textit{ for } \theta \leq \widehat{\theta}. \tag{PC}$$

When the set $\left\{\theta \in \left[\underline{\theta}, \overline{\theta}\right] : (4) \text{ binding}\right\}$ is non-empty, we call θ^* its *infimum*. The threshold θ^* is therefore the minimum productivity level for which participation constraints are active and, by definition, θ^* is less than the highest productivity level within the resident population $\widehat{\theta}$.

Proposition 1 (The Curse of the Middle-Skilled)**.** *The participation constraints are binding from* $\theta^* < \overline{\theta}$*. When* $\rho < \infty$*, the optimum indirect utility in* A *is V-shaped in* θ*, minimum at* θ^**. When* $\rho \to \infty$*, the optimum indirect utility in* A *is constant up to* θ^* *and then increasing.*

Proof. See A.1 in the Appendix. □

Figure 1 illustrates Proposition 1. On panel (a), the government's aversion to income inequality is finite. The θ^*-individuals are the worse-off when potential mobility is taken into account. On panel (b), the government is Rawlsian. The utility levels of the individuals with productivity below θ^* are reduced compared to the closed economy.

The participation constraints (PC) separate the population into two intervals: they are inactive below θ^* and active above. Consequently, it is no longer possible to require the most talented individuals to work as much as without mobility, i.e. to require them to keep working even though labour disutility exceeds the gains from the increase in income. The productive rent is thus extracted to the maximum from the most productive individuals among those threatening to emigrate. Redistribution in A is reduced and the situation of the low-skilled individuals gets worse.

It is therefore from the most productive individuals among those insufficiently talented to threaten to leave the country that the productive rent is extracted to the maximum. However, this rent cannot be extracted at will because of the participation constraints. Redistribution in A is thus reduced and the situation of the low-skilled individuals deteriorates.

3.2. Citizen Criterion

We examine if it is socially optimal to prevent emigration of the highly skilled individuals under the Citizen criterion.

Problem 2 (Citizen Criterion, First-Best)**.** *Find* $(x, \ell) \in \mathcal{X}$ *and* $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$ *to maximize* $W^C_{A,\rho}\left(\widehat{\theta}\right)$ *subject to (PC) and (TR).*

Proposition 2. *Under the Citizen criterion, the optimal tax policy is the same as that chosen under the National criterion.*

The proof proceeds by contradiction. Assume $\widehat{\theta} < \overline{\theta}$ is socially optimal. The individuals with productivity $\widehat{\theta}$ are indifferent between A and B, i.e. $R\left(\widehat{\theta}\right) = 0$, and those with productivity greater than $\widehat{\theta}$ emigrate to B. It is always *feasible* to make the latter relocate to A, without reducing the indirect utilities of A's residents, in giving them their laissez-faire utility V_B (or a bit more than their reservation utility). Since $c(.) > 0$ and $\phi'_\rho(.) > 0$, one gets $\phi_\rho(V_B(.) - c(.)) < \phi_\rho(V_B(.))$ and thus

$$\int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A(\tau))\, dF(\tau) + \int_{\widehat{\theta}}^{\overline{\theta}} \phi_\rho(V_B(\tau))\, dF(\tau) > \int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A(\tau))\, dF(\tau) + \int_{\widehat{\theta}}^{\overline{\theta}} \phi_\rho(V_B(\tau) - c(\tau))\, dF(\tau)\,, \quad (8)$$

the RHS of which is $W^C_{A,\rho}\left(\widehat{\theta}\right)$. Hence, making the highly-skilled emigrate from B to A results in a feasible increase in social welfare, which contradicts the premise. The social optimum corresponds therefore to the corner solution as regards the allocation of individuals between A and B.

3.3. Resident Criterion

The basic difference between the citizen criterion and the resident one is that the latter does not take the welfare of A's citizens living in B. Hence, it might be socially desirable to let some individuals emigrate to B.

Problem 3 (Resident Criterion, First-Best)**.** *Find* $(x,\ell)\in\mathcal{X}$ *and* $\widehat{\theta}\in\left[\underline{\theta},\overline{\theta}\right]$ *to maximize* $W^R_{A,\rho}\left(\widehat{\theta}\right)$ *subject to (PC) and (TR).*

By the envelope theorem and Leibnitz's rule, the impact of a small increase in $\widehat{\theta}$ on the social objective is given by

$$\frac{\partial W^R_{A,\rho}\left(\widehat{\theta}\right)}{\partial\widehat{\theta}}=\gamma\frac{T\left(\widehat{\theta}\right)f\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)}+\frac{\left[\phi_\rho\left(V_A\left(\widehat{\theta}\right)\right)-W^R_{A,\rho}\left(\widehat{\theta}\right)\right]f\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)}. \tag{9}$$

The first term, which is strictly positive, corresponds to the fiscal contribution of the $\widehat{\theta}$-individuals $T\left(\widehat{\theta}\right)f\left(\widehat{\theta}\right)$, converted in social welfare per capita. The second term (9) basically compare the social utility of the marginal individuals, whose number is represented by $f\left(\widehat{\theta}\right)$, to the average social welfare. It is divided by the size of the population in order to obtain a per capita measurement. Its sign is ambiguous. As a consequence, a trade-off appears between the "tax effect" and the "utility effect" of the presence in A of the marginal $\widehat{\theta}$-individuals. In the optimum, (9) must be non-negative. As a consequence, a sufficient condition for emigration of the most productive citizens to be socially optimal is obtained.

Proposition 3. *Under the Resident criterion, letting the most productive citizens emigrate increases social welfare when* $\partial\overline{W}^R_{A,\rho}\left(\widehat{\theta}\right)/\partial\widehat{\theta}\Big|_{\widehat{\theta}=\overline{\theta}}<0$.

In order to determine the optimal upper bound of the resident population, $\widehat{\theta}$, the analogue of Problem 3 in which $\widehat{\theta}$ is arbitrarily given in $\left[\underline{\theta},\overline{\theta}\right]$ is first considered. Let $\overline{W}^R_{A,\rho}\left(\widehat{\theta}\right)$ be the social value function. The optimal value of $\widehat{\theta}$ is that for which $\overline{W}^R_{A,\rho}\left(\widehat{\theta}\right)$ is maximum.

4. SECOND-BEST OPTIMAL ALLOCATIONS

The distribution of characteristics in the economy remains common knowledge, but individual productivity is now private information. A's government is thus restricted to setting taxes as a function of earnings, i.e. $T(\theta,\ell)=T(z)$. Hence, it has to ensure that the tax schedule is incentive compatible.

4.1. Statement of the Problem

T is an *incentive compatible* tax schedule if and only if individuals living in A have an incentive to reveal their type truthfully when it is implemented. By the revelation principle, the incentive compatibility conditions read

$$u\left(x_A\left(\theta'\right),z_A\left(\theta'\right);\theta\right)\leq u\left(x_A\left(\theta\right),z_A\left(\theta\right);\theta\right)\ \text{for all}\ \left(\theta,\theta'\right)\in\left[\underline{\theta},\widehat{\theta}\right]^2. \tag{IC}$$

To deal with this uncountable infinity of constraints, the Spence-Mirrlees property is assumed to hold:

Assumption 5 (Single-Crossing)**.** $s'_\theta(x, z; \theta) < 0$.

Under Assumption 5, (IC) is equivalent to:

$$V'_A(\theta) = -\frac{z_A(\theta)}{\theta} u'_z(x_A(\theta), z_A(\theta); \theta) \text{ for } \theta \leq \widehat{\theta}, \tag{FOIC}$$

$$z_A(\theta) \text{ non-decreasing for } \theta \leq \widehat{\theta}. \tag{SOIC}$$

The proof of this equivalence is standard and is omitted. (FOIC) is an envelope condition specifying how the indirect utility V_A must locally change. Since $V'_A \geq 0$, V_A cannot be V-shaped as in the first-best. (SOIC) is a global monotonicity condition of gross income. The analysis will herein focus on continuous mechanisms which possibly exhibit kinks at a finite number of points corresponding to jumps of the marginal tax rate. In this case, $R(\theta)$ is continuous and (SOIC) is equivalent to

$$z'_A(\theta) \geq 0 \text{ for } \theta \leq \widehat{\theta}. \tag{SOIC'}$$

Since A's government does not know who are the agents for whom the location rent $R(\theta)$ is zero, the participation constraints and the incentive compatibility conditions have to be taken simultaneously into account for all A's residents[4]. The second-best optimal non-linear income tax problems read thus as follows.

Problem 4 (Second-Best)**.** *Find $T(z_A)$ to maximize $W^i_{A,\rho}$, $i = \{N, C, R\}$, subject to (i) (FOIC), (SOIC'), (PC), (TR); (ii) $\widehat{\theta} = \overline{\theta}$ when $i = N$ and $\widehat{\theta} \in [\underline{\theta}, \overline{\theta}]$ otherwise.*

In the closed economy version of Problem 4, $\widehat{\theta}$ is equal to $\overline{\theta}$ and the tax revenue constraints (TR) are not taken into account. Let $V^{c\ell}_A(\theta)$ be the optimum indirect utility. If $V^{c\ell}_A(\theta) \geq V_B(\theta) - c(\theta)$ for every $\theta \in [\underline{\theta}, \overline{\theta}]$, allowing individuals to vote with their feet does not alter the social optimum. Therefore, we herein place ourselves in the case where there are individuals for whom $V^{c\ell}_A(\theta) < V_B(\theta) - c(\theta)$ since otherwise the participation constraints would never be active.

Problem 4 raises three main difficulties compared to its closed-economy analogue. First, (PC) can a priori bind on any subset of the resident population, even at isolated points, because $R(\theta)$ is not necessarily monotonic. Second, (PC) are pure state constraints in the optimization problems. The adjoint variables may thus have jump discontinuities. Third, under the Citizen and Resident criteria, $\widehat{\theta}$ is free to vary between $\underline{\theta}$ and $\overline{\theta}$.

4.2. Optimal Tax Schedule for the Individuals Threatening to Emigrate

Before looking at a specific social criterion, we derive properties which are satisfied by all optimal tax schemes for the individuals threatening to emigrate.

[4] If the participation constraints (PC) were not type-dependent, it would be necessary and sufficient to check that they are satisfied at $\underline{\theta}$ since (FOIC) ensures that the optimal utility path is non-decreasing.

For this purpose, let I be an interval of positive length where (4) is active. By definition, for $\theta \in I$, $R(\theta) \equiv 0$ and thus $V_A'(\theta) = V_B'(\theta) - c'(\theta)$. Hence the rate of increase of the indirect utility the government has to give to the individuals so that they reveal their private information, is equal to the slope of the reservation utility on I. In addition, employing (FOIC) and rearranging yield

$$z_A(\theta) = -\theta \frac{V_B'(\theta) - c'(\theta)}{u_z'(x_A(\theta), z_A(\theta); \theta)} \text{ for } \theta \in I, \tag{10}$$

and by differentiation,

$$z_A'(\theta) = \frac{\left[V_B'(\theta) - c'(\theta)\right]\left\{\theta\left(u_{xz}'' x_A' + u_{\theta z}''\right) - \left(1 + \theta \frac{V_B''(\theta) - c''(\theta)}{V_B'(\theta) - c'(\theta)}\right) u_z'\right\}}{\left(u_z'\right)^2 - \theta\left(V_B'(\theta) - c'(\theta)\right) u_{zz}''} \text{ for } \theta \in I. \tag{11}$$

The second-order condition for incentive compatibility (SOIC') can only be satisfied on I if the curly bracket in (11) is non-negative. When preferences are separable ($u_{xz}'' = 0$), one gets

$$z_A'(\theta) \geq 0 \Leftrightarrow \frac{\theta u_{\theta z}''}{u_z'} \leq \left(1 + \theta \frac{V_B''(\theta) - c''(\theta)}{V_B'(\theta) - c'(\theta)}\right), \tag{12}$$

the LHS of which is negative since $u_{\theta z}'' > 0$ and $u_z' < 0$.

Property 1. *Let preferences be separable ($u_{xz}'' = 0$) and consider an interval I of positive length where (PC) is active. Then, there is* no bunching *on I when*

$$\theta \frac{V_B''(\theta) - c''(\theta)}{V_B'(\theta) - c'(\theta)} > -1 \text{ for } \theta \in I. \tag{13}$$

Condition (13) states that the elasticity of the marginal reservation utility to the wage rate, evaluated at θ, is greater than -1 for every $\theta \in I$. To have further insight, we now turn to quasilinear-in-consumption preferences,

$$u(x, z; \theta) = x - v(z/\theta), \text{ with } v(\ell) = \ell^{1+1/e} / (1 + 1/e). \tag{14}$$

The Hicksian elasticity of labour supply is thus constant ($e^H(\theta) = e$) and, by (10),

$$\ell_A(\theta) = \theta^{\frac{e}{1+e}} \left[\theta^e - c'(\theta)\right]^{\frac{e}{1+e}}. \tag{15}$$

Since $V_A(\theta) = \theta \ell_A(\theta) - T(\theta \ell_A(\theta)) - v(\ell_A(\theta))$ and $V_B(\theta) = \theta^{1+e} / (1+e)$, $R(\theta) = 0$ on I results in

$$T(z_A(\theta)) = \theta \ell_A(\theta) - V_B(\theta) + c(\theta) - v(\ell_A(\theta)). \tag{16}$$

Property 2. *Let preferences be quasilinear in consumption, e be the constant elasticity of labour* *supply and I be an interval of positive length where (PC) is active. Then,*

$$T(z_A(\theta)) = \theta(\theta^e - c'(\theta)) \left[\theta^{\frac{e}{1+e}} (\theta^e - c'(\theta))^{-\frac{1}{1+e}} - \frac{e}{1+e}\right] - \frac{\theta^{1+e}}{1+e} + c(\theta). \tag{17}$$

On I, the optimal tax liability depends on productivity, on the cost of migration and its slope as well as on the elasticity of labour supply. From now on, we concentrate on migration cost functions $c(\theta)$ which satisfy (13); hence there is no bunching on I. The first-order condition for individual utility maximization in A yields $\ell_A(\theta) = \theta^e [1 - T']^e$. Combining this expression of $\ell_A(\theta)$ with (15) and solving for T', the following property is obtained.

Property 3. *Let preferences be quasilinear in consumption, e be the constant elasticity of labour supply, $c(\theta)$ satisfy (13), and I be an interval of positive length where (PC) is active. Then,*

$$\frac{T'}{1-T'} = \theta^{\frac{e}{1+e}} [\theta^e - c'(\theta)]^{-\frac{1}{1+e}} - 1 \text{ for } \theta \in I. \tag{18}$$

In this case, the optimal marginal tax rate on I depends on the productivity level, on the elasticity of labour supply and on the slope of the costs of migration. Its sign is as follows.

Property 4. *Consider the same situation as in Property 3. Then,*

$$T'(z_A(\theta)) \gtreqless 0 \Leftrightarrow c'(\theta) \gtreqless 0 \text{ for } \theta \in I. \tag{19}$$

When the costs of migration are *non-increasing*, the theorem stating that the optimal tax function is strictly increasing at *all* income levels (Seade, 1982) does no longer hold. When the costs of migration are *strictly decreasing* in productivity, the optimal marginal tax rates faced by the individuals threatening to emigrate are *strictly negative*[5]. This property contrasts with two results obtained in closed economy, stating that: (i) the optimal marginal tax rates are non-negative (Mirrlees, 1971); (ii) the optimal marginal tax rate is zero at the top (Sadka, 1976, Seade, 1977). The next corollaries of Property 4 provide further details about these significant changes.

The first one considers constant migration costs on I. By Property 4, $\ell_A(\theta) = \theta^e$ on I. Moreover, $V_A = V_B - c = \theta^{1+e}/(1+e) - c$. Then, by (16), $T(z_A(\theta)) = c(\theta)$ on I.

Property 5. *Consider the same situation as in Property 3 and let $c'(\theta) = 0$ on I. Then, the optimal tax function has a flat section corresponding to potentially mobile individuals paying* taxes equal to their positive costs of migration*.*

Hence, because of the threat of migration, the optimal tax schedule becomes regressive: highly skilled individuals for whom the participation constraints are binding pay less taxes in proportion to gross income than lower skilled individuals. The situation is even more acute when the costs of migration are strictly decreasing.

[5] An example of optimal income tax schedule with strictly negative marginal tax rates is provided in the simulation section.

Property 6. *Consider the same situation as in Property 3 and let* $c'(\theta) < 0$ *on* I*. Then, the optimal average tax rate and the optimal tax function are strictly decreasing in productivity on* I.[6]

Here, progressivity of the optimal tax schedule does not only collapse because of potential mobility; the *tax liability* itself becomes strictly decreasing. This means that there are middle-skilled individuals insufficiently talented to leave the country which pay higher taxes than more productive individuals. This is a *second-best counterpart of the curse of the middle-skilled*, in which taxes replace utility levels.

4.3. National Criterion

We study the impact of the threat of migration on the optimum tax scheme in A when A's government adopts the National criterion.

In solving Problem 4, we assume that the adjoint variables have a finite number of jump discontinuities and are $\mathcal{C}^1$ elsewhere. For later reference, we call ι the adjoint variable associated with (FOIC) and $\pi' \geq 0$ the Lagrange multiplier of (PC), which corresponds to the shadow price of a marginal increase in the reservation utility at θ.

In order to characterize the optimal income tax schedule, it is useful to introduce a few additional definitions We denote by π the shadow price of a uniform marginal increase in the reservation utility for all $\theta' \geq \theta$. By definition, it is the non-decreasing function, with derivative π' almost everywhere, satisfying

$$\pi(\theta) := \pi\left(\overline{\theta}\right) - \int_{\theta}^{\overline{\theta}} \pi'(\tau)\, d\tau. \tag{20}$$

We also call e^H and e^M the Hicksian and Marshallian elasticities of labour supply with respect to the net-of-tax wage rate. Moreover, as shown by Saez (2001), the magnitude of the uncompensated behavioural response of the τ-individuals to a small increase in the marginal tax rate at θ, with $\theta < \tau$, is summarized by

$$\Psi_{\theta\tau} = \exp \int_{\theta}^{\tau} \left(1 - \frac{e^M(\delta)}{e^H(\delta)}\right) \frac{z'_A(\delta)}{z_A(\delta)} d\delta.$$

Proposition 4. *Under the National criterion and in the absence of bunching, the optimal marginal tax rates are given by*

$$\frac{T'(z_A(\theta))}{1 - T'(z_A(\theta))} = A(\theta)\, B(\theta)\, C(\theta)\,, \qquad \textit{for } \theta < \overline{\theta}, \tag{21}$$

[6] The fact that the optimal average tax rate $T(z_A(\theta))/z_A(\theta)$ is strictly decreasing on I follows from $T'(z_A(\theta)) < 0$ and $z'_A(\theta) > 0$ on I.

where

$$A(\theta) := \frac{1+e^{M}(\theta)}{e^{H}(\theta)},\ B(\theta) := B_1(\theta) - B_2(\theta)\ \textit{and}\ C(\theta) := \frac{1-F(\theta)}{\theta f(\theta)},$$

with

$$\begin{aligned} B_1(\theta) &:= \frac{1}{1-F(\theta)}\int_{\theta}^{\overline{\theta}}\left[1-\frac{\phi'_{\rho}(V_A(\tau))\,u'_x(x_A,z_A;\tau)}{\gamma}\right]\Psi_{\theta\tau}dF(\tau),\\ B_2(\theta) &:= \frac{1}{1-F(\theta)}\left[\int_{\theta}^{\overline{\theta}}\frac{\pi'(\tau)\,u'_x(x_A,z_A;\tau)}{\gamma}\Psi_{\theta\tau}d\tau+\frac{\iota(\overline{\theta})\,u'_x(x_A,z_A;\theta)}{\gamma}\right],\\ &\pi'(\tau)\geq 0\ \ (=0\ \textit{if}\ R(\tau)>0)\ \ \textit{and}\ \iota(\overline{\theta})\geq 0\ \ \left(=0\ \textit{if}\ R(\overline{\theta})>0\right). \end{aligned}$$

At the top,

$$\frac{T'\left(z_A\left(\overline{\theta}\right)\right)}{1-T'\left(z_A\left(\overline{\theta}\right)\right)}=\frac{A\left(\overline{\theta}\right)}{\overline{\theta}f\left(\overline{\theta}\right)}\frac{\iota\left(\overline{\theta}\right)u'_x\left(x_A,z_A;\overline{\theta}\right)}{\gamma}\leq 0\ \ \left(=0\ \textit{if}\ R\left(\overline{\theta}\right)>0\right). \tag{22}$$

Proof. See A.2 in the Appendix. □

Proposition 4 extends Mirrlees's (1971) optimal income tax formula to take the threat of migration into account, using behavioural elasticities as in Saez (2001). It reflects the trade-off between efficiency and equity when the government has decided to maintain the national productive capacity to the maximum in preventing its citizens from leaving the country. $A(\theta)$ and $C(\theta)$ are the usual efficiency and demographic factors, respectively. However, the value of $A(\theta)$ is usually not the same whether the individuals can or cannot vote with their feet since it depends on gross income which is endogenous. The factor $B(\theta)$, which combines efficiency and equity, is the only factor which is not written as in Mirrlees's formula, in which the RHS of (21) reduces to $A(\theta)B_1(\theta)C(\theta)$. As previously stated, the optimal marginal tax rates can be strictly negative at the top, and therefore non-positive at interior points of the schedule.

We now turn to the different channels captured in Formula (21). We consider a small tax reform perturbation around the optimal income tax schedule. A small increase dT for gross income between z and $z+dz$ has *four* effects on social welfare. Three effects are already observed in closed economy and have been thoroughly examined by Saez (2001).

- *The three "usual" effects* allow us to grasp $A(\theta)$, $B_1(\theta)$ and $C(\theta)$.

First, the local increase in the marginal rate of tax mechanically results in individuals with gross income greater than z paying additional taxes. Second, the elasticity response from the taxpayers with gross income between z and $z+dz$ decreases their labour supply and reduces tax revenue. Third, under Assumption 2, the increase in taxes paid by these individuals has an income effect, leading them to work more, which is good for tax receipts.

- *The new* participation *effect* illuminates $B_2(\theta)$.

The tax reform perturbation *mechanically* results in an increase in taxes paid by all individuals with gross income strictly above z. Consequently, those among them for whom the participation constraints were already active receive now a utility level below their reservation utility. Then the participation constraints (PC) are no longer satisfied. So, these individuals have to be compensated for the increase in taxes they face.

We first examine the compensation for the individuals whose gross income is strictly below $z_{\overline{\theta}}$. The *compensation effect* leads A's government to *totally* compensate them for staying in A. Each of them is thus given $u'_x(x_A, z_A; \tau) \times dTdz$ additional units of utility. Since $\pi'(\tau)$ is the shadow price of the participation constraint at τ and γ the Lagrange multiplier of the tax revenue constraint (TR), the cost in terms of social welfare of the compensation of the τ-individuals amounts to

$$\pi'(\tau) \times \frac{u'_x(x_A, z_A; \tau)}{\gamma} \times dTdz. \tag{23}$$

The compensation effect combines with the usual *income effect*. Because leisure is a normal good under Assumption 2, the increase in the tax burden paid by all individuals with income greater than z induces them to work more. This allows A's government to increase the taxes they face. As a result, it is not required to compensate the potentially mobile individuals as high as the increase in taxes they face. We know from Saez (2001) that the magnitude of the uncompensated behavioural response is summarized by $\Psi_{\theta\tau} \geq 1$, which converts the social marginal utility of consumption of the τ-individuals, $u'_x(x_A, z_A; \tau)$, into that of the θ_z-individuals, $u'_x(x_A, z_A; \theta_z)$. Using (23), the social cost of the compensation of the τ-individuals, including income effects, is

$$\pi'(\tau)\,\Psi_{\theta\tau} \times \frac{u'_x(x_A, z_A; \tau)}{\gamma} \times dTdz. \tag{24}$$

By integration of (24), we get the cost of compensating the individuals with productivity below $\overline{\theta}$. For the individuals on the upper bound of the population, the social cost is directly obtained as

$$\left.\frac{\partial W^N_{A,\rho}}{\partial V_A}\right|_{\overline{\theta}} \times \frac{u'_x(x_A, z_A; \theta_z)}{\gamma} \times dTdz = \iota\left(\overline{\theta}\right) \times \frac{u'_x(x_A, z_A; \theta_z)}{\gamma} \times dTdz. \tag{25}$$

Finally, by (24) and (25), the average social cost of the compensation of all potentially mobile individuals with gross income above z is

$$\frac{1}{1-F(\theta_z)}\left[\int_{\theta_z}^{\overline{\theta}} \frac{\pi'(\tau)\,u'_x(x_A, z_A; \tau)}{\gamma}\Psi_{\theta_z\tau}d\tau + \frac{\iota\left(\overline{\theta}\right)u'_x(x_A, z_A; \theta_z)}{\gamma}\right] \times dTdz$$
$$= B_2(\theta_z) \times dTdz. \tag{26}$$

$B_2(\theta_z)$ is positive as soon as there are individuals with productivity above θ_z for whom the participation constraints are binding. This term goes therefore against progressivity on a range of gross income levels preceding that on which individuals hesitate to leave the country. This is because increasing the marginal tax rates at θ_z makes the compensation of all more productive individuals threatening to emigrate more expensive in terms of social welfare.

Eventually, the participation effect results in the adjustment of the optimal marginal tax rates to make the *average* tax rates compatible with the participation constraints. In consequence, A's government should be particularly cautious about increasing marginal tax rates even at productivity levels where individuals do not hesitate to vote with their feet.

4.4. Citizen and Resident Criteria

Under the National criterion, the whole population is constrained to stay in A. We now relax this constraint to examine whether keeping everybody in the home country is not too expensive in terms of social welfare. For this purpose, we separate Problem 4 into two subproblems to determine the optimal $\widehat{\theta}$. In the first subproblem, $\widehat{\theta}$ is arbitrarily chosen by A's government.

Subproblem 1. *Given* $\widehat{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]$, *find* (x_A, z_A) *to maximize* $W^i_{A,\rho}\left(\widehat{\theta}\right)$, $i = \{C, R\}$, *subject to* $(FOIC)$, $(SOIC')$, (PC), (TR).

Let $\mathcal{W}^i_{A,\rho}\left(\widehat{\theta}\right)$ be the social value function of this subproblem, $\iota^i_{\widehat{\theta}}(\theta)$ the shadow price of incentive-compatibility constraint (FOIC), and $\pi^i_{\widehat{\theta}}(\theta)$ the shadow price of a uniform marginal increase in the reservation utility for all $\theta' \geq \theta$. The solution in $\widehat{\theta}$ to Problem 4 is then obtained as:

Subproblem 2. *Find* $\widehat{\theta}^i \in \left[\underline{\theta}, \overline{\theta}\right]$ *solution to* $\max_{\widehat{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]} \mathcal{W}^i_{A,\rho}\left(\widehat{\theta}\right)$, $i = \{C, R\}$.

Subproblem 1 is a generalization of the second-best National problem where the upper productivity in A is exogenously given. Consequently, the optimal marginal tax rates share qualitative properties irrespective of the chosen social criterion. The only differences come from changes in the size of A's resident population.

Property 7. *Proposition 4 applies for the Citizen and Resident criteria provided:*
(i) $\overline{\theta}$ *is replaced by* $\widehat{\theta}^i$ *and* $1 - F(\theta)$ *by* $F\left(\widehat{\theta}^i\right) - F(\theta)$, $i = \{C, R\}$;
(ii) in $B_1(\theta)$, $\phi'_\rho(V_A)$ *is divided by* $F\left(\widehat{\theta}^R\right)$ *for the Resident criterion.*

Proof. See A.3 in the Appendix. □

We are now prepared to examine the allocation of individuals between A and B resulting from the implementation of the Citizen and Resident optimal income tax schedules. For all $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$, the $\widehat{\theta}$-individuals are indifferent between living in A or B. Let us assume $\widehat{\theta} < \overline{\theta}$. Hence, individuals with productivity above $\widehat{\theta}$ are in B. Making them relocate to A requires adjustments to prevent them from imitating less productive individuals. It also brings about a *new upward mimicking behaviour*. Indeed, A's residents can now have an incentive to mimicking them since they have the most appealing outside options. We now explain the different terms for the Citizen criterion. The terms for the Resident criterion are the same up to a scale factor $F\left(\widehat{\theta}\right)$.

The upward mimicking behaviour is crucial to understanding the interactions between the incentive-compatibility conditions and the type-dependent participation constraints. In closed economy, individuals have the usual incentive to understate their productivity θ to obtain greater social benefit whilst enjoying more leisure.[7] When type-dependent participation constraints are taken into account, the individuals may also be tempted to overstate their productivity θ, in working harder, to obtain greater compensation for staying in A. This behaviour reflects *countervailing incentives*. An asymmetry in terms of informational constraints between the individuals with productivities below $\widehat{\theta}$ and the $\widehat{\theta}$-individuals may therefore arise. Indeed, contrary to the former, the latter can only have the usual incentives. The cost of making the θ-individuals reveal their private information, represented by $\iota^i_{\widehat{\theta}}(\theta) \geq 0$, can thus have a downward jump discontinuity at $\widehat{\theta}$. However, making them reveal their private knowledge requires the gap between $V_B\left(\widehat{\theta}\right) - c\left(\widehat{\theta}\right)$ and $V_A(\theta)$ to be reduced. This increase in $V_A(\theta)$ reduces the social cost of a uniform increase in the reservation utility at θ and above, which is captured by $\pi(\theta)$. This effect stops suddenly when θ tends to $\widehat{\theta}$. Consequently, an *upward* jump discontinuity in $\pi^i_{\widehat{\theta}}$ corresponds to the *downward* jump discontinuity in $\iota^i_{\widehat{\theta}}$ at $\widehat{\theta}$. It turns out that these discontinuities have the same magnitude.

Property 8. *At* $\widehat{\theta}$,

$$\iota^i_{\widehat{\theta}}\left(\widehat{\theta}^-\right) - \iota^i_{\widehat{\theta}}\left(\widehat{\theta}\right) = \pi^i_{\widehat{\theta}}\left(\widehat{\theta}\right) - \pi^i_{\widehat{\theta}}\left(\widehat{\theta}^-\right) \geq 0 \left(= 0 \text{ if (PC) inactive at } \widehat{\theta}\right), \tag{27}$$

where $\iota^i_{\widehat{\theta}}\left(\widehat{\theta}^-\right) := \lim_{\theta \underset{\to}{\leq} \widehat{\theta}} \iota^i_{\widehat{\theta}}(\theta)$.

Proof. See (A.29) in the Appendix. □

A variational analysis provides insights into the costs and benefits of the presence in the home country of the marginal $\widehat{\theta}$-individuals.

Proposition 5. *Given* $\widehat{\theta} < \overline{\theta}$,

$$\frac{\partial \mathcal{W}^C_{A,\rho}\left(\widehat{\theta}\right)}{\partial \widehat{\theta}} = \gamma T\left(z_A\left(\widehat{\theta}\right)\right) f\left(\widehat{\theta}\right) + \iota\left(\widehat{\theta}^-\right) V'_A\left(\widehat{\theta}\right) - \left[\iota\left(\widehat{\theta}^-\right) - \iota\left(\widehat{\theta}\right)\right] \left|R'\left(\widehat{\theta}\right)\right|. \tag{28}$$

[7] In the discrete population model of Guesnerie and Seade (1982), a sufficient condition for incentive-compatibility of the tax scheme is that only the downward adjacent incentive-compatibility constraints are binding (see also Weymark (1986, 1987)). Hellwig (2007) has recently established, in both discrete and continuous models that under "desirability of redistribution" only the downward incentive-compatibility constraints are binding in a closed economy.

and

$$\frac{\partial \mathcal{W}^R_{A,\rho}\left(\widehat{\theta}\right)}{\partial\widehat{\theta}} = \frac{\gamma T\left(z_A\left(\widehat{\theta}\right)\right) f\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)} + \frac{\left[\phi_\rho\left(V_A\left(\widehat{\theta}\right)\right) - \mathcal{W}^R_{A,\rho}\left(\widehat{\theta}\right)\right] f\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)} + \frac{\iota\left(\widehat{\theta}^-\right) V_A'\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)} - \frac{\left[\iota\left(\widehat{\theta}^-\right) - \iota\left(\widehat{\theta}\right)\right]\left|R'\left(\widehat{\theta}\right)\right|}{F\left(\widehat{\theta}\right)}. \tag{29}$$

Proof. See A.3 in the Appendix. □

In (28) and (29), the first terms of the RHS corresponds to the contribution of the marginal individuals to tax revenue. The second term in (29) compares the marginal and average social welfare. In contrast to the first-best, this term cannot be negative because (FOIC) ensures that the indirect utility is non-decreasing. The last two terms in (28) and (29) are specific to the second-best setting. They reflect the marginal costs and benefits with regard to incentives of the presence in A of the $\widehat{\theta}$-individuals.

First, there is a cost due to countervailing incentives. Indeed, individuals to the very left of $\widehat{\theta}$ have the possibility to mimic the $\widehat{\theta}$-individuals to benefit from their higher outside options. They can therefore claim an increase in their utility at the margin, equal to $V_B'\left(\widehat{\theta}\right) - c'\left(\widehat{\theta}\right) - V_A'\left(\widehat{\theta}\right) = \left|R'\left(\widehat{\theta}\right)\right|$. The shadow price of this upward-mimicking behaviour is given by the excess of $\iota\left(\widehat{\theta}^-\right)$ over $\iota\left(\widehat{\theta}\right)$, which is non-negative by Property 8. The corresponding marginal social cost is thus

$$\left[\iota\left(\widehat{\theta}^-\right) - \iota\left(\widehat{\theta}\right)\right]\left|R'\left(\widehat{\theta}\right)\right| \geq 0. \tag{30}$$

If $\widehat{\theta}$ is a non-isolated point where (PC) is active, this cost is nil because $R'\left(\widehat{\theta}\right) = 0$.

Second, when countervailing incentives arise, the individuals to the very left of $\widehat{\theta}$ have greater utility. They thus are less inclined to mimic less productive individuals. The slope of the indirect utility V_A' at $\widehat{\theta}$ required for them to reveal their type truthfully is therefore reduced at the margin. Since $\iota\left(\widehat{\theta}^-\right)$ is the shadow price of (FOIC), the marginal social benefit of this slackening of the downward incentive compatibility constraints is $\iota\left(\widehat{\theta}^-\right) V_A'$. This is the second implication of countervailing incentives due to the presence of the marginal individuals. Finally, the *net* marginal social cost incurred to restore incentives at the top amounts to

$$\left[\iota\left(\widehat{\theta}^-\right) - \iota\left(\widehat{\theta}\right)\right]\left|R'\left(\widehat{\theta}\right)\right| - \iota\left(\widehat{\theta}^-\right) V_A'\left(\widehat{\theta}\right). \tag{31}$$

In the light of Proposition 5, the marginal individuals positively contribute to social welfare when (i) they pay positive taxes and (ii) either ι is continuous at $\widehat{\theta}$ or $\widehat{\theta}$ is a non-isolated point where (4) is active. These situations correspond to cases where the usual downward mimicking behaviour predominates for highly skilled individuals. However, they do not exhaust all possible cases. In particular, the trade-off between maintaining national capacity to the maximum and sustaining the redistribution programme become more complex when $\widehat{\theta}$ is an isolated point where (4) is active and ι has a jump discontinuity. In the solution to Subproblem 2, $\partial \mathcal{W}_{A,\rho}^{i}\left(\widehat{\theta}\right)/\partial\widehat{\theta}$ must be non-negative, $i = \{C, R\}$. Otherwise, emigration of the most productive individuals would increase social welfare. This leads to the following sufficient condition under which emigration of the most productive individuals initially living in A is socially optimal.

Proposition 6. *Consider the National optimal allocation. Assume $\overline{\theta}$ is an isolated point where (PC) is active and ι has a jump discontinuity. Then, it is optimal to allow some highly skilled to emigrate if*
(i) Citizen case ($\widehat{\theta}^{C} < \overline{\theta}$):

$$\gamma T\left(z_A\right) f < \left[\iota\left(\overline{\theta}^{-}\right) - \iota\right] |R'| - \iota\left(\overline{\theta}^{-}\right) V_A'; \tag{32}$$

(ii) Resident case ($\widehat{\theta}^{R} < \overline{\theta}$):

$$\gamma T\left(z_A\right) f + \left[\phi_\rho\left(V_B - c\right) - W_{A,\rho}^{N}\right] f < \left[\iota\left(\overline{\theta}^{-}\right) - \iota\right] |R'| - \iota\left(\overline{\theta}^{-}\right) V_A', \tag{33}$$

where all functions are evaluated at $\overline{\theta}$ except otherwise stated.

Proof. See A.3 in the Appendix. □

This proposition exploits the fact that both Citizen and Resident criteria coincide with the National one when $\widehat{\theta} = \overline{\theta}$. It is then possible to apply the previous cost/benefit analysis to the National criterion. (32) and (33) correspond to cases where $\partial \mathcal{W}_{A,\rho}^{i}\left(\widehat{\theta}\right)/\partial\widehat{\theta}_{\widehat{\theta}=\overline{\theta}} < 0$, $i = \{C, R\}$.

The basic intuition behind Proposition 6 is that the choice of $\widehat{\theta}$ by A's government can be regarded as a means of revealing private information. Indeed, if A's government designs a tax policy such that the individuals with productivity greater than $\widehat{\theta}$ do not receive in A their reservation utility, it *knows* that $\widehat{\theta}$ is the maximum productivity in its resident population and consequently that the individuals with productivity greater than $\widehat{\theta}$ are in B. Proposition 6 tells us in which cases using this means improves social welfare.

5. CONCLUSION

This paper provides a first example of the introduction of type-dependent participation constraints in the optimal income tax framework. These constraints interact with the standard constraints in a non-trivial way and make the structure of the mimicking behaviour more complex than in closed economy. Since they induce substantial changes, it might be worth introducing them in other classic models of taxation theory, like those devoted to capital taxation.

In this extended framework, the issue of the optimal allocation of individuals between the home country and abroad is embedded in the optimal income tax problem. Consequently, a new trade-off between maintaining the redistribution programme and preserving national productive capacities adds to the traditional trade-off between equity and efficiency. Emigration of highly skilled individuals should not always be prevented to maximize social welfare.

Key qualitative features of the optimal income tax policy obtained in closed economy do no longer hold. The participation effect does not only favour a decrease in the optimal marginal tax rates; it can also make them strictly negative. Consequently, the optimal average tax rates as well as the optimal tax liabilities can be decreasing. More generally, our results convey a curse of the middle-skilled workers: this curse is expressed in terms of utility in the first-best, and in terms of taxes in the second-best.

A. APPENDIX

A.1. First-Best

Proof of Proposition 1. Let π' and γ be the Lagrange multipliers of (PC) and (TR) respectively. Under Assumption 1, the solution is interior and the SOC are satisfied. Hence, the necessary and sufficient FOC are

$$\left(\phi'_\rho + \pi'\right) U'_x = \gamma \text{ and } \left(\phi'_\rho + \pi'\right) U'_\ell = -\gamma\theta, \tag{A.1}$$

with

$$\pi' \geq 0,\ U(x,\ell) - V_B + c \geq 0,\ \pi'\left[U(x,\ell) - V_B + c\right] = 0,\ \forall\theta \in \left[\underline{\theta}, \overline{\theta}\right]. \tag{A.2}$$

Since $\phi'_\rho > 0$, (A.1) and (A.2) implies $\gamma > 0$. The following Lemma is shown in Mirrlees (1974) for $\rho = 0$.

Lemma 1. *Let J be a non-empty open interval where $\pi' \equiv 0$. Then for all $\theta \in J$, (a) $V'_A(\theta) < 0$ when $0 \leq \rho < \infty$, (b) $V'_A(\theta) = 0$ when $\rho \to \infty$.*

Proof. Assumption 2 holds if and only if $d\ell/dT > 0$. Since $\pi' \equiv 0$, applying the implicit function theorem to (A.1) yields

$$\frac{d\ell(\theta)}{dT(\theta)} = -\frac{\theta U''_{xx} + U''_{x\ell}}{\theta\left[2U''_{x\ell} - \frac{U'_\ell}{U'_x}U''_{xx} - \frac{U'_x}{U'_\ell}U''_{\ell\ell}\right]}, \tag{A.3}$$

where the square bracket is strictly positive because U is strictly quasi-concave under Assumption 1. Therefore, Assumption 2 is equivalent to

$$\theta U''_{xx} + U''_{x\ell} < 0. \tag{A.4}$$

(a) Since $\pi' \equiv 0$, (A.1) yields $\theta = U'_x/U'_\ell$ and, by differentiation,

$$\begin{pmatrix} U''_{xx} - \gamma\rho U^{\rho-1}U'_x & U''_{x\ell} - \gamma\rho U^{\rho-1}U'_\ell \\ U''_{x\ell} + \theta\gamma\rho U^{\rho-1}U'_x & U''_{\ell\ell} + \theta\gamma\rho U^{\rho-1}U'_\ell \end{pmatrix} \begin{pmatrix} x'(\theta) \\ \ell'(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ -\gamma U^\rho \end{pmatrix}. \tag{A.5}$$

As $|A| > 0$ under Assumption 1,

$$\begin{cases} x'(\theta) = \gamma U^\rho\left[U''_{x\ell} - \gamma\rho U^{\rho-1}U'_\ell\right] / |A| \\ l'(\theta) = -\gamma U^\rho\left[U''_{xx} - \gamma\rho U^{\rho-1}U'_x\right] / |A| \end{cases} \tag{A.6}$$

from which

$$V'_A(\theta) = u'_x x'(\theta) + u'_\ell \ell'(\theta) = -\gamma U^\rho U'_\ell\left[U''_{xx} - U''_{x\ell}U'_x/U'_\ell\right] / |A|, \tag{A.7}$$

which has the same sign as $U''_{xx} - U''_{x\ell}U'_x/U'_\ell$, i.e. as $\theta U''_{xx} - U''_{x\ell}$. Hence, by (A.4), $V'_A(\theta) < 0$.
(b) The result directly follows from duality. □

Step 1: The existence of θ^* is obvious. Indeed, since $V_A^{c\ell}\left(\overline{\theta}\right) < V_B\left(\overline{\theta}\right) - c\left(\overline{\theta}\right)$, the closed-economy solution violates (4); so there are θ such that $\pi' > 0$ at the solution to Problem 1.

Step 2: $\pi'(\theta) > 0$ for all $\theta > \theta^*$.

By (A.1), $\pi'(\theta) = \gamma/U'_x - \phi'_\rho$, which implies under Assumption 1 and the continuity of T, the continuity of π'. Assume $\theta' := \min\left\{\theta \in \left[\theta^*, \overline{\theta}\right] : \pi'(\theta) = 0\right\}$ exists. Then, by continuity of π', there exists $\theta'' > \theta'$ such that $\pi' = 0$ on $[\theta', \theta'']$. By continuity of R, $R(\theta') = 0$. On $[\theta', \theta'']$, $V'_A \leq 0$ by Lemma 1 and $V'_B - c' > 0$ under Assumption 3. Then $R < 0$ for $\theta \in (\theta', \theta'')$, contradicting (PC). Hence, θ' does not exist. □

A.2. Second-Best: National Criterion

Proof of Proposition 4. z_A is control variable; V_A and $G(\theta) := \int_{\underline{\theta}}^{\theta} T(z_A(\tau))\,dF(\tau)$ are state variables. Since $T := z_A - x_A$, Leibnitz's rule yields $G'(\theta) = (z_A(\theta) - x_A(\theta))\,f(\theta)$. The isoperimetric constraint (TR) is taken into account through G' and the boundary conditions $G(\underline{\theta}) = 0$ and $G\left(\overline{\theta}\right) = 0$. It is not necessary to take x_A explicitly into account because it is uniquely determined by V_A and z_A. Let $x_A = h(V_A, z_A; \theta)$; differentiating shows $\partial x_A/\partial V_A = 1/u'_x$ and $\partial x_A/\partial z_A = s$. The Hamiltonian and Lagrangian are respectively

$$\begin{aligned} H^N &= \phi_\rho(V_A)\,f + \iota u'_\theta + \gamma(z_A - x_A)\,f, \\ L^N &= H^N + \pi' R. \end{aligned}$$

As $\partial u'_\theta/\partial z_A = u''_{\theta z} + s u''_{\theta x} = -u'_x s'_\theta$, and $\partial u'_\theta/\partial V_A = u''_{\theta x}/u'_x$, necessary conditions are:

$$\partial H^N/\partial z_A = 0 \Leftrightarrow \iota u'_x s'_\theta - \gamma(1-s)\,f = 0, \tag{A.8}$$

$$\partial L^N/\partial V_A = -\iota' \Leftrightarrow \iota'(\theta) = -\phi'_\rho(V_A)\,f - \iota u''_{\theta x}/u'_x - \pi' + \gamma f/u'_x, \tag{A.9}$$

$$\partial L^N/\partial G = -\gamma' \Leftrightarrow \gamma' = 0, \tag{A.10}$$

$$\iota\left(\overline{\theta}\right) \geq 0 \ \left(= 0 \text{ when } R\left(\overline{\theta}\right) > 0\right), \tag{A.11}$$

$$\iota(\underline{\theta}) \leq 0 \ \left(= 0 \text{ when } R(\underline{\theta}) > 0\right), \tag{A.12}$$

$$\pi'(\theta) \geq 0, \ R(\theta) \geq 0, \ \pi'(\theta)\,R(\theta) = 0, \tag{A.13}$$

$$\iota\left(\theta_j^-\right) - \iota\left(\theta_j^+\right) = \pi\left(\theta_j^+\right) - \pi\left(\theta_j^-\right) \geq 0 \ (= 0 \text{ if } R(\theta_j) > 0). \tag{A.14}$$

$\gamma(\theta)$ is constant, equal to $\gamma > 0$. As $s = 1 - T'$, $T' = \iota u'_x s'_\theta/(\gamma f)$ by (A.8). In addition, using basic calculus, $\left[1 + e^M(\theta)\right]/e^H(\theta) = -\theta s'_\theta/s$. Hence,

$$\frac{T'}{1-T'} = -\frac{\iota u'_x}{\gamma \theta f}\frac{1 + e^M(\theta)}{e^H(\theta)}. \tag{A.15}$$

When $\theta = \overline{\theta}$, (A.15) and (A.11) yield (22). When $\theta < \overline{\theta}$, (A.15) can be rewritten as

$$\frac{T'}{1-T'} = -\frac{\iota u'_x}{\gamma(1 - F(\theta))}\frac{1 + e^M(\theta)}{e^H(\theta)}\frac{1 - F(\theta)}{\theta f(\theta)}, \tag{A.16}$$

If $(.;\tau)$ means evaluation at $(x_A(\tau), z_A(\tau);\tau)$, integrating (A.9) between θ and $\overline{\theta}$ yields

$$\iota(\theta) = \iota\left(\overline{\theta}\right) + \int_{\theta}^{\overline{\theta}} \left(\phi'_\rho\left(V_A(\tau)\right) f(\tau) + \pi'(\tau) - \frac{\gamma f(\tau)}{u'_x(.;\tau)}\right) \widetilde{\Psi}_{\theta\tau} d\tau, \tag{A.17}$$

with $\widetilde{\Psi}_{\theta\tau} := \exp \int_{\theta}^{\tau} u''_{\theta x}(.;\tau') / u'_x(.;\tau') d\tau'$. The following relation has been proved by Saez (2001, p. 227):

$$\Psi_{\theta\tau} := \frac{u'_x(.;\theta)}{u'_x(.;\tau)} \widetilde{\Psi}_{\theta\tau} = \exp \int_{\theta}^{\tau} \left(1 - \frac{e^M(\tau')}{e^H(\tau')}\right) \frac{z'_A(\tau')}{z_A(\tau')} d\tau'. \tag{A.18}$$

Using (A.17) and (A.18),

$$-\frac{\iota(\theta) u'_x(.;\theta)}{\gamma} = \int_{\theta}^{\overline{\theta}} \left[1 - \left(\phi'_\rho\left(V_A(\tau)\right) + \frac{\pi'(\tau)}{f(\theta)}\right) \frac{u'_x(.;\tau)}{\gamma}\right] \Psi_{\theta\tau} dF(\tau) - \frac{\iota\left(\overline{\theta}\right) u'_x(.;\theta)}{\gamma}, \tag{A.19}$$

and plug the obtained expression in (A.16). □

A.3. Second-Best: Citizen and Resident Criteria

Proof of Property 7. (a) *Citizen criterion.* By definition, $W^C_{A,\rho}\left(\widehat{\theta}\right)$ is maximum when $\widehat{\theta} = \widehat{\theta}^C$, i.e. when $W^C_{A,\rho}\left(\widehat{\theta}^C\right)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (PC), (TR). The FOC are the same as (A.8)–(A.14), except that $\overline{\theta}$ is replaced by $\widehat{\theta}^C$. We then proceed as in the proof of Proposition 4.

(b) *Resident criterion.* By definition, $W^R_{A,\rho}\left(\widehat{\theta}\right)$ is maximum when $\widehat{\theta} = \widehat{\theta}^R$, i.e. when $W^R_{A,\rho}\left(\widehat{\theta}^R\right)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (PC), (TR). The FOC are the same as (A.8)–(A.14), except that (i) $\overline{\theta}$ is replaced by $\widehat{\theta}^R$ and (ii) $\phi'_\rho(V_A)$ is divided by $F\left(\widehat{\theta}^R\right)$. We then proceed as in the proof of Proposition 4. □

Proof of Proposition 5. We proceed in two steps.

Step 1: We first state necessary conditions for a maximum in Subproblem 1. These conditions are the same under the National and Resident criteria since $\widehat{\theta}$ is given. $\zeta_A := z'_A$ is control variable; z_A, V_A and G are state variables; η, ι and γ are adjoint variables. (SOIC) is transformed into $g(\zeta_A) \geq 0$ to avoid dealing with singular solutions, where g is a $\mathcal{C}^2$-function such that $g' > 0$ and $g(0) = 0$. The Hamiltonian and Lagrangian are

$$\begin{aligned} H^i &= \phi_\rho(V_A) f + \eta\zeta_A + \iota u'_\theta + \gamma(z_A - x_A) f, \\ L^i &= H^R + \pi' R + \kappa g(\zeta_A), \end{aligned}$$

with $i = \{N, R\}$. A solution to Subproblem 1 must satisfy:

$$\partial L^i / \partial \zeta_A = 0 \Leftrightarrow \eta + \kappa g' (\zeta_A) = 0, \tag{A.20}$$

$$\eta' = -\partial L^i / \partial z_A \Leftrightarrow \eta' = \iota u'_x s'_\theta - \gamma (1 - s) f, \tag{A.21}$$

$$\iota' = -\partial L^i / \partial V_A \Leftrightarrow \iota' = -\phi'_\rho f - \iota u''_{\theta x} / u'_x + \gamma f / u'_x - \pi', \tag{A.22}$$

$$\gamma' = -\partial L^i / \partial G \Leftrightarrow \gamma' = 0, \tag{A.23}$$

$$\pi' \geq 0, \ R \geq 0, \ \pi' R = 0, \tag{A.24}$$

$$\kappa \geq 0, \ g (\zeta_A) \geq 0, \ \kappa g (\zeta_A) = 0, \tag{A.25}$$

$$\eta (\underline{\theta}) = \eta \left(\widehat{\theta}\right) = 0, \tag{A.26}$$

$$\iota (\underline{\theta}) \leq 0 \ \ (= 0 \text{ if } R (\underline{\theta}) > 0) , \tag{A.27}$$

$$\iota \left(\widehat{\theta}\right) \geq 0 \ \ \left(= 0 \text{ if } R \left(\widehat{\theta}\right) > 0\right), \tag{A.28}$$

$$\iota \left(\theta_j^-\right) - \iota \left(\theta_j^+\right) = \pi \left(\theta_j^+\right) - \pi \left(\theta_j^-\right) \geq (= 0 \text{ if } R (\theta_j) > 0) . \tag{A.29}$$

η is continuous (see Eq. (75), p. 375, in S-S). We check that $\gamma > 0$. In addition, by continuity of η and (A.26),

$$\eta \left(\widehat{\theta}^-\right) \zeta_A \left(\widehat{\theta}\right) = \eta \left(\widehat{\theta}\right) \zeta_A \left(\widehat{\theta}\right) = 0. \tag{A.30}$$

Step 2: We now turn to Subproblem 2. By Leibnitz's rule,

$$\partial \mathcal{W}^C_{A,\rho} \left(\widehat{\theta}\right) / \partial \widehat{\theta} = \frac{\partial}{\partial \widehat{\theta}} \left[\int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho (V_A) \, dF (\theta) \right] - \phi_\rho \left(V_B \left(\widehat{\theta}\right) - c \left(\widehat{\theta}\right)\right) f \left(\widehat{\theta}\right), \tag{A.31}$$

$$\partial \mathcal{W}^R_{A,\rho} \left(\widehat{\theta}\right) / \partial \widehat{\theta} = \frac{1}{F\left(\widehat{\theta}\right)} \left[\frac{\partial}{\partial \widehat{\theta}} \int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho (V_A) \, dF (\theta) \right] - \frac{f\left(\widehat{\theta}\right)}{F\left(\widehat{\theta}\right)} \mathcal{W}^R_{A,\rho} \left(\widehat{\theta}\right). \tag{A.32}$$

Eq. (79), p. 376, in (S-S) gives the value of the square brackets on the RHSs of (A.31) and (A.32):

$$H \left(\widehat{\theta}^-\right) + \left[\pi \left(\widehat{\theta}\right) - \pi \left(\widehat{\theta}^-\right)\right] R' \left(\widehat{\theta}\right). \tag{A.33}$$

Using the continuity of x_A, z_A, f, V_A, (A.30), (A.29), $T = z_A - x_A$, and the fact that (4) is active at $\widehat{\theta}$, (A.31) and (A.32) reduce to (28) and (29) respectively. □

Proof of Proposition 6. The results of Proposition 5 are also valid for $\widehat{\theta} = \overline{\theta}$ provided $R\left(\widehat{\theta}\right) = 0$. We compute $\partial \mathcal{W}^i_{A,\rho} \left(\widehat{\theta}\right) / \partial \widehat{\theta} \Big|_{\widehat{\theta} = \overline{\theta}} < 0$ and note that $\mathcal{W}^i_{A,\rho} \left(\overline{\theta}\right) = W^N_{A,\rho}$, $i = \{C, R\}$. □

References

Blackorby, C., W. Bossert, and D. Donaldson (2005): *Population Issues In Social Choice Theory, Welfare Economics and Ethics.* Cambridge University Press.

Bodin, J. (1578): *Réponse Au Paradoxe de Monsieur de Malestroit.* Jean du Pays, Paris.

Borjas, G. (1999): "The Economic Analysis of Immigration," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, chap. 28, pp. 1697–1760. North-Holland.

Brett, C., and J. A. Weymark (2008): “The Impact of Changing Skill Levels on Optimal Nonlinear Income Taxes,” *Journal of Public Economics*, 92, 1765–1771.

Brito, D., and W. Oakland (1977): “Some Properties of the Optimal Income-Tax,” *International Economic Review*, 18, 407–423.

Docquier, F., and A. Marfouk (2005): “International Migration by Education Attainment, 1990-2000,” in *International Migration, Remittances and the Brain Drain*, pp. 151–200. World Bank.

——— (2006): “International migration by educational attainment (1990-2000) - Release 1.1,” in *International Migration, Remittances and Development*, ed. by C. Ozden, and M. Schiff, chap. 5. Palgrave Macmillan, New-York.

Gordon, H., and B. McCormick (1981): “Do Council Housing Policies Reduce Migration Between Regions?,” *The Economic Journal*, 61, 919–937.

Guesnerie, R., and J. Seade (1982): “Nonlinear Pricing in a Finite Economy,” *Journal of Public Economics*, 57, 157–179.

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.

Hanson, G. (2005): “International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States,” *Journal of Political Economy*, 113, 239–281.

Hellwig, M. (2007): “A Contribution to the Theory of Optimal Utilitarian Income Taxation,” *Journal of Public Economics*, 91, 1449–1477.

Huber, B. (1999): “Tax Competition and Tax Coordination in an Optimal Income Tax Model,” *Journal of Public Economics*, 71, 441–458.

Inoki, T., and T. Surugan (1981): “Migration, Age and Education,” *Journal of Regional Science*, 21, 507–517.

Jullien, B. (2000): “Participation Constraints in Adverse Selection Models,” *Journal of Economic Theory*, 93, 1–47.

Krause, A. (2008): “Education and Taxation Policies in the Presence of Countervailing Incentives,” *Economica*, Forthcoming.

Leite-Monteiro, M. (1997): “Redistributive Policy with Labour Mobility Accross Countries,” *Journal of Public Economics*, 65, 229–244.

Lewis, T., and D. Sappington (1989): “Countervailing Incentives in Agency Problems,” *Journal of Economic Theory*, 49, 294–313.

Maggi, G., and A. Rodriguez-Clare (1995): “On Countervailing Incentives,” *Journal of Economic Theory*, 66, 238–263.

Mirrlees, J. A. (1971): “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38, 175–208.

——— (1974): “Notes on Welfare Economics, Information and Uncertainty,” in *Essays in economic behaviour under uncertainty*, ed. by M. Balch, D. McFadden, and S. Wu, pp. 243–258. North-Holland, Amsterdam.

——— (1982): "Migration and Optimal Income Taxes," *Journal of Public Economics*, 18, 319–341.

Morelli, M., H. Yang, and L. Ye (2008): "Competitive Nonlinear Taxation and Constitutional Choice," Unpublished Manuscript, Columbia University and Ohio State University.

Nakosteen, R., and M. Zimmer (1980): "Migration and Income: The Question of Self Selection," *Southern Economic Journal*, 46, 840–851.

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

Parfit, D. (1984): *Reasons and Persons*. Oxford University Press.

Piaser, G. (2003): *Labor Mobility and Income Tax Competition*. UCL.

Piketty, T. (1997): "La redistribution fiscale face au chômage," *Revue Française d'Economie*, 12, 157–201.

Sadka, E. (1976): "On Income Distribution, Incentive Effects and Optimal Income Taxation," *Review of Economic Studies*, 43, 261–267.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–229.

Sahota, G. (1968): "An Economic Analysis of Internal Migration in Brazil," *Journal of Political Economy*, 76, 218–245.

Schwartz, A. (1973): "Interpreting the Effect of Distance on Migration," *Journal of Political Economy*, 81, 1153–1169.

Seade, J. (1982): "On the Sign of the Optimum Marginal Income Tax," *Review of Economic Studies*, 49, 637–643.

Seade, J. K. (1977): "On the shape of optimal tax schedule," *Journal of Public Economics*, 7, 203–235.

Seierstad, A., and K. Sydsaeter (1987): *Optimal Control Theory with Economic Applications*. North-Holland, Amsterdam.

Simula, L., and A. Trannoy (2006): "Optimal Linear Income Tax When Agents Vote with their Feet," *FinanzArchiv*.

——— (2009): "Optimal Income Tax under the Threat of Migration by Top-Income Earners," IDEP and Uppsala Center for Fiscal Studies.

Stiglitz, J. (1982): "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17(2), 213–240.

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

Weymark, J. A. (1986): "A Reduced-Form Optimal Nonlinear Income Tax Problem," *Journal of Public Economics*, 30, 199–217.

——— (1987): "Comparative Static Properties of Optimal Nonlinear Income Taxes," *Econometrica*, 55, 1165–1185.

WORKING PAPERS
Uppsala Center for Fiscal Studies
Editor: Håkan Selin

2009:1 Sören Blomquist and Håkan Selin, Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates. 31 pp.

2009:2 Luca Micheletto, Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects. 26 pp.

2009:3 Håkan Selin, The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971. 38 pp.

2009:4 Håkan Selin, Marginal tax rates and tax-favoured pension savings of the self-employed Evidence from Sweden. 32 pp.

2009:5 Tobias Lindhe and Jan Södersten, Dividend taxation, share repurchases and the equity trap. 27 pp.

2009:6 Che-Yuan Liang, Nonparametric Structural Estimation of Labor Supply in the Presence of Censoring. 48 pp.

2009:7 Sören Blomquist , Vidar Christiansen and Luca Micheletto, Public Provision of Private Goods and Nondistortionary Marginal Tax Rates: Some Further Results. 42 pp.

2009:8 Laurent Simula and Alain Trannoy, Optimal Income Tax under the Threat of Migration by Top-Income Earners. 26 pp.

2009:9 Laurent Simula and Alain Trannoy, Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective. 26 pp.