CESifo Working Papers

www.cesifo.org/wp

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

Laurent Simula
Alain Trannoy

CESIFO WORKING PAPER NO. 3326
CATEGORY 1: PUBLIC FINANCE
JANUARY 2011

CESifo *Working Paper No. 3326*

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

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, under the Citizen and Resident criteria, preventing emigration of the highly skilled is not necessarily optimal because the interaction between the incentive-compatibility and participations constraints may cause countervailing incentives. In important cases, a Rawlsian policymaker should decrease top marginal tax rates to keep everyone at home.

JEL-Code: H210, H310, D820, F220.

Keywords: optimal income tax, top-income earners, migration, incentive constraints, participation constraints.

Laurent Simula
Uppsala University
Department of Economics
P.O. Box 513
Sweden - 75120 Uppsala
laurent.simula@nek.uu.se

Alain Trannoy
EHESS
Centre de la Vieille Charité
2 rue de la Charité
France – 13236 Marseille cedex 02
alain.trannoy@univmed.fr

Revised, 7 January 2011
We are grateful the editor and two anonymous rewievers for their very useful comments and suggestions. We also would like to thank Thomas Aronsson, David Bevan, Chuck Blackorby, Philippe Choné, David de la Croix, Jeremy Edwards, Nicolas Gravel, Jonathan Hamilton, Jean Hindriks, Mathias Hungerbühler, Laurence Jacquet, Karolina Kaiser, Etienne Lehmann, Jean-Marie Lozachmeur, Hamish Low, François Maniquet, Michael S. Michael, Gareth Myles, Frank Page, Pierre Pestieau, Panu Poutvaara, Emanuela Sciubba, Gwenola Trotin, Joanna Waschek, John Weymark and David Wildasin. The usual caveats apply.

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." Nowadays in the OECD, many governments, notably in continental Europe, fear the departure of their top-income earners to less redistributive countries. In this context, France and Germany – among other countries – reduced their top income tax rates from 48.1% to 40% and from 48.5% to 45% respectively between 2003 and 2008. However, it is always socially optimal to lower taxes in order to prevent top-income earners from emigrating?

This article examines the *asymmetric* situation in which the redistributive income tax policy of a highly redistributive country is challenged by the low tax policy of one of its neighbours. We adopt the viewpoint of optimal taxation (Mirrlees, 1971) to address this issue. The world consists of a highly redistributive country ("home") and a less redistributive country ("abroad"). The government of the former wants to improve the well-being of the low skilled by taxing highly skilled individuals and redistributing incomes. However, it must recognize that taxpayers will emigrate to the latter if taxes are too high and thus 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)), participation constraints are type-dependent. We borrow these constraints from recent articles in contract theory (see Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995), and Jullien (2000)) and introduce them in the optimal income tax problem. To cast light on the main economic effects and keep the analysis sufficiently manageable, we consider the simple case in which the tax policy abroad is the laissez-faire, and foreigners are immobile. It is possible to motivate this approach as a building block to a tax competition model or to see the home country as part of a "competitive fringe" to make an analogy from industrial organization.

The social objective is complex to specify when individuals are allowed to vote with their feet, because the set of agents whose welfare is to count depends on the income tax itself (Mirrlees, 1982). 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. We therefore address a *population problem* in combination with the optimum income tax problem. As far as we know, the previous literature on optimum nonlinear taxation and individual mobility always restricted itself to fixed-population social criteria.

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 described by 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 maximally 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. This effect 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 the home country of the highly-skilled individuals. When the indirect cost due to countervailing incentives prevails over the benefits of them staying in the home country, implementing a tax schedule inducing them to emigrate increases social welfare. We provide several conditions under which it is not the case. A Rawlsian policymaker should decrease top marginal tax rates to impede emigration of taxpayers in many relevant situations. In particular, this is true for quasilinear and separable preferences under reasonable assumptions on the migration costs. A policymaker which wants to maximise the welfare of the worst-off should design the income tax to keep the best-off at home.

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. On average, 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 regimes are identified but no optimal tax scheme is characterized. In Simula and Trannoy (2006), we address the same issue as in the present paper, where 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 and interpret this result as a lack of degrees of freedom offered by a linear tax when agents can vote with their feet. In Simula and Trannoy (2010), 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 that we implement using French data. In particular, we only consider the National social objective and thus migration never actually occurs. 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). By contrast, Blackorby, Brett, and Cebreiro (2007) address the spatial distribution of the population under optimal income taxes when governments cooperate (or, equivalently, when there is a central tax authority).

Our article is organized as follows. Section 2 sets up the model. Section 3 examines the first-best allocations. Section 4 sets up the second-best income tax problem. Second 5 studies the properties of the second-best optimal allocations under the National Criterion, whilst Section 6 is devoted to the Citizen and Resident criteria. Section 7 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 a laissez-faire country. 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 productivity θ. Individual productivity is public knowledge in the first best and private information in the second best. The cumulative distribution function of θ, denoted F, is common knowledge. It is defined on $\Theta \equiv \left[\underline{\theta}, \overline{\theta}\right] \subseteq \mathbb{R}^+$, 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 C^2 strictly concave function with $U'_x > 0$, $U'_\ell < 0$, $U'_x \to \infty$ as $x \overset{>}{\to} 0$ and $U'_\ell \to -\infty$ as $\ell \to \overline{\ell}$.

Assumption 2. Leisure is a normal good.

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

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

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 the (x, z)-bundle 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 country 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 country 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 country 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 country B is $V_B\left(\theta\right) := U\left(x_B\left(\theta\right), \ell_B\left(\theta\right)\right)$, which is strictly increasing in θ.

2.2. Emigration and Participation Constraints

An individual leaving country A pays a strictly positive *migration cost*, denoted c. Given the cardinality of individual preferences, this cost can be expressed as a "time-equivalent" loss in utility, due to various 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. $c : \Theta \longrightarrow \mathbb{R}^{++}$ *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 (see, e.g., Sahota (1968), Schwartz (1973), Gordon and McCormick (1981), Nakosteen and Zimmer (1980), Inoki and Surugan (1981), Hanson (2005) or Docquier and Marfouk (2005)).

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

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

An individual stays in country A if and only if

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

and therefore leaves country 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. Because the focus is on the mobility of highly skilled individuals, we consider that there is a partition of citizens between country A and country B, with the less skilled individuals being immobile and staying in country A. We therefore introduce a restriction on the class of feasible tax schedules.

Assumption 4. *Country A's resident population is a closed interval of types $[\underline{\theta}, \hat{\theta}]$, with $\hat{\theta} \in \Theta$.*

Without this assumption, the resident population in the home country might in principle consist of several disjoint intervals. This possibility does not seem to be relevant, given the issue we want to examine. Moreover, from a technical viewpoint, we then would not be able to rely on classical optimal control theory to solve the policymaker's problem (Seierstad and Sydsaeter, 1987). We will see later that there are cases in which the fact that country A's resident population is a closed interval of types $[\underline{\theta}, \hat{\theta}]$ is a *property* of the optimum solution.[1]

We consider that country A's government is not able to levy taxes in country B, because the fiscal prerogative is closely linked to national sovereignty, and A is not willing to redistribute income to the individuals staying in country B. Consequently, the tax function in country A is such that $T : \mathcal{T} \to \mathbb{R}$ with $\mathcal{T} = [\underline{\theta}, \hat{\theta}] \times [0, \overline{\ell}]$. Because the tax paid by an individual is equal to the difference between his gross income and his net income, a tax policy is *budget balanced* if and only if it satisfies the tax revenue constraint

$$\int_{\underline{\theta}}^{\hat{\theta}} \left(z_A(\theta) - x_A(\theta)\right) dF(\theta) \geq 0. \tag{5}$$

In the rest of the paper, we denote by γ the Lagrange multiplier associated with the budget constraint (5). Hence, one euro of tax revenue corresponds to γ units of social welfare.

2.3. Social Criteria

Country A's government is a benevolent policymaker 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 $\rho \geqq 0$. A zero aversion corresponds to utilitarianism and an infinite one to the Rawlsian maximin. The social objective is more difficult to specify than in a closed economy. Indeed, it does not only depend on ρ which is captured through an isoelastic function à la Atkinson 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.

(i) Should we maximize total or average social welfare? Classical utilitarianism has been criticized on the grounds that it leads to the so-called repugnant conclusion and this is a significant shortcoming (for details, see Blackorby, Bossert, and Donaldson

[1] This is the case in Propositions 3, 7, 8 and 9.

(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.

(ii) Who are the agents whose welfare is to count? At least three social criteria can be proposed, each of them corresponding to a specific answer. Under the *National criterion*, country A's government cares about the welfare of all its citizens and wants each citizen to choose to stay in country A. The social objective is

$$W_{A,\rho}^{N}(\hat{\theta}) := \int_{\underline{\theta}}^{\hat{\theta}} \phi_{\rho}\left(V_A\left(\theta\right)\right) dF\left(\theta\right), \text{ with } W_{A,\rho}^{N} = -\infty \text{ for } \hat{\theta} < \overline{\theta}. \tag{6}$$

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 country prosperous. This social criterion provides a building block for the solutions of the following Citizen and Resident criteria.

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

$$W_{A,\rho}^{C}(\hat{\theta}) := \int_{\underline{\theta}}^{\hat{\theta}} \phi_{\rho}\left(V_A\left(\theta\right)\right) dF\left(\theta\right) + \int_{\hat{\theta}}^{\overline{\theta}} \phi_{\rho}\left(V_B\left(\theta\right) - c\left(\theta\right)\right) dF\left(\theta\right). \tag{7}$$

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 $\hat{\theta}$ and determines an allocation of country A's citizens between country A and country B. Hence, country A's resident population is endogenous while the set of agents the welfare of whom matters is exogenously fixed.

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

$$W_{A,\rho}^{R}(\hat{\theta}) := \frac{1}{F(\hat{\theta})} \int_{\underline{\theta}}^{\hat{\theta}} \phi_{\rho}\left(V_A\left(\theta\right)\right) dF\left(\theta\right). \tag{8}$$

[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". Twelve senators represent the French citizens living abroad.

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 country B does not count. When this objective is chosen, the tax function as well as the set of agents whose welfare is taken into account, depend on the choice of $\hat{\theta}$.[3] $W^R_{A,\rho}(\hat{\theta})$ 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 by the highly skilled individuals initially living in country A, whose utility is greater than the maximum utility in A.

3. FIRST-BEST ALLOCATIONS

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

We call $V_A^{c\ell}(\theta)$ the indirect utility if country A, and use it as a benchmark. When ρ is finite, the latter 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*. When ρ is infinite, all individuals receive the same utility level.

In this section, we assume that $V_A^{c\ell}(\overline{\theta}) < V_B(\overline{\theta}) - c(\overline{\theta})$. Otherwise, no one would have an incentive to move abroad and the solution in open economy would be the same as in closed economy.

Note that, in the first best, individuals for whom the participation constraints are active pay strictly positive taxes. Indeed, since $V_B(\theta) - c(\theta) < V_B(\theta)$ under Assumption 3, the tangency point between their highest possible indifference curve and their budget constraint must be below the 45°-line through the origin in the gross-income/consumption space.

3.1. National criterion

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

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

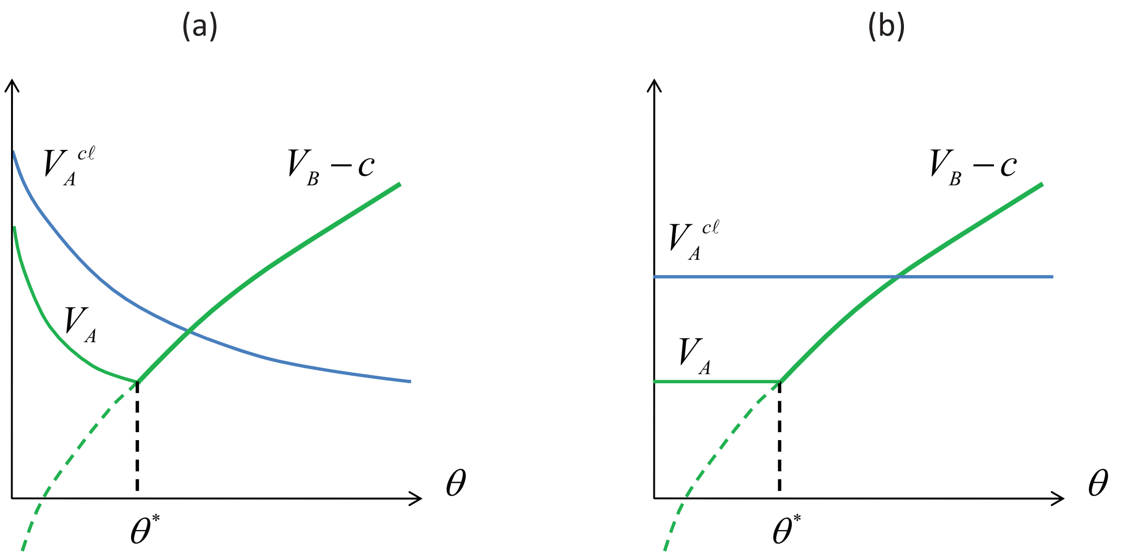

Figure 1: The curse of the middle-skilled

Problem 1 (National Criterion, First-Best)**.** Choose (x, ℓ) in $\mathcal{X}$ to maximize $W^N_{A,\rho}$, with $\hat{\theta} = \overline{\theta}$, subject to (5) and

$$R(\theta) \geq 0 \text{ for } \theta \leq \hat{\theta}. \tag{9}$$

The set $\{\theta \in \Theta\colon\ R(\theta) = 0\}$ is non-empty because $V_A^{c\ell}\left(\overline{\theta}\right) < V_B\left(\overline{\theta}\right) - c\left(\overline{\theta}\right)$. We call θ^* the minimum productivity level for which the participation constraint (9) is binding. A priori, there may exist larger skill levels for which the participation constraint (9) is slack. The following proposition shows that this is not the case in the first-best optimum under the National criterion.

Proposition 1 (The Curse of the Middle-Skilled)**.** *The participation constraints are binding between θ^* and $\overline{\theta}$.*

- *When the government's aversion to income inequality is finite, the optimum indirect utility in country A is V-shaped in θ, minimum at θ^*.*
- *When the government's aversion to income inequality is infinite (maximin), the optimum indirect utility in country A is constant up to θ^* and then increasing.*

Proof. See 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. In both cases, the participation constraint separates the population into two intervals: it is 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. 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. The policymaker solves the following programme.

Problem 2 (Citizen Criterion, First-Best)**.** Choose an allocation (x, ℓ) in $\mathcal{X}$ and $\hat{\theta}$ in Θ to maximize $W^C_{A,\rho}(\hat{\theta})$ subject to the participation constraint (9) and the tax revenue constraint (5).

For later use, we note that by the envelope theorem

$$\frac{\partial W^C_{A,\rho}(\hat{\theta})}{\partial \hat{\theta}} = \underbrace{\gamma T(\hat{\theta}) \frac{f(\hat{\theta})}{F(\hat{\theta})}}_{\text{Tax effect}}. \tag{10}$$

We first investigate whether it is optimal to keep everyone in the home country. For this purpose, we assume that $\hat{\theta} < \overline{\theta}$ is socially optimal. Individuals with productivity $\hat{\theta}$ are indifferent between living in country A or in country B, which means that their location rent $R(\hat{\theta})$ is zero. Individuals with productivity above $\hat{\theta}$ emigrate to country B. Because there is no net subsidy in country B, it is always *feasible* to make the latter relocate to country A, without reducing the indirect utilities of country A's residents, for example in giving them their laissez-faire utility V_B (or a bit more than their reservation utility). Because $c > 0$ and $\phi'_\rho > 0$, one gets $\phi_\rho(V_B - c) < \phi_\rho(V_B)$ for every skill level.

Therefore,

$$\int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho \left(V_A\left(\tau\right)\right) dF\left(\tau\right) + \int_{\hat{\theta}}^{\overline{\theta}} \phi_\rho \left(V_B\left(\tau\right)\right) dF\left(\tau\right) > \int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho \left(V_A\left(\tau\right)\right) dF\left(\tau\right) + \int_{\hat{\theta}}^{\overline{\theta}} \phi_\rho \left(V_B\left(\tau\right) - c\left(\tau\right)\right) dF\left(\tau\right). \quad (11)$$

The right-hand side of (11) is equal to $W^C_{A,\rho}(\hat{\theta})$. Hence, letting the highly-skilled emigrate from country B to country A is both feasible and welfare beneficial, which contradicts the premise. The optimum allocation has thus the following feature.

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

3.3. Resident Criterion

The basic difference between the Citizen and Resident criteria is that the latter does not take the welfare of citizens living abroad into account. Hence, it might be socially desirable to let some individuals emigrate to country B under the Resident criterion.

Problem 3 (Resident Criterion, First-Best)**.** Choose an allocation (x, ℓ) in $\mathcal{X}$ and $\hat{\theta}$ in Θ to maximize $W^R_{A,\rho}(\hat{\theta})$ subject to the participation constraint (9) and the tax revenue constraint (5).

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

$$\frac{\partial W^R_{A,\rho}(\hat{\theta})}{\partial \hat{\theta}} = \underbrace{\gamma T(\hat{\theta}) \frac{f(\hat{\theta})}{F(\hat{\theta})}}_{\text{Tax effect}} + \underbrace{\left[\phi_\rho\left(V_A(\hat{\theta})\right) - W^R_{A,\rho}(\hat{\theta})\right] \frac{f(\hat{\theta})}{F(\hat{\theta})}}_{\text{Utility effect}} \times \mathbf{1}_{\rho<\infty}, \quad (12)$$

where $\mathbf{1}_{\rho<\infty}$ is the indicator function. The first term of (12) corresponds to the fiscal contribution of the individuals with productivity $\hat{\theta}$, expressed in social welfare per capita. It is strictly positive. The second term is the contribution to social welfare of the marginal individuals, in excess to average social welfare. It is divided by the size of the population to obtain a per capita measurement.

When ρ is infinite (maximin), the second term is equal to zero. Therefore, for any $\hat{\theta} > \underline{\theta}$, the marginal individuals' net contribution to social welfare is strictly positive. It is thus socially optimal to keep everyone in the home country.

When the aversion to income inequality ρ is finite, the sign of utility effect is ambiguous. There is a trade-off between the tax and the utility effects of the marginal $\hat{\theta}$-individuals. For example, let us consider that $\hat{\theta} = \overline{\theta}$. The utility effect of the presence in country A of the $\hat{\theta}$ -individuals is given by $\{\phi_\rho \left[V_B(\overline{\theta}) - c(\overline{\theta})\right] - W^R_{A,\rho}\left(\overline{\theta}\right)\} \times (f(\overline{\theta})/F(\overline{\theta}))$. From the left-hand side of Figure 1, we can see that $\phi_\rho \left[V_B\left(\overline{\theta}\right) - c\left(\overline{\theta}\right)\right]$ may be small compared to the average social welfare $W^R_{A,\rho}\left(\overline{\theta}\right)$ when the reservation utility is quite flat. If the utility effect is sufficiently negative to counterbalance the positive tax effect, letting the most productive citizens leave the home country increases social welfare.[4]

4. SECOND-BEST ALLOCATIONS: PRELIMINARY RESULTS

We now turn to the characterization of second-best optimum allocations. The distribution of characteristics in the economy remains common knowledge, but individual productivity is now private information. Country 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

We first state the optimal income tax problem. A tax schedule T is *incentive compatible* if and only if individuals living in country A have an incentive to reveal their type truthfully when it is implemented. By the revelation principle, the incentive-compatibility conditions are the following:

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

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, the incentive-compatibility conditions (IC) are equivalent to:

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

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

[4]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 Θ is first considered. Let $\overline{W}^R_{A,\rho}(\widehat{\theta})$ be the social value function. The optimal value of $\widehat{\theta}$ is that for which $\overline{W}^R_{A,\rho}(\widehat{\theta})$ is maximum.

The proof of this equivalence is standard (see, e.g., Myles (1995)) and thus omitted. The first-order condition for incentive compatibility (FOIC) is an envelope condition specifying how the indirect utility V_A must locally change. Because $V_A' \geq 0$, V_A cannot be V-shaped as in the first-best. The second-order condition for incentive compatibility (SOIC) is a global monotonicity condition of gross income. The analysis will focus on continuous mechanisms which possibly exhibit kinks at a finite number of points corresponding to jumps in the marginal tax rates. In this case, the location rent $R(\theta)$ is continuous and the second-order condition (SOIC) is equivalent to:

$$z_A'(\theta) \geq 0 \text{ for } \theta \leq \hat{\theta}. \tag{SOIC'}$$

Because country A's government does not know who are the agents for whom the location rent $R(\theta)$ is zero, we have to take into account both the participation constraint and the condition for incentive compatibility for every resident staying in the home country [5]. The second-best optimal non-linear income tax problem can thus be written as follows.

Problem 4 (Second-Best)**.** Choose a tax schedule $T(z_A)$ to maximize social welfare $W^i_{A,\rho}$, $i = \{N, C, R\}$, subject to the following constraints : (i) (FOIC), (SOIC'), (5), (9); (ii) $\hat{\theta} = \overline{\theta}$ when $i = N$ and $\hat{\theta}$ in Θ otherwise.

In the closed-economy version of Problem 4, $\hat{\theta}$ is equal to $\overline{\theta}$ and the participation constraint (9) is not taken into account. Let $V_A^{c\ell}(\theta)$ be the (second-best) optimum indirect utility. If $V_A^{c\ell}(\theta) \geq V_B(\theta) - c(\theta)$ for every θ in Θ, allowing individuals to vote with their feet does not alter the social optimum. We place ourselves in the case where there are individuals for whom $V_A^{c\ell}(\theta) < V_B(\theta) - c(\theta)$ because the participation constraint would otherwise never be active.

Problem 4 raises three important difficulties, to which we are not confronted in a closed economy. First, the participation constraint (9) can a priori bind on any subset of the resident population, even at isolated points, because the location rent $R(\theta)$ is not necessarily monotonic. Second, this constraint is a pure state constraint. The adjoint variable may thus have jump discontinuities. Third, under the Citizen and Resident criteria, the upper bound $\hat{\theta}$ is free to vary between $\underline{\theta}$ and $\overline{\theta}$.

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

[5] If the participation constraints (9) 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.

variable associated with (FOIC) and $\pi' \geq 0$ the Lagrange multiplier of (9), 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

$$\pi(\theta) := \pi\left(\overline{\theta}\right) - \int_{\theta}^{\overline{\theta}} \pi'(\tau)\, d\tau. \tag{13}$$

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

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

4.2. Optimal Tax Schedule for 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.

For this purpose, let I be an interval of positive length where the participation constraint (4) is active. By definition, for θ in I, we have $R(\theta) \equiv 0$ and thus $V_A'(\theta) = V_B'(\theta) - c'(\theta)$. Hence, the rate of increase in 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 interval I. In addition, employing (FOIC) and rearranging yield:

$$z_A(\theta) = -\theta \frac{V_B'(\theta) - c'(\theta)}{u_z'(x_A(\theta), z_A(\theta); \theta)} \text{ for } \theta \text{ in } I, \tag{14}$$

and by differentiation:

$$z_A'(\theta) = \frac{\left[V_B'(\theta) - c'(\theta)\right]\left\{\theta\left(u_{xz}'' x_A' + u_{\theta z}''\right) - \left(1 + \theta \frac{V_B''(\theta) - c''(\theta)}{V_B'(\theta) - c'(\theta)}\right) u_z'\right\}}{\left(u_z'\right)^2 - \theta\left(V_B'(\theta) - c'(\theta)\right) u_{zz}''} \text{ for } \theta \text{ in } I. \tag{15}$$

The second-order condition for incentive compatibility (SOIC') can only be satisfied on interval I if the curly bracket in (15) is non-negative. When preferences are separable

($u''_{xz} = 0$), one gets:

$$z'_A(\theta) \geq 0 \Leftrightarrow \frac{\theta u''_{\theta z}}{u'_z} \leq \left(1 + \theta \frac{V''_B(\theta) - c''(\theta)}{V'_B(\theta) - c'(\theta)}\right), \tag{16}$$

the LHS of which is negative because $u''_{\theta z} > 0$ and $u'_z < 0$.

Property 1. *Let preferences be separable ($u''_{xz} = 0$) and consider an interval I of positive length where the participation constraint (9) is active. There is no bunching on this interval when*

$$\theta \frac{V''_B(\theta) - c''(\theta)}{V'_B(\theta) - c'(\theta)} > -1 \text{ for } \theta \text{ in } I. \tag{17}$$

Condition (17) states that the elasticity of the marginal reservation utility to the wage rate, evaluated at θ, is greater than -1 for every θ in I. To have further insight, we now turn to quasilinear-in-consumption preferences

$$u(x, z; \theta) = x - v(z/\theta). \tag{18}$$

We examine the curvature of the tax schedule along the participation constraint. We consider that individual preferences are quasilinear in consumption as in (18), but we do not specify the desutility of labour $v(.)$. We call $D = 1 - T'$ and we restrict ourselves to cases in which $0 < D < 1$. The first-order condition of the individual utility maximisation programme yields $\ell_A = v'^{-1}(\theta D)$. The first-order condition for incentive compatibility is $V'_A(\theta) = Dv'^{-1}(\theta D)$. Along I, we have $R'(\theta) = 0$, i.e., $Dv'^{-1}(\theta D) = V'_B(\theta) - c'(\theta)$, which implicitly defines D as a function of the marginal outside option $V'_B(\theta) - c'(\theta)$. By the implicit function theorem, we obtain:

$$T'' = -\frac{dD}{d\theta} = \frac{D^2/v''(v'^{-1}(\theta D)) - (V''_B(\theta) - c''(\theta))}{v'^{-1}(\theta D) + D/v''(v'^{-1}(\theta D))}, \tag{19}$$

where $V''_B(\theta) = 1/[v''(v'^{-1}(\theta)]$. The sign of T'' is thus as follows:

$$T'' < 0 \Leftrightarrow \frac{D^2 v''(v'^{-1}(\theta)) - v''(v'^{-1}(\theta D)}{v''(v'^{-1}(\theta) v''(v'^{-1}(\theta D))} + c''(\theta) < 0. \tag{20}$$

Property 2. *Let individual preferences be quasilinear in consumption, the marginal disutility of labour be concave, the marginal cost of migration be non-increasing in productivity and consider $0 < T' < 1$. Then the optimal marginal tax rate decreases along the participation constraints.*

A sufficient condition for the optimal marginal tax rate to be negative along the participation constraint is therefore that $v''' < 0$ and $c'' \leqq 0$: the desutility of working an extra hour increases at a decreasing rate while the cost of migration is concave. The first condition corresponds to a psychological law, which can be tested empirically. The second condition depends on the nature of the migration costs faced by the individuals.

5. SECOND-BEST ALLOCATIONS UNDER NATIONAL CRITERION

We now study the impact of the threat of migration on the optimum tax scheme in country A when its government adopts the National criterion. This analysis provides a building block for the analysis of the optimal tax schedule under the Citizen and Resident criteria, that will be carried out in the next subsection. We start by examining conditions under which the participation constraints generate a partition of the population in two intervals. We then characterize the optimal income tax schedule, providing a formula for the optimal marginal tax rates.

5.1. When the participation constraints generate a partition of the population: sufficient conditions

We assumed in the last section that there is a non-degenerated interval I on which the participation constraint is active. We now show that there are cases in which this is actually the case. Moreover, we establish that under certain conditions the National solution is such that the participation constraint generates a partition of the population in two intervals.

To this end, we consider that individuals have quasilinear-in-consumption preferences and that the policymaker's aversion to income inequality in infinite (maximin). We choose the exogenous parameters of the model so that (i) the autarkic second-best indirect utility crosses the reservation utility only once, from above, and (ii) the reservation utility is convex. Assumption (i) is in line with the focus of the article in which the threat of migration comes from individuals in the upper part of the skill distribution. Assumption (ii) is equivalent to $V_B'' > c''$. It ensures that gross income is nondecreasing along an interval where the participation constraint is active. Note that the indirect utility abroad V_B is convex when individual preferences are quasilinear.

Under these assumptions, we establish that – under the National criterion – the participation constraint is binding on the interval $[\theta^*, \overline{\theta}]$, where θ^* is the minimum productivity level at which individuals threaten to emigrate. We know that the tax function is continuous, and differentiable except on a set of measure zero. Consequently, the indirect utility V_A and the location rent R are both piecewise continuously differentiable.

We consider an allocation $\theta \longrightarrow (\hat{x}_A(\theta), \hat{z}_A(\theta))$ which satisfies the incentive, participation and tax revenue constraints. If the participation constraint is slack at one point above θ^* (i.e., $R(\theta) > 0$ for $\theta > \theta^*$), then this constraint is slack on an interval. We call $[\theta_-, \theta^+]$ this interval.

In the Appendix, we prove by contradiction that this allocation cannot be socially optimal. The idea is that it is possible to construct another feasible allocation for which tax revenue is higher. This allocation $\theta \longrightarrow (\check{x}_A(\theta), \check{z}_A(\theta))$ is such that $(\check{x}_A(\theta), \check{z}_A(\theta)) = (\hat{x}_A(\theta), \check{z}_A(\theta))$. Because the maximin objective is equivalent to maximising tax revenue, this allocation is welfare-improving and thus the initial allocation cannot be optimum. The following proposition is obtained.

Proposition 3. *Let individuals have quasilinear-in-consumption preferences and consider the National criterion with an infinite aversion to income inequality (maximin). If the autarkic second-best indirect utility crosses the reservation utility only once, from above and the reservation utility is convex, then the optimum allocation is such that:*

$$V_A(\theta') = V_B(\theta') - c(\theta') \Longrightarrow V_A(\theta) = V_B(\theta) - c(\theta) \text{ for every } \theta > \theta'. \tag{21}$$

Proof. See the Appendix. □

We will use this proposition later to investigate cases in which the participation constraint is binding on an interval at the optimum under the Citizen and Resident criteria.

5.2. Optimal income tax rates

We now investigate the impact of the potential threat of migration on the optimal income tax schedule.

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

$$\frac{T'(z_A(\theta))}{1 - T'(z_A(\theta))} = A(\theta)\, B(\theta)\, C(\theta) \qquad \text{for } \theta < \overline{\theta}, \tag{22}$$

where

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

with

$$B_1(\theta) := \frac{1}{1-F(\theta)} \int_{\theta}^{\overline{\theta}} \left[1 - \frac{\phi'_\rho (V_A(\tau)) u'_x (x_A, z_A; \tau)}{\gamma}\right] \Psi_{\theta\tau} dF(\tau),$$

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

$$\pi'(\tau) \geq 0 \ \ (= 0 \text{ if } R(\tau) > 0) \ \text{ and } \iota(\overline{\theta}) \geq 0 \ \ \left(= 0 \text{ if } R(\overline{\theta}) > 0\right).$$

At the top,

$$\frac{T'(z_A(\overline{\theta}))}{1 - T'(z_A(\overline{\theta}))} = \frac{A(\overline{\theta})}{\overline{\theta} f(\overline{\theta})} \frac{\iota(\overline{\theta}) u'_x (x_A, z_A; \overline{\theta})}{\gamma} \leq 0 \ \ \left(= 0 \text{ if } R(\overline{\theta}) > 0\right). \tag{23}$$

Proof. See 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 maximum national productive capacity by 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, because it depends on gross income which is endogenous. The factor $B(\theta)$, which combines efficiency and equity, is the only factor not written as in Mirrlees's formula, in which the right-hand side of (22) 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 (22). 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 now receive a utility below their reservation level. Then the participation constraint (9) is no longer satisfied. These individuals have to be compensated for the increase in taxes they face.

We first examine compensation for the individuals with gross income between z and $z_{\overline{\theta}}$. The *compensation effect* leads A's government to *totally* compensate them for staying in A. Each of them is given extra $dTdz$ euros, which generate $u'_x(x_A, z_A; \tau) \times dTdz$ 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 (5), the cost in terms of social welfare of the compensation of the τ -individuals amounts to

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

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 country A's government to increase their taxes. As a result, it is not necessary to fully compensate the potentially mobile individuals forthe 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 (24), the social cost of the compensation of the τ-individuals, including income effects, is

$$\pi'(\tau)\,\Psi_{\theta\tau} \times \frac{u'_x(x_A, z_A; \tau)}{\gamma} \times dTdz. \tag{25}$$

By integration of (25), we get the cost of compensating the individuals with productivity between θ_z and $\overline{\theta}$. For the individuals on the upper bound of the population, the social cost is directly obtained as

$$\left.\frac{\partial W^N_{A,\rho}}{\partial V_A}\right|_{\overline{\theta}} \times \frac{u'_x(x_A, z_A; \theta_z)}{\gamma} \times dTdz = \iota\left(\overline{\theta}\right) \times \frac{u'_x(x_A, z_A; \theta_z)}{\gamma} \times dTdz. \tag{26}$$

Finally, by (25) and (26), the average social cost of the compensation of all potentially mobile individuals with gross income above z is

$$\frac{1}{1-F(\theta_z)}\left[\int_{\theta_z}^{\overline{\theta}} \frac{\pi'(\tau)\, u'_x(x_A, z_A; \tau)}{\gamma}\Psi_{\theta_z \tau} d\tau + \frac{\iota\left(\overline{\theta}\right) u'_x(x_A, z_A; \theta_z)}{\gamma}\right] \times dTdz$$
$$= B_2(\theta_z) \times dTdz. \quad (27)$$

$B_2(\theta_z)$ is positive as soon as there are individuals with productivity above θ_z for whom the participation constraints are binding. This term counters 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 costly 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. Consequently, country 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.

6. SECOND-BEST ALLOCATIONS UNDER CITIZEN AND RESIDENT CRITERIA

Under the National criterion, the whole population is constrained to stay in country A. We now relax this constraint, in order to examine whether keeping everybody in the home country is not too costly in terms of social welfare.

6.1. A Two-Step Problem

For this purpose, we separate Problem 4 into two subproblems to determine the optimal upper bound $\hat{\theta}$. In the first subproblem, $\hat{\theta}$ is arbitrarily chosen by country A's government.

Subproblem 1. *For a given $\hat{\theta}$ in Θ and $i = \{C, R\}$, choose an allocation (x_A, z_A) to maximize social welfare $W^i_{A,\rho}(\hat{\theta})$ subject to the conditions for incentive compatibility (FOIC) and (SOIC'), the tax revenue constraint (5) and the participation constraint (9).*

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

Subproblem 2. *For $i = \{C, R\}$, choose $\hat{\theta}^i$ in Θ which maximises $\mathcal{W}^i_{A,\rho}(\hat{\theta})$.*

Subproblem 1 is a generalization of the second-best National problem where the upper productivity in country A is given exogenously . This implies that the optimal tax schedules obtained under the National, Citizen and Resident criteria all share qualitative properties.

Proposition 5. *Under the Citizen and Resident criteria, the optimal marginal tax rates are given by Proposition 4 with:*

- *$\overline{\theta}$ is replaced by $\hat{\theta}^i$ and $1 - F(\theta)$ by $F\left(\hat{\theta}^i\right) - F(\theta)$, $i = \{C, R\}$.*
- *$\phi'_\rho(V_A)$ is divided by $F(\hat{\theta}^R)$ under the Resident criterion.*

Proof. See the Appendix. □

6.2. Countervailing Incentives and Upward Mimicking Behaviour

We are now prepared to examine the allocation of individuals between country A and country B resulting from the implementation of the Citizen and Resident optimal income tax schedules. For every $\hat{\theta}$ in Θ, the $\hat{\theta}$-individuals are indifferent between living in country A or in country B. Let us assume $\hat{\theta} < \overline{\theta}$. Hence, individuals with productivity above $\hat{\theta}$ are in country B. Making them relocate to country A requires adjustments to prevent them from imitating less productive individuals. It also brings about a *new upward mimicking behaviour*: country A's residents may now have an incentive to mimicking the $\hat{\theta}$-individuals because they have the most appealing outside options.

The upward mimicking behaviour is crucial to understanding the interactions between the conditions for incentive compatibility and the type-dependent participation constraint. In closed economy, individuals have the usual incentive to *understate* their productivity to obtain greater social benefit whilst enjoying more leisure.[6] When type-dependent participation constraints are taken into account, the individuals may also be tempted to *overstate* their productivity, in working harder, to obtain greater compensation for staying in the home country. This behaviour reflects *countervailing incentives.*

[6] 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) and Simula (2010)). Hellwig (2007) has established, in both discrete and continuous models that under "desirability of redistribution" only the downward incentive-compatibility constraints are binding in a closed economy.

An asymmetry in terms of informational constraints between the individuals with productivity below $\hat{\theta}$ and the marginal $\hat{\theta}$-individuals may therefore arise. Indeed, contrary to the former, the latter can only have the usual downward incentives. The cost of making the θ-individuals reveal their private information may therefore drop at $\hat{\theta}$. This cost is represented by $\iota(\theta)$, which may thus have a downward jump discontinuity $\iota(\hat{\theta}^-) - \iota(\hat{\theta}) \geqq 0$ at this productivity level (we denote by minus the limit to the left).[7]

To make the individuals to the (very) left of $\hat{\theta}$ reveal their type, their utility in A must be increased. This rise *mechanically* reduces the shadow cost of participation for these types, which is captured by $\pi^i_{\hat{\theta}}(\theta)$. This effect stops suddenly at $\hat{\theta}$. Consequently, the shadow cost $\pi(\theta)$ may have an *upward* jump discontinuity $\pi(\hat{\theta}) - \pi(\hat{\theta}^-) \geq 0$, which corresponds to the *downward* jump discontinuity in $\iota(\theta)$ at $\hat{\theta}$. These discontinuities have the same magnitude and vanish when the participation constraint is inactive at $\hat{\theta}$.

Lemma 1. *$\iota(\widehat{\theta^-}) - \iota(\hat{\theta}) = \pi(\hat{\theta}) - \pi(\hat{\theta}^-) \geq 0$ ($= 0$ if the participation constraint is inactive at $\hat{\theta}$).*

Proof. See (A.34) in the Appendix. □

They illustrate the non-trivial interactions that may arise between the incentive-compatibility and participation constraints.

6.3. Effects of the Presence of the Marginal Individuals on Social Welfare

A variational analysis provides insights into the costs and benefits of the presence in the home country of the marginal $\hat{\theta}$-individuals. It is useful to introduce the following expressions.

- Tax Effect: $TE(\hat{\theta}) = \gamma T\left(z_A(\hat{\theta})\right)$.
- Utility Effect: $UE(\hat{\theta}) = \left[\phi_\rho(V_A(\hat{\theta})) - \mathcal{W}^R_{A,\rho}(\hat{\theta})\right] \mathbf{1}_{\rho<\infty}$.
- Net Information Externality: $IE(\hat{\theta}) = [\iota(\hat{\theta}) - \iota(\hat{\theta}^-)]R'(\hat{\theta}) - \iota(\hat{\theta}^-)V'_A(\hat{\theta})$.

Proposition 6. *Using a variational analysis,*

$$\begin{aligned} &\textit{Citizen criterion:} \quad && \frac{\partial \mathcal{W}^C_{A,\rho}(\hat{\theta})}{\partial \hat{\theta}} = f(\hat{\theta}) \times TE(\hat{\theta}) - IE(\hat{\theta}), \\ &\textit{Resident criterion:} \quad && \frac{\partial \mathcal{W}^R_{A,\rho}(\hat{\theta})}{\partial \hat{\theta}} = \frac{f(\hat{\theta})}{F(\hat{\theta})}\left[TE(\hat{\theta}) + UE(\hat{\theta})\right] - \frac{IE(\hat{\theta})}{F(\hat{\theta})}. \end{aligned} \tag{28}$$

[7] Note that ι – as well as π in the next paragraph – depend on the upper bound $\hat{\theta}$ and on the chosen welfare criterion. We do not show it explicitly to keep notations simple.

Proof. See the Appendix. □

We first see that we recover the tax and utility effects identified in the first-best (10) and (12): the channels identified in the first-best still play a key part in the second-best. However, note that the utility effect $UE(\hat{\theta})$ is now necessarily positive because the indirect utility in country A is nondecreasing in the second-best.

A new term is specific to the second-best. It corresponds to an information externality and its expression is – at first sight – rather complicated. We will see that it reflects the marginal costs and benefits with regard to incentives of the presence in country A of the marginal $\hat{\theta}$-individuals. These costs and benefits arise from the upward and downward mimicking behaviours.

Let us assume that the $\hat{\theta}$-individuals were living in country B and now relocate to country A.

- If the participation constraint is active for the $\hat{\theta}$-individuals but not for individuals to the very left of $\hat{\theta}$, then the latter are left with a strictly positive location rent. Note that this implies a negative *marginal* rent ($R'(\hat{\theta}) < 0$). In that case, individuals in the upper tail of the productivity distribution have an incentive to mimic the most productive agents living in country A to benefit from their higher outside options. The location rent of the individuals to the very left of $\hat{\theta}$ must be increased at the margin, i.e. by an amount $-R'(\hat{\theta})$, to induce truthtelling. The shadow price of this increase is $\iota(\hat{\theta}^-) - \iota(\hat{\theta}) \geqq 0$. Hence, the social cost of the *upward* mimicking behaviour is $[\iota(\hat{\theta}) - \iota(\hat{\theta}^-)] \times R'(\hat{\theta}) \geqq 0$, the first part of the information externality $IE(\hat{\theta})$. But this is not the end of the story.

- The increase in the location rent also brings about a positive effect on social welfare: because individuals to the very left of $\hat{\theta}$ have now greater utility, they are less inclined to mimic less productive individuals. The slope of the indirect utility V_A' required for them to reveal their type truthfully is therefore reduced at the margin. Because $\iota(\hat{\theta}^-)$ is the shadow price of the first-order condition for incentive compatibility (FOIC), the social benefit of this slackening of the *downward* incentive compatibility constraints is $\iota(\hat{\theta}^-)V_A'$. This is the second part of the information externality $IE(\hat{\theta})$.

Combining the positive and negative effects, we obtain the *net* marginal social cost incurred to restore an incentive tax scheme at the top. If the information externality is sufficiently large, it may prevail over the tax (and utility) effects. Then, it would be optimum to let the marginal $\hat{\theta}$-individuals go abroad.

6.4. Preventing emigration of top earners may be optimum

We now examine cases in which we can establish that it is optimal to keep everyone at home. We start with the same setting as in Proposition 3: individuals have quasilinear-in-consumption preferences, the government's aversion to income inequality is infinite (maximin), the autarkic second-best indirect utility crosses the reservation utility only once, from above and the reservation utility is convex. Moreover, the disutility of labour is isoelastic. Note that the Citizen and Resident criteria coincide under the maximin.

We initially consider the population of measure μ_1 described by the CDF $F_1(\theta)$ over $[\underline{\theta}, \hat{\theta}]$. The optimal allocation is a mapping $\theta \in [\underline{\theta}, \hat{\theta}] \longrightarrow (x_1(\theta), z_1(\theta))$. We call T_1 the corresponding tax schedule. We know from Proposition 3 that the optimum allocation involves a partition of the population into two connected intervals, the participation constraint being binding for the top of the population. We now consider a second population obtained from the first one, through the addition of an interval $(\hat{\theta}, \overline{\theta}]$ of measure μ_2 over which the distribution of θ is given by the CDF $F_2(\theta)$. The measure of the second population is thus $\mu_1 + \mu_2$. The CDF of the second population is:

$$\theta \in \left[\underline{\theta}, \overline{\theta}\right] \longrightarrow F_{1+2}(\theta) = \frac{\mu_1}{\mu_1 + \mu_2} F_1(\theta) + \frac{\mu_2}{\mu_1 + \mu_2} F_2(\theta). \tag{29}$$

We construct the following allocation:

- For $\theta \in (\underline{\theta}, \hat{\theta}]$, consumption and gross income are unaltered. Individuals pay taxes given by T_1.
- For $\theta \in (\hat{\theta}, \overline{\theta}]$, the participation constraint is binding and the incentive compatibility constraint is satisfied. This is equivalent to the equation of motion $u'_\theta(x_A(\theta), z_A(\theta); \theta) = V'_B(\theta) - c'(\theta)$ and the boundary condition $u(x_A(\hat{\theta}), z_A(\hat{\theta}); \hat{\theta}) = V_B(\hat{\theta}) - c(\hat{\theta})$. We have established – in the proof of Proposition 3 – that this system of two equations has a solution, which defines the optimum tax over $(\hat{\theta}, \overline{\theta}]$, with $T_2 : (\hat{\theta}, \overline{\theta}] \to T_2(\theta) = z_2(\theta) - x_2(\theta)$. T_2 is non-decreasing over $(\hat{\theta}, \overline{\theta}]$ because the participation constraint is binding and $c' \geqq 0$ (see Corollary 1 in Simula and Trannoy (2010)).

For $\theta \in (\underline{\theta}, \overline{\theta})$, we have $0 \leqq T' < 1$ on the participation constraint because $c' \geqq 0$ (see Corollary 1 in Simula and Trannoy (2010)). Outside the participation constraint, we have $0 < T' < 1$ (see Simula and Trannoy 2010, formula (32) with $\gamma = 1$). Then, $\theta \in (\underline{\theta}, \hat{\theta}) \longrightarrow T_1(\theta)$ is an increasing function of θ. Because $V_B - c$ is a C^2-function, we have $\lim_{\theta \to \hat{\theta}} T_2(\theta) = T_1\left(\hat{\theta}\right)$. Therefore,

$$T_{1+2}(\theta) := \begin{cases} T_1(\theta), & \theta \in [\underline{\theta}, \hat{\theta}] \\ T_2(\theta), & \theta \in (\hat{\theta}, \overline{\theta}] \end{cases} \tag{30}$$

is increasing. In addition, $T_1(\hat{\theta}) > 0$. Consequently, $T_{1+2}(\theta) > 0$ for every $\theta \in (\hat{\theta}, \overline{\theta}]$. We have built a new allocation which increases tax revenue and thus social welfare. Tax receipts are an increasing function of $\hat{\theta}$. Consequently, everyone must stay in A in the social optimum. This result can be summarized as follows.

Proposition 7. *Let individuals have quasilinear-in-consumption preferences, with constant elasticity of labour supply, the government's objective be the maximin, the autarkic second-best indirect utility cross the reservation utility only once, from above, and let the reservation utility be convex with nondecreasing migration cost. Under the Citizen and Resident criteria, we have* $\hat{\theta} = \overline{\theta}$ *in the optimum.*

We continue with loglinear preferences, given by $U(x, \ell) = \log x + \log(1 - \ell)$, and constant migration costs. It is possible to replace a nonlinear tax schedule $\theta \longrightarrow T(z_A(\theta))$ by a collection of linearized schedules $\theta \longrightarrow (t(\theta), m(\theta))$ such that $x_A(\theta) = (1 - t(\theta))\theta\ell_A(\theta) + m(\theta)$. This is illustrated in Figure 2. The first-order condition of the individual utility maximisation programme yields:

$$\ell_A(\theta) = \frac{1}{2} - \frac{m(\theta)}{2\theta(1 - t(\theta))}. \tag{31}$$

We assume that the costs of migration are constant. To satisfy the participation constraint on a nondegenerated interval I, we must have $R'(\theta) = 0$, which is equivalent to $\ell_A(\theta) = 1/2$ and implies $m(\theta) = 0$. Moreover, we must have $R(\theta) = 0$ on this interval, i.e., $t(\theta) = 1 - \exp(-c)$. Note that $t(\theta)$ is constant and belongs to $(0, 1)$. The implications are threefold. First, it is always possible to modify a tax schedule in such a way that the participation constraint is binding on a given interval. Second, if individuals were on this interval, they would pay positive taxes, given by $t(\theta)\theta\ell_A(\theta) = \theta[1 - \exp(-c)]/2$. Third and consequently, country A's policymaker should prevent emigration of the highly skilled under the maximin ($\hat{\theta}^R = \hat{\theta}^C = \overline{\theta}$ in the optimum).

Proposition 8. *Let individuals have loglinear preferences* $U(x, \ell) = \log x + \log(1 - \ell)$, *the policymaker adopt the maximin, and let the costs of migration be constant. In this case, it is socially optimal to keep everyone in the home country under the Citizen and Resident criteria.*

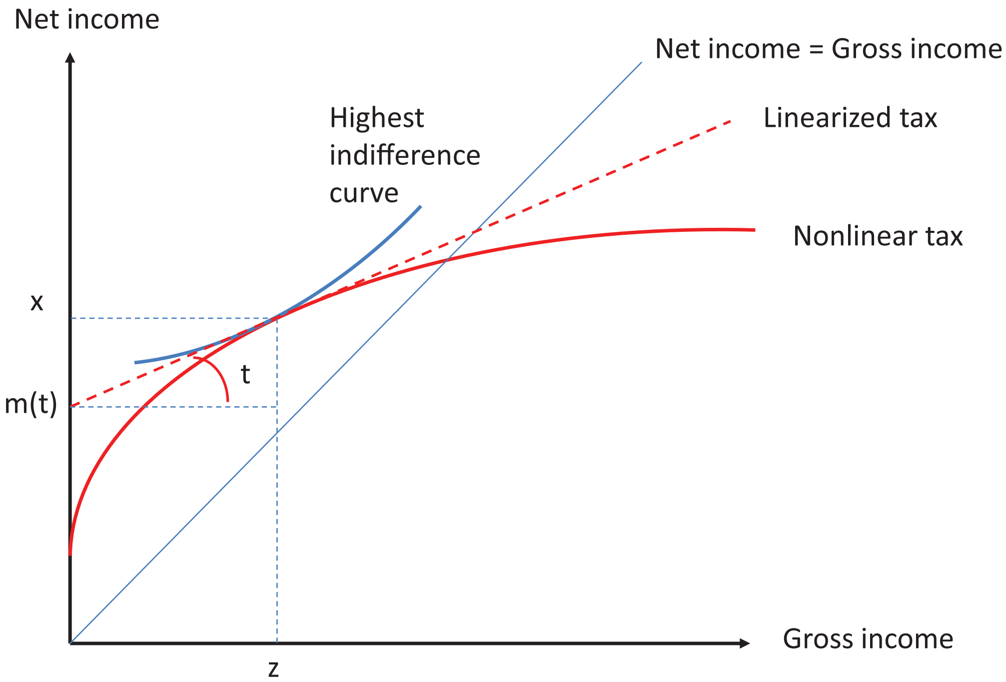

Figure 2: Linearized schedule

We next consider *separable* preferences, given by $U(x,\ell) = h(x) - v(\ell)$ and constant migration costs. Considering separable preferences is interesting because of the Atkinson-Stiglitz theorem. We assume that the labour supply is such that the substitution effect prevails over the income effect. We replace a nonlinear tax schedule $\theta \longrightarrow T(z_A(\theta))$ by a collection of linearized schedules $\theta \longrightarrow (t(\theta), m(\theta))$ such that $x_A(\theta) = (1-t(\theta))\theta\ell_A(\theta) + m(\theta)$. To satisfy the participation constraint on a nondegenerated interval I, we must have $R'(\theta) = 0$, which is equivalent to $\ell_A(\theta) = \ell_B(\theta)$. In the t, m space $([0,1] \times \mathbb{R})$ and given θ, we represent $m_{R'}(t)$ defined by $\ell_A(t,m) = \ell_B(\theta)$. We know that $\ell_A(0,0) = \ell_B(\theta)$. Because leisure is a normal good and the substitution effect prevails over the income effect $(\partial\ell_A/\partial t < 0)$, we have $dm_{R'}(t)/dt < 0$. Hence $m_{R'}(t)$ is decreasing on $[0,1]$, equal to 0 for $t = 0$ and strictly negative for $t = 1$. Moreover, we must have $R(\theta) = 0$ on I. In the t, m space $([0,1] \times \mathbb{R})$ and given θ, we also represent $m_R(t)$ defined by $R(\theta) = 0$. Because $\ell_B(\theta)$ is given, we obtain $dm_R(t)/dt = \theta\ell_B(\theta) > 0$. We now examine the intersection points between $m_R(t)$ and the axes of the t, m space. Note that (i) when $t = 0$, we have $R(\theta) = 0 \Leftrightarrow h(\theta\ell_B + m) = h(\theta\ell_B) - c$ and (ii) when $m = 0$, we have $R(\theta) = 0 \Leftrightarrow h(\theta\ell_B(1-t)) = h(\theta\ell_B) - c$, which is equivalent to $t = 1 - h^{-1}(h(\theta\ell_B) - c)/(\theta\ell_B) < 1$. Hence, $m_R(t)$ is increasing on $[0,1]$, negative for

$t = 0$ and strictly positive for $t = 1$. By continuity and monotonicity, there is a unique junction point $(\tilde{t}, \tilde{m})$ between $m_R(t)$ and $m_{R'}(t)$ for every θ. It is such that $\tilde{m} < 0$ and $0 < \tilde{t} < 1$. Consequently, it is always possible to modify the tax schedule in such a way that the participation constraint is binding on a given interval. Individuals for whom the participation constraints are binding pay strictly positive taxes, $\tilde{t}\theta\ell_B(\theta) - \tilde{m}$. It is therefore socially optimum for country A's policymaker to prevent emigration of the highly skilled under the maximin ($\hat{\theta}^R = \hat{\theta}^C = \overline{\theta}$ in the optimum).

Proposition 9. *Let individuals have separable preferences $U(x, \ell) = h(x) - v(\ell)$, let the policymaker adopt the maximin, and let the costs of migration be constant. When the substitution effect on labour supply prevails over the income effect, it is socially optimal to design the tax schedule so that everyone decides to stay in the home country, under the Citizen and Resident criteria.*

In other words, we have exhibited three cases in which the optimum tax schedule is the same under the National, Citizen and Resident criteria. The first case is not implied by the third one: we consider nondecreasing migration costs in the first and constant costs in the third. The second case is not included in the third one: with loglinear preferences, the labour supply is backwards bending and thus the substitution effect does not prevail over the income effect as in the third case.

7. 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 incentive constraints in a non-trivial way and make the structure of the mimicking behaviour more complex than in closed economy.

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. We are not able to establish that emigration of the highly skilled individuals should be prevented to maximize social welfare in all cases. However, in the important Ralwsian case, we show that in the interest of the worst-off the best-off must stay at home. It remains an open question to know whether the statement is still valid for other social preferences.

APPENDIX

Proof of Proposition 1. Let π' and γ be the Lagrange multipliers of (9) and (5) 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) imply $\gamma > 0$.

The proof employs the following lemma.

Lemma 2. *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 of the Lemma. 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,

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

Let A be the first matrix on the LHS. 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. □

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. It remains to show that $\pi'(\theta) > 0$ 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 2 and $V'_B - c' > 0$ under Assumption 3. Then $R < 0$ for $\theta \in (\theta', \theta'')$, contradicting (9). Hence, θ' does not exist. □

Proof of Proposition 3. Step 1: We use hats for the initial allocation and breves for the new one. We show that the new allocation is incentive-compatible. It is such that:

$$\breve{V}_A(\theta) = \begin{cases} V_B(\theta) - c(\theta) & \text{for every } \theta \in [\theta_-, \theta^+], \\ \hat{V}_A(\theta) & \text{otherwise.} \end{cases} \tag{A.8}$$

By construction, the participation constraint is satisfied.

(1) The allocation $(\hat{x}_A, \breve{z}_A)$ satisfies the first-order condition for incentive compatibility:

a) for $\theta \notin [\theta^*, \theta^{**}]$: because the indirect utility is unaltered, the envelope condition remains satisfied.

b) for $\theta \in [\theta_-, \theta^+]$: By definition of $\breve{V}_A$, the individuals are on the participation constraint. The envelope condition for incentive-compatibility must be satisfied. For that to be the case, we must have

$$\frac{v'(\ell_A(\theta))\,\ell_A(\theta)}{\theta} = V'_B(\theta) - c'(\theta) \text{ for } \theta \in \left[\theta_-, \theta^+\right]. \tag{A.9}$$

In addition, using the first-order condition of the individual utility maximisation programme, we note that $v'(\ell_A(\theta)) = \theta(1 - \breve{T}'(\theta\ell_A(\theta)))$ and, thus,

$$\ell_A(\theta) = v'^{-1}[\theta(1 - \breve{T}'(\theta\ell_A(\theta)))]. \tag{A.10}$$

Combining (A.9) and (A.10), we obtain

$$(1 - \breve{T}'(\theta)) \times v'^{-1}[\theta(1 - \breve{T}'(\theta))] = V'_B(\theta) - c'(\theta). \tag{A.11}$$

We study the monotonicity of the LHS with respect to $\breve{T}'$:

$$-v'^{-1}[\theta(1 - \breve{T}'(\theta))] - \frac{\theta(1 - \breve{T}'(\theta))}{v''[\theta(1 - \breve{T}'(\theta))]} < 0. \tag{A.12}$$

If $\breve{T}' = 0$, the LHS of (A.11) is equal to $v'^{-1}(\theta)$ while the RHS is equal to $v'^{-1}(\theta) - c'(\theta)$.

For every θ, there exists a solution in $\breve{T}'$. There are indeed three cases. For constant migration costs, $\breve{T}' = 0$ is the solution along the participation constraint. For decreasing migration costs, $\breve{T}' < 0$ is the solution along participation constraints. For increasing migration costs, the RHS is strictly larger than the LHS. We know that the solution is such that $\breve{T}' > 0$. It remains to establish that $\breve{T}' < 1$. If $\breve{T}' = 1$, $\ell_A(\theta) = 0$ and by assumption $V'_B(\theta) - c'(\theta) > 0$. Then, by the mean value theorem, there exists a $\breve{T}'$ in $(0, 1)$.

(2) The allocation $(\hat{x}_A, \breve{z}_A)$ satisfies the second-order condition for incentive compatibility.

Because the initial allocation $(\hat{x}_A, \hat{z}_A)$ is incentive-compatible, $\breve{z}_A$ is non-decreasing in θ outside $[\theta_-, \theta^+]$. Because the reservation utility is convex, $\breve{z}_A$ is non-decreasing in θ on $[\theta_-, \theta^+]$. Moreover, by construction, $\breve{z}_A(\theta_-) = \hat{z}_A(\theta_-)$ and $\breve{z}_A(\theta^+) = \hat{z}_A(\theta^+)$. Therefore, $\breve{z}_A$ is non-decreasing for the whole population.

Step 2: We show that the new tax schedule increases tax revenue.

For $\theta \notin [\theta_-, \theta^+]$, gross income is constant by construction. Hence, $\breve{T}(\theta) = \hat{T}(\theta)$.

For $\theta \in [\theta_-, \theta^+]$, utility is decreased. Because consumption remained fixed and the utility is decreasing in gross income, a reduction in utility is associated with an increase in gross income for everyone in this interval. Hence, $\breve{T}(\theta) = \breve{z}_A(\theta) - \hat{x}_A(\theta) > \hat{T}(\theta) = \hat{z}_A(\theta) - \hat{x}_A(\theta)$.

Therefore, $\int_\theta \breve{T}(\theta)\, dF(\theta) > \int_\theta \hat{T}(\theta)\, dF(\theta)$.

The tax adjusment is incentive-compatible and increases tax revenue, i.e., social welfare. Therefore, the initial schedule cannot be socially optimal. $\square$

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 (5) is taken into account through G' and the boundary conditions $G(\underline{\theta}) = 0$ and $G(\overline{\theta}) = 0$. 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:

$$
\begin{aligned}
\partial H^N/\partial z_A = 0 &\Leftrightarrow \iota u'_x s'_\theta - \gamma (1-s) f = 0, && \text{(A.13)}\\
\partial L^N/\partial V_A = -\iota' &\Leftrightarrow \iota'(\theta) = -\phi'_\rho(V_A) f - \iota u''_{\theta x}/u'_x - \pi' + \gamma f/u'_x, && \text{(A.14)}\\
\partial L^N/\partial G = -\gamma' &\Leftrightarrow \gamma' = 0, && \text{(A.15)}\\
\iota(\overline{\theta}) \geq 0 \ &\left(= 0 \text{ when } R(\overline{\theta}) > 0\right), && \text{(A.16)}\\
\iota(\underline{\theta}) \leq 0 \ &\left(= 0 \text{ when } R(\underline{\theta}) > 0\right), && \text{(A.17)}\\
\pi'(\theta) \geq 0,\ R(\theta) &\geq 0,\ \pi'(\theta) R(\theta) = 0, && \text{(A.18)}\\
\iota(\theta_j^-) - \iota(\theta_j^+) &= \pi(\theta_j^+) - \pi(\theta_j^-) \geq 0 \ (= 0 \text{ if } R(\theta_j) > 0). && \text{(A.19)}
\end{aligned}
$$

γ is constant and strictly positive. Because $s = 1 - T'$, $T' = \iota u'_x s'_\theta/(\gamma f)$ by (A.13). In addition, using basic calculus, $\left[1 + e^M(\theta)\right]/e^H(\theta) = -\theta s'_\theta/s$. Hence,

$$
\frac{T'}{1-T'} = -\frac{\iota u'_x}{\gamma \theta f} \frac{1 + e^M(\theta)}{e^H(\theta)}. \tag{A.20}
$$

When $\theta = \overline{\theta}$, (A.20) and (A.16) yield (23). When $\theta < \overline{\theta}$, (A.20) can be rewritten as

$$\frac{T'}{1-T'} = -\frac{\iota u'_x}{\gamma\left(1-F\left(\theta\right)\right)} \frac{1+e^M\left(\theta\right)}{e^H\left(\theta\right)} \frac{1-F\left(\theta\right)}{\theta f\left(\theta\right)}, \tag{A.21}$$

If $(.;\tau)$ means evaluation at $\left(x_A\left(\tau\right), z_A\left(\tau\right); \tau\right)$, integrating (A.14) between θ and $\overline{\theta}$ yields

$$\iota\left(\theta\right) = \iota\left(\overline{\theta}\right) + \int_{\theta}^{\overline{\theta}} \left(\phi'_\rho\left(V_A\left(\tau\right)\right) f\left(\tau\right) + \pi'\left(\tau\right) - \frac{\gamma f\left(\tau\right)}{u'_x\left(.;\tau\right)}\right) \widetilde{\Psi}_{\theta\tau} d\tau, \tag{A.22}$$

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

$$\Psi_{\theta\tau} := \frac{u'_x\left(.;\theta\right)}{u'_x\left(.;\tau\right)} \widetilde{\Psi}_{\theta\tau} = \exp_{\theta}^{\tau}\left(1 - \frac{e^M\left(\tau'\right)}{e^H\left(\tau'\right)}\right) \frac{z'_A\left(\tau'\right)}{z_A\left(\tau'\right)} d\tau'. \tag{A.23}$$

Using (A.22) and (A.23), we get

$$-\frac{\iota\left(\theta\right) u'_x\left(.;\theta\right)}{\gamma} = \int_{\theta}^{\overline{\theta}} \left[1 - \left(\phi'_\rho\left(V_A\left(\tau\right)\right) + \frac{\pi'\left(\tau\right)}{f\left(\theta\right)}\right) \frac{u'_x\left(.;\tau\right)}{\gamma}\right] \Psi_{\theta\tau} dF\left(\tau\right) - \frac{\iota\left(\overline{\theta}\right) u'_x\left(.;\theta\right)}{\gamma}, \tag{A.24}$$

that we plug in (A.21). □

Proof of Proposition 5. Citizen criterion: By definition, $W^C_{A,\rho}(\hat{\theta})$ is maximum when $\hat{\theta} = \hat{\theta}^C$, i.e. when $W^C_{A,\rho}(\hat{\theta}^C)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (9), (5). The FOC are the same as (A.13)–(A.19), except that $\overline{\theta}$ is replaced by $\hat{\theta}^C$. We then proceed as in the proof of Proposition 4.

Resident criterion: By definition, $W^R_{A,\rho}(\hat{\theta})$ is maximum when $\hat{\theta} = \hat{\theta}^R$, i.e. when $W^R_{A,\rho}(\hat{\theta}^R)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (9), (5). The FOC are the same as (A.13)–(A.19), except that (i) $\overline{\theta}$ is replaced by $\hat{\theta}^R$ and (ii) $\phi'_\rho\left(V_A\right)$ is divided by $F(\hat{\theta}^R)$. We then proceed as in the proof of Proposition 4. □

Proof of Proposition 6. 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 $\hat{\theta}$ 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 $g\left(\zeta_A\right) \geq 0$ to avoid dealing with singular solutions, where g is a $\mathcal{C}^2$-function such that $g' > 0$ and $g\left(0\right) = 0$. The Hamiltonian and Lagrangian are

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

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

$$\partial L^i/\partial \zeta_A = 0 \Leftrightarrow \eta + \kappa g'(\zeta_A) = 0, \tag{A.25}$$

$$\eta' = -\partial L^i/\partial z_A \Leftrightarrow \eta' = \iota u'_x s'_\theta - \gamma(1-s) f, \tag{A.26}$$

$$\iota' = -\partial L^i/\partial V_A \Leftrightarrow \iota' = -\phi'_\rho f - \iota u''_{\theta x}/u'_x + \gamma f/u'_x - \pi', \tag{A.27}$$

$$\gamma' = -\partial L^i/\partial G \Leftrightarrow \gamma' = 0, \tag{A.28}$$

$$\pi' \geq 0, \; R \geq 0, \; \pi' R = 0, \tag{A.29}$$

$$\kappa \geq 0, \; g(\zeta_A) \geq 0, \; \kappa g(\zeta_A) = 0, \tag{A.30}$$

$$\eta(\underline{\theta}) = \eta\left(\hat{\theta}\right) = 0, \tag{A.31}$$

$$\iota(\underline{\theta}) \leq 0 \; (= 0 \text{ if } R(\underline{\theta}) > 0), \tag{A.32}$$

$$\iota(\hat{\theta}) \geq 0 \; (= 0 \text{ if } R(\hat{\theta}) > 0), \tag{A.33}$$

$$\iota(\theta_j^-) - \iota(\theta_j^+) = \pi(\theta_j^+) - \pi(\theta_j^-) \geq (= 0 \text{ if } R(\theta_j) > 0). \tag{A.34}$$

η is continuous (see Eq. (75), p. 375, in S-S). We check that $\gamma > 0$. In addition, by continuity of η and (A.31),

$$\eta(\hat{\theta}^-)\zeta_A(\hat{\theta}) = \eta(\hat{\theta})\zeta_A(\hat{\theta}) = 0. \tag{A.35}$$

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

$$\partial \mathcal{W}^C_{A,\rho}(\hat{\theta})/\partial\hat{\theta} = \frac{\partial}{\partial\hat{\theta}}\left[\int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A)\, dF(\theta)\right] - \phi_\rho\left(V_B(\hat{\theta}) - c(\hat{\theta})\right) f(\hat{\theta}), \tag{A.36}$$

$$\partial \mathcal{W}^R_{A,\rho}(\hat{\theta})/\partial\hat{\theta} = \frac{1}{F(\hat{\theta})}\left[\frac{\partial}{\partial\hat{\theta}}\int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A)\, dF(\theta)\right] - \frac{f(\hat{\theta})}{F(\hat{\theta})}\mathcal{W}^R_{A,\rho}(\hat{\theta}). \tag{A.37}$$

Eq. (79), p. 376, in (S-S) gives the value of the square brackets on the RHS of (A.36) and (A.37):

$$H(\hat{\theta}^-) + \left[\pi(\hat{\theta}) - \pi(\hat{\theta}^-)\right] R'(\hat{\theta}). \tag{A.38}$$

Using the continuity of x_A, z_A, f, V_A, (A.35), (A.34), $T = z_A - x_A$, and the fact that (4) is active at $\hat{\theta}$, (A.36) and (A.37), we obtain the expressions in the proposition. □

References

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

Blackorby, C., C. Brett, and A. Cebreiro (2007): "Nonlinear Taxes for Spatially Mobile Workers," *International Journal of Economic Theory*, 3, 57–74.

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): "Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility," Vanderbilt University Working Paper 08-W12.

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.

MYLES, G. (1995): *Public Economics*. Cambridge University Press, 2002 edn., Chapter 5.

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.

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.

SEIERSTAD, A., AND K. SYDSAETER (1987): *Optimal Control Theory with Economic Applications*. North-Holland, Amsterdam.

Simula, L. (2010): "Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties," *Social Choice and Welfare*, 35, 199–220.

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

——— (2010): "Optimal Income Tax under the Threat of Migration by Top-Income Earners," *Journal of Public Economics*, 94, 163–173.

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.