

Optimal Non-Linear Income Tax when Highly Skilled Individuals Vote with their Feet*

Laurent SIMULA
EHESS, GREQAM-IDEP[†]

Alain TRANNOY
EHESS, GREQAM-IDEP[‡]

October 13, 2006

Abstract

This paper examines how allowing individuals to emigrate to pay lower taxes changes the optimal non-linear income tax scheme in a Mirrleesian economy. Type-dependent participation constraints are borrowed from contract theory. 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 social criteria are distinguished according to the agents whose welfare matters. Mobility significantly alters the closed-economy results qualitatively, but also quantitatively as verified by simulations. A curse of the middle-skilled occurs in the first-best. In the second-best, the middle-skilled can support the highest average tax rates and the marginal tax rates can be negative. Moreover, preventing emigration of the highly-skilled is not necessarily optimal.

Keywords: Optimal Taxation, Income Tax, Emigration, Participation Constraints

JEL classification: H21; H31; D82; F22.

*We are grateful to Thomas Aronsson, David Bevan, Chuck Blackorby, Philippe Choné, David de la Croix, Jeremy Edwards, Jonathan Hamilton, Jean Hindriks, Mathias Hungerbühler, Laurence Jacquet, Karolina Kaiser, Jean-Marie Lozachmeur, Hamish Low, François Maniquet, Michael S. Michael, James Mirrlees, Gareth Myles, Frank Page, Pierre Pestieau, Panu Poutvaara, Emanuela Sciubba, John Weymark, David Wildasin for their suggestions. Thanks are due to Hélène Couprie and Gwenola Trotin. Our work has benefited from comments of seminar participants at GREQAM, the University of Cambridge, PET, ESWC, ASSET, HECER Workshop on Fiscal Federalism, Séminaire Economique de Louvain, Doctorales ADRES, Journées Louis-André Gérard-Varet, ESEM, IIPF Congress, EDGE Jamboree. Laurent Simula would like to thank the Faculty of Economics, University of Cambridge, for excellent research facilities during his stays as an EDGE Fellow. The usual caveats apply.

[†]Paris-Jourdan Sciences Economiques, ENS, 48 boulevard Jourdan, 75014 Paris, France. Email: Laurent.Simula@ehess.fr

[‡]GREQAM-EHESS, Centre de la Vieille Charité, 2 rue de la Charité, 13236 Marseille cedex 02, France. Email: Alain.Trannoy@ehess.univ-mrs.fr

1. INTRODUCTION

In Mirrlees's (1971) seminal article, migration is supposed to be impossible. However, according to Mirrlees himself, "since the threat of migration is a major influence on the degree of progression in actual tax systems, at any rate outside the United States, this is [an] assumption one would rather not make". This threat is certainly even more topical after 25 years of increasing international openness. A first source of worry in many OECD countries is the departure of some of their highly skilled individuals for tax havens. It is considered that the US\$11.5 trillion of assets which are held offshore worldwide by high net-worth individuals represent a loss in taxes exceeding US\$ 255 billion each year (TRL, 2005). A second source of concern is the challenge faced by some developed high-tax countries because their neighbours have less redistributive objectives. For instance, about 34 000 income taxpayers have left France each year since 2000. These individuals paid three times more taxes than the average taxpayer and 70% of them chose to relocate to another EU country, like the UK, Belgium, Luxembourg, Switzerland, or to North America (DGI, 2005), mainly to countries where income taxes are lower. This suggests that international differences in income taxes are *one* of the determinants of the migration decision. This motivation for leaving the home country is in accordance with John Hicks's idea that migration decisions are based on the comparison of earnings opportunities across countries, net of moving costs, which is the cornerstone of practically all modern economic studies of migration (Sjaastad, 1962, Borjas, 1999).

The mobility of highly skilled individuals for tax purposes induces both losses in taxes and in productive capacity in the left countries. It differs from the brain drain because its key parameter is not the change in individual productivity resulting from emigration. Governments have also a more limited set of instruments than when they face tax evasion (e.g. Chander and Wilde (1998)). They have indeed few alternatives but to reduce taxes to prevent the departure of highly skilled individuals: in a nutshell, they can use "carrots" but no "sticks". As a result, a specific conflict arises between the desire to maintain national income per capita in keeping taxes down and the aim to sustain the redistribution programme. The possibility that highly skilled individuals vote with their feet with a view to paying lower taxes appears therefore as a *new constraint* on the design of the optimal income tax schemes.

This paper studies the optimal non-linear income tax in a Mirrleesian economy ("home country") the citizens of which have type-dependent outside options consisting in emigrating to a *less* redistributive country ("foreign country"). The government wants to redistribute incomes from the more to the less productive individuals as in Mirrlees's (1971) model, but has also to take account of participation constraints for the individuals it wants to keep at home. The optimal income tax papers taking individual mobility into account have used models with no leisure-consumption choice (Mirrlees, 1982, Hindriks, 1999, Osmundsen, 1999), considered a world with two classes of individuals and lump-sum taxes (Leite-Monteiro, 1997), focused on linear taxes (Wilson, 1980, 1982, Simula and Trannoy, 2006), or employed Stiglitz's (1982) self-selection approach with two types of individuals (Huber, 1999, Hamilton and Pestieau, 2005, Piaser, 2003).

Among them, Leite-Monteiro (1997), Hindriks (1999), Huber (1999) and Piaser (2003) have adopted the point of view of tax competition. Hamilton and Pestieau (2005) have concentrated on migration equilibria.

This paper considers optimal non-linear income taxation when there is a continuum of individuals differing in productivity as well as migration costs and facing consumption-leisure choices in the absence of unemployment. It examines how the foreign income tax policy influences the optimal income tax schedule implemented at home when agents vote with their feet. Since at this stage we do not want to study the reverse question, the foreign tax policy has to be exogenous with respect to the policy implemented at home¹. The only coherent case is that in which the foreign government chooses the *laissez-faire*. Otherwise, the tax revenue constraint abroad would usually be slack or violated after the arrival of individuals from the home country, so the income tax scheme abroad should be adapted². In addition, we consider that all individuals are initially in the home country because we are not interested herein in emigration of low-skilled individuals to the home country. Finally, we assume that both countries have the same constant-returns-to-scale production function as we do not want individual productivity to depend on the country of residence.

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

We consider that an individual chooses to emigrate if his indirect utility at home is lower than his best outside option. Since many empirical studies have shown that the propensity to migrate increases with the skill level (Sahota, 1968, Schwartz, 1973, Gordon and McCormick, 1981, Nakosteen and Zimmer, 1980, Inoki and Surugan, 1981), it is sensible to assume that more productive individuals should have more attractive outside options. Consequently, the reservation utility, i.e. the minimum utility the domestic government should give to keep an individual at home, should be increasing in productivity. We ensure it is the case by assuming that the costs of migration, expressed in terms of utility, monotonically depend on productivity and do not increase faster than the indirect utility in B . Monotonicity of these costs implies that productivity is in fact the sole parameter of heterogeneity within the population.

Since individuals have type-dependent outside options, A 's optimal income tax scheme must satisfy *type-dependent participation constraints*. We borrow these constraints from contract theory (see Lewis and Sappington (1989), Maggi and Rodriguez-Clare (1995), and Jullien (2000)) and introduce them in the optimal non-linear income tax problem à la Mirrlees to examine at

¹Studying the interaction between two non-linear income tax countries raises difficult problems. For instance the revelation principle would generally vanish in A and B (See Page and Monteiro (2003) for non-linear pricing).

²It should be noted that the main properties derived below remain valid when B 's government implements a given non-linear income tax schedule. What really matters is that B 's tax policy is given. We only assume that B is a *laissez-faire* country for theoretical coherence.

which productivity levels they should be binding to obtain the highest social welfare from our three social criteria³. We have shown in Simula and Trannoy (2006) that linear taxes lack degrees of freedom when such constraints are taken into account, which explains why we consider non-linear taxes herein.

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 from both Citizen and Resident criteria, which coincide therefore with the National criterion at the optimum. There is a curse of the middle-skilled workers at the optimum, instead of the curse of the highly skilled obtained in closed economy (Mirrlees, 1974). Indeed, it is no longer possible to demand as much work as without mobility from the highly skilled individuals, so the productive rent is extracted to the maximum from the most productive individuals among those insufficiently talented to threaten to emigrate. However, these middle-skilled workers cannot be taxed at will because they would otherwise threaten to emigrate. Consequently, the redistribution in favour of the low-skilled individuals has to be reduced.

When each individual's productivity is private information (*second-best*), two qualitative properties of the optimal marginal tax rates are lost: they can be non-positive at interior points and strictly negative at the top. Consequently, individual mobility does not only render the tax schedule less progressive, but can also make the tax function *decreasing*. In fact, the small tax reform perturbation around the optimal tax scheme used by Piketty (1997) and Saez (2001) has an additional *participation effect* on social welfare, which favours a decrease in the optimal marginal tax rates even for individuals below the productivity levels where the individuals threaten to emigrate. This new effect results in changes in Mirrlees's (1971) and Diamond's (1998) formulae to ensure that the optimal *average* tax rates are compatible with the participation constraints of the individuals threatening to emigrate. In addition, the interaction between the type-dependent participation constraints and the incentive compatibility conditions can give rise to countervailing incentives, in which case less skilled individuals want to mimic more skilled individuals because the latter have more appealing outside options. Countervailing incentives cause an indirect social cost of the presence in A of the highly-skilled individuals. The Citizen and Resident criteria allow us to consider whether it is not too expensive in terms of social welfare to implement a tax scheme which prevents emigration of the highly skilled workers. When the indirect cost due to countervailing incentives prevails over the benefits of them staying in A , implementing a tax schedule inducing them to emigrate increases social welfare.

Numerical simulations calibrated with French data are provided to *quantify* to which extent individual mobility alters the whole optimal non-linear income tax schedule. They emphasize that the optimal marginal and average tax rates are significantly altered even if there are very few people threatening to emigrate. In particular, the optimal average tax rates can start to

³See Osmundsen (1999) for a direct application of the framework developed by Maggi and Rodriguez-Clare (1995) to a tax problem in a country the individuals of which share their working time between home and abroad. By construction, this model does not allow to investigate how individual mobility alters the issues raised in the closed-economy optimal income tax literature.

decrease far below the income level from which potential mobility occurs. There is consequently a *second-best curse of the middle-skilled*, consisting in them being taxed the most in proportion to gross income. This curse is even stronger when migration costs are decreasing in productivity: the optimal income tax schedule is not only less progressive but also such that the highly-skilled pay taxes lower than the middle-skilled.

The paper is organized as follows. Section 2 sets up the model. Section 3 examines the first-best optimal allocations. Section 4 studies the properties of the second-best optimal allocations. Section 5 provides numerical simulations on French data. Section 6 concludes. Most proofs are relegated to the Appendix.

2. THE MODEL

The world consists of two countries, A and B . All individuals are initially living in A . A 's government implements a redistributive tax policy and B is committed to being a laissez-faire country. The governments provide no public goods. Both countries have the same production function with constant returns to scale. Hence, productivity levels, equal to wages in the absence of taxation, are independent of the country in which the individuals are working.

Individuals differ in productivity θ . An individual with productivity θ is called a θ -individual. The cumulative distribution function of θ , denoted F , is common knowledge. It is defined on a closed interval $[\underline{\theta}, \bar{\theta}] \subseteq \mathbb{R}^+$ where it admits a continuous and strictly positive density f^4 .

2.1. Individual Behaviour

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

Assumption 1. U is a \mathcal{C}^2 strictly concave function such that $U_x > 0$, $U_l < 0$ and $U \rightarrow -\infty$ as $x \searrow 0$ or $l \searrow \bar{l}$.

Assumption 2. *Leisure is a normal good.*

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

$$u(x, z; \theta) := U(x, z/\theta) \quad (1)$$

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

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

⁴Optimization programmes consisting in choosing the socially optimal upper productivity level in A are addressed below. If $\bar{\theta}$ were allowed to be $+\infty$, we would have to choose one convergence criterion among the different available ones and introduce additional assumptions.

Each individual decides about the optimal amount of consumption and labour to maximize his utility subject to his budget constraint. Using a tax function $T(\theta, l)$, A 's government can arrange that an individual with gross income z has disposable income $z - T(\theta, l)$ in A . Consequently, the utility maximization programme in A , $\max_{(x,l) \in \mathcal{X}} \{U(x, l) \text{ s.t. } x = z - T(\theta, l)\}$, defines implicitly the consumption and labour supply functions in A , $x_A(\theta)$ and $l_A(\theta)$ respectively⁵. The indirect utility in A is

$$V_A(\theta) := U(x_A(\theta), l_A(\theta)). \quad (3)$$

We call e^H and e^M the Hicksian and Marshallian elasticities of labour supply with respect to the net-of-tax wage rate. The utility maximization programme in B , $\max_{(x,l) \in \mathcal{X}} \{U(x, l) \text{ s.t. } x = z\}$, defines implicitly the consumption and labour supply functions in B , $x_B(\theta)$ and $l_B(\theta)$ respectively. The indirect utility in B is

$$V_B(\theta) := U(x_B(\theta), l_B(\theta)), \quad (4)$$

which is strictly increasing in θ .

2.2. Emigration and Participation Constraints

An individual leaving A pays a strictly positive *migration cost* c . This cost corresponds to a "time-equivalent" loss in utility, due to different material and psychic costs of moving: application fees, transportation of persons and household's goods, forgone earnings, costs of speaking a different language and adapting to another culture, costs of leaving one's family and friends, etc. "[These migration] costs probably vary among persons [but] the sign of the correlation between costs and wages is ambiguous" (Borjas, 1999, p. 12). We consider that they depend on productivity and that their distribution is known to A 's government. In addition:

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

Monotonicity implies that A 's government knows $c(\theta)$ when it knows θ . Migration costs are allowed to be either non-increasing or non-decreasing provided they do not increase faster than the laissez-faire utility V_B . No restrictions are placed on their level.

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 thus to considering that the outside opportunities are increasing in productivity. The *location rent* of a θ -individual is the excess of his indirect utility in A over his reservation utility, i.e.

$$R(\theta) = V_A(\theta) - V_B(\theta) + c(\theta). \quad (5)$$

⁵It is assumed that the function $x(\theta, l) := z(\theta, l) - T(\theta, l)$ is upper semi-continuous.

An individual stays in A if and only if

$$R(\theta) \geq 0, \quad (6)$$

and therefore leaves A if and only if $R(\theta) < 0$. When $\{\theta \in [\underline{\theta}, \bar{\theta}] : (6) \text{ binding}\}$ is non-empty, we call θ^* its *infimum*.

A *citizen* is defined as an individual born in A , so all individuals have A 's citizenship. Individuals are committed to working in the country where they live. Since the focus is on emigration of highly skilled individuals, we consider that there is a partition of citizens between A and B , with the low-skilled individuals being in A . In addition, emigration of a set of individuals of measure zero does not capture any economic phenomenon, so we assume that A 's resident population is compact.

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

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

$$\int_{\underline{\theta}}^{\hat{\theta}} (z_A - x_A) dF(\theta) \geq 0. \quad (\text{TR})$$

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

2.3. Social Criteria

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

The social objective is more difficult to specify than in closed economy. Indeed, it does not only depend on ρ which is captured through an isoelastic function defined by $\phi_\rho : \mathbb{R}^{++} \rightarrow \mathbb{R}$, $\phi_\rho(U) = U^{1-\rho}/(1-\rho)$ for $\rho \neq 1$ and $\phi_1(U) = \ln U$ for $\rho = 1$, but also on the answers to the following questions. First, should we maximize total or average social welfare? We consider that the government is interested in social welfare per capita because we want to be able to compare allocations differing in population size. Second, who are the agents whose welfare is to count? At least three social criteria can be proposed, each of which corresponds to a specific answer.

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

$$W_{A,\rho}^N := \int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A(\theta)) dF(\theta) \text{ and } \hat{\theta} = \bar{\theta}. \quad (7)$$

This objective corresponds to the mercantilist idea, formulated by Bodin (1578), that "the only source of welfare is mankind itself". Emigration should therefore be prevented to keep the state prosperous.

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

$$W_{A,\rho}^C(\hat{\theta}) := \int_{\underline{\theta}}^{\hat{\theta}} \phi_{\rho}(V_A(\theta)) dF(\theta) + \int_{\hat{\theta}}^{\bar{\theta}} \phi_{\rho}(V_B(\theta) - c(\theta)) dF(\theta). \quad (8)$$

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 residence⁶. When this objective is chosen, the optimal tax function depends on the choice of $\hat{\theta}$ and determines an allocation of A 's citizens between A and B . Hence, A 's resident population is endogenous while the set of agents the welfare of whom matters is exogenously fixed.

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

$$W_{A,\rho}^R(\hat{\theta}) := \frac{1}{F(\hat{\theta})} \int_{\underline{\theta}}^{\hat{\theta}} \phi_{\rho}(V_A(\theta)) dF(\theta). \quad (9)$$

This criterion is based on the idea that a public policy should take the welfare of all taxpayers into account. Consequently, the welfare of the citizens living in B does not count. When this objective is chosen, the tax function as well as the set of agents whose welfare is to count depend on the choice of $\hat{\theta}$. $W_{A,\rho}^R(\hat{\theta})$ 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. However, this paradox is not a matter herein because we are focusing on emigration of the highest skilled individuals initially living in A , whose utility will be shown to be greater than average utility in A .

3. FIRST-BEST OPTIMAL ALLOCATIONS

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

⁶In France, the 14th Article of the Declaration of the Rights of Man and of the Citizen, which has constitutional value, provides that: "All citizens have the right to vote, by themselves or through their representatives, for the need for the public contribution, to agree to it voluntarily, to allow implementation of it, and to determine its appropriation, the amount of assessment, its collection and its duration". For example, twelve senators represent the French citizens living abroad.

⁷In other words, a population problem consisting in "different number choices" (Parfit, 1984) is embedded in the optimal income tax problem.

and differentiable almost everywhere.

The indirect utility if A were a closed economy, $V_A^{cl}(\theta)$, is used as a benchmark. When ρ is finite, it is decreasing in θ at the social optimum (Mirrlees, 1974): there is therefore a *curse of the highly skilled workers*. When ρ is infinite, all individuals receive the same utility level. In this section, we assume $V_A^{cl}(\bar{\theta}) < V_B(\bar{\theta}) - c(\bar{\theta})$ since otherwise the participation constraints would never be active.

3.1. National criterion

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

Problem 1 (National Criterion, First-Best). *Find $(x, l) \in \mathcal{X}$ to maximize $W_{A,\rho}^N$, with $\hat{\theta} = \bar{\theta}$, subject to*

$$R(\theta) \geq 0 \text{ for } \theta \leq \hat{\theta} \quad (\text{PC})$$

and (TR).

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

Proof. See A.1 in the Appendix. \square

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

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

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

3.2. Citizen and Resident Criteria

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

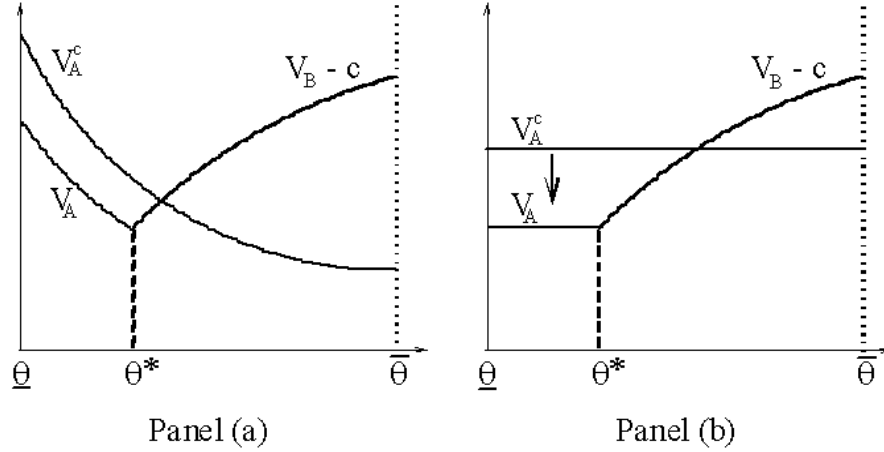


Figure 1: The curse of the middle-skilled workers

Problem 2 (Citizen and Resident Criteria, First-Best). Find $(x, l) \in \mathcal{X}$ and $\hat{\theta} \in [\underline{\theta}, \bar{\theta}]$ to maximize $W_{A,\rho}^i(\hat{\theta})$, $i = \{C, R\}$, subject to (PC) and (TR).

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

Proof. See A.1 in the Appendix. □

The flavour of the proof is as follows. Let $\hat{\theta} < \bar{\theta}$ be socially optimal. The individuals with productivity $\hat{\theta}$ are indifferent between A and B , i.e. $R(\hat{\theta}) = 0$, and those with productivity greater than $\hat{\theta}$ emigrate to B . Note that it is always *feasible* to make the latter relocate to A , without reducing the indirect utilities of A 's residents, in giving them their laissez-faire utility V_B (or a bit more than their reservation utility). We show in the Appendix that the laissez-faire utility of individuals with productivity $\theta > \hat{\theta}$ is greater than the average utility of A 's resident population. This feasible change increases social welfare, which contradicts the premise. The social optimum corresponds therefore to the corner solution as regards the allocation of individuals between A and B .

4. SECOND-BEST OPTIMAL ALLOCATIONS

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

4.1. Statement of the Problem

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

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

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

Assumption 5 (Single-Crossing). $s'_\theta(x_A(\theta), z_A(\theta); \theta) < 0$.

Under Assumption 5, (IC) is equivalent to:

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

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

The proof of this equivalence is standard and is omitted. (FOIC) is an envelope condition specifying how the indirect utility V_A must locally change. Since $V'_A \geq 0$, V_A cannot be V -shaped. (SOIC) is a global monotonicity condition of gross income.

(SOIC) implies that the most general class of direct revelation mechanisms (x_A, z_A) to consider is the class of almost everywhere differentiable functions. Here, we restrict the analysis to the class of functions which are continuous and piecewise differentiable. Hence, we look for the optimal tax function among the admissible functions which are continuous but can exhibit kinks at a finite number of points corresponding to jumps of the marginal tax rate. Consequently, (SOIC) is equivalent to

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

In addition $R(\theta)$ is continuous. This implies that θ^* is actually the minimum productivity level at which individuals threaten to emigrate and that the $\hat{\theta}$ -individuals are indifferent between living in A or in B provided $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$.

Since A 's government does not know who are the agents for whom the location rent $R(\theta)$ is zero, the participation constraints and the incentive compatibility conditions have to be taken simultaneously into account for all A 's residents⁸. The second-best optimal non-linear income tax problems read thus as follows.

Problem 3 (Second-Best). Find $T(z_A)$ to maximize $W_{A,\rho}^i$, $i = \{N, C, R\}$, subject to (i) (FOIC), (SOIC'), (PC), (TR); (ii) $\hat{\theta} = \bar{\theta}$ when $i = N$ and $\hat{\theta} \in [\underline{\theta}, \bar{\theta}]$ otherwise.

⁸If the participation constraints (PC) were not type-dependent, it would be necessary and sufficient to check that they are satisfied at $\underline{\theta}$ since (FOIC) ensures that the optimal utility path is non-decreasing.

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

In solving Problem 3, we assume that the adjoint variables have a finite number of jump discontinuities and are \mathcal{C}^1 elsewhere. This ensures that the necessary conditions are equivalent to the sufficient ones, under appropriate concavity restrictions. Without this assumption, it would be very difficult to say anything about the optimal tax schemes. For later reference, we call ι the adjoint variable associated with (FOIC) and $\pi' \geq 0$ the Lagrange multiplier of (PC), which corresponds to the shadow price of a marginal increase in the reservation utility at θ . Let also

$$\pi(\theta) := \pi(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} \pi'(\tau) d\tau \quad (10)$$

be the shadow price of a uniform marginal increase in the reservation utility for all $\theta' \geq \theta$. The function π , with derivative π' almost everywhere, is non-decreasing.

4.2. Optimal Tax Schedule for the Individuals Threatening to Emigrate

Before looking at a specific social criterion, we derive properties which are satisfied by all optimal tax schemes for the individuals threatening to emigrate. The first one deals with the sign of the tax function.

Property 1. *Let (PC) be active at θ . Then, $T(z_A(\theta)) > 0$.*

It is always possible to levy taxes on the individuals who threaten to emigrate. Indeed, let (PC) be active at some θ . Since $c(\theta) > 0$ under Assumption 3, $V_B(\theta) > V_A(\theta)$ and thus $T(z_A(\theta)) > 0$.

We now examine the case where the participation constraints are active on an interval of positive length $I \subset [\underline{\theta}, \bar{\theta}]$. By definition,

$$R(\theta) \equiv 0 \text{ for } \theta \in I. \quad (11)$$

Consequently, $V'_A(\theta) = V'_B(\theta) - c'(\theta)$ for $\theta \in I$. Hence the rate of increase of the indirect utility the government has to give to the individuals so that they reveal their private information, is equal to the slope of the reservation utility on I . In addition, employing (FOIC) and rearranging yield

$$z_A(\theta) = -\theta [V'_B(\theta) - c'(\theta)] / u'_z \text{ for } \theta \in I, \quad (12)$$

and by differentiation,

$$z'_A(\theta) = \frac{[V'_B(\theta) - c'(\theta)] \left\{ \theta (u''_{xz} x'_A + u''_{\theta z}) - \left(1 + \theta \frac{V''_B(\theta) - c''(\theta)}{V'_B(\theta) - c'(\theta)} \right) u'_z \right\}}{(u'_z)^2 - \theta (V'_B(\theta) - c'(\theta)) u''_{zz}} \text{ for } \theta \in I. \quad (13)$$

The second-order condition for incentive compatibility (SOIC') can only be satisfied on I if the curly bracket in (13) is positive. It thus restricts the set of intervals where (PC) can be binding.

Property 2. *(PC) can be active on an interval of positive length I only if*

$$\frac{\theta (u''_{xz} x'_A + u''_{\theta z})}{u'_z} \leq 1 + \theta \frac{V''_B(\theta) - c''(\theta)}{V'_B(\theta) - c'(\theta)} \text{ for } \theta \in I. \quad (14)$$

The elasticity of the marginal reservation utility, evaluated at θ , appears on the RHS. The LHS captures the behavioural response of the θ -individuals to a slight change in their reservation utility.

To have further insight, we turn to quasilinear-in-consumption preferences,

$$u(x_A, z_A; \theta) = x_A - v(z_A/\theta), \text{ with } v' > 0 \text{ and } v'' > 0. \quad (15)$$

Since $e^H = v'/(lv'')$, $u''_{xz} = 0$, $u''_{\theta z} = (1 + lv''/v') v'/\theta^2$ and $u'_z = -v'/\theta$, (14) reads

$$-1 - \frac{1}{e^H(\theta)} \leq 1 + \theta \frac{V''_B(\theta) - c''(\theta)}{V'_B(\theta) - c'(\theta)},$$

the LHS of which is strictly negative.

Property 3. *Let preferences be quasilinear in consumption and consider an interval of positive length where (PC) is active. Then, there is no bunching on this interval when $V_B - c$ is convex.*

When the elasticity of labour supply is constant ($e^H(\theta) = e$), the disutility of labour can be described by

$$v(l) = l^{1+1/e} / (1 + 1/e). \quad (16)$$

Hence, $V_B(\theta) = \theta^{1+e} / (1 + e)$. The first-order condition for individual utility maximization in A yields $l_A(\theta) = \theta^e [1 - T']^e$. In addition, by (12), $l_A(\theta) = \theta^{\frac{e}{1+e}} [\theta^e - c'(\theta)]^{\frac{e}{1+e}}$. Using both expressions of $l_A(\theta)$ and solving for T' , one gets:

Property 4. *Let preferences be quasilinear in consumption, e be the constant elasticity of labour supply and I be an interval of positive length where (PC) is active. Then,*

$$T' = 1 - \theta^{-\frac{e}{1+e}} [\theta^e - c'(\theta)]^{\frac{1}{1+e}} \text{ for } \theta \in I. \quad (17)$$

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

Property 5. *Consider the same situation as in Property 4. Then,*

$$T'(z_A(\theta)) \gtrless 0 \Leftrightarrow c'(\theta) \gtrless 0 \text{ for all } \theta \in I. \quad (18)$$

When the costs of migration are *non-increasing*, the theorem stating that the optimal tax function is strictly increasing at *all* income levels (Seade, 1982) does no longer hold. When the costs of migration are *strictly decreasing* in productivity, the optimal marginal tax rates faced by the individuals threatening to emigrate are *strictly negative*⁹. This property contrasts with two results obtained in closed economy, stating that: (i) the optimal marginal tax rates are non-negative (Mirrlees, 1971); (ii) the optimal marginal tax rate is zero at the top (Sadka, 1976, Seade, 1977). The next corollaries of Property 5 provide further details about these significant changes.

The first one considers constant migration costs on I . Then, by Property 5, $l_A(\theta) = \theta^e$ on I . In addition, $V_A = V_B - c = \theta^{1+e}/(1+e) - c$ and, by (15), $T = \theta l_A - V_A - l_A^{1+1/e}/(1+1/e)$. Consequently, $T(z_A(\theta)) = c(\theta)$ on I .

Corollary 1. *Consider the same situation as in Property 4 and let $c'(\theta) = 0$ on I . Then, the optimal tax function has a flat section corresponding to potentially mobile individuals paying taxes equal to their costs of migration.*

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

Corollary 2. *Consider the same situation as in Property 4 and let $c'(\theta) < 0$ on I . Then, the optimal average tax rate and the optimal tax function are strictly decreasing in productivity on I .*

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

4.3. National Criterion

We study the impact of the threat of migration on the optimum tax scheme in A when A 's government adopts the National criterion. To this aim, Mirrlees's formula is extended to the case where agents are allowed to vote with their feet. This formula gives the optimal marginal tax rates in the absence of bunching.

⁹An example of optimal income tax schedule with strictly negative marginal tax rates is provided in the simulation section.

4.3.1. A First Pass

We first look at a very simple situation to illuminate the basic economic relations which determine the optimal marginal tax rates¹⁰. A 's government adopts the Rawlsian maximin and there are agents with zero productivity ($\theta = 0$). The social objective is thus to maximize the social benefit given to the latter individuals. Preferences are quasilinear in consumption, which captures the absence of income effects on labour supply, and the elasticity of labour supply is constant ($e^H(\theta) = e$). Migration costs are constant, equal to $c(\theta) = \bar{c}$. In addition, attention is restricted to the cases where (PC) is only active on non-degenerate intervals; by Property 1, taxes paid by individuals threatening to emigrate amount to \bar{c} .

We adopt the methodology employed by Piketty (1997) and Saez (2001) to derive the optimal marginal tax rates. We consider the effects of a small increase dT in the optimal marginal tax rates for income between z and $z + dz$. This tax perturbation has three effects on social welfare, captured through changes in tax revenue G . The first two effects are the same as in a closed economy. The third one is new.

Mechanical effect: All individuals with income greater than z pay additional taxes $dTdz$. Since their proportion is given by $1 - F(\theta_z)$, the effect on tax revenue is

$$dG^+ = (1 - F(\theta_z)) \times dTdz. \quad (19)$$

Elasticity effect: The net-of-tax wage rate of the individuals with income between z_A and $z_A + dz_A$ decreases from $\theta_z(1 - T')$ to $\theta_z(1 - T' - dT)$, i.e. by $dT/(1 - T')$ %. The reduction in gross income z for the $f d\theta$ individuals is therefore $e \times dT/(1 - T') \times z f d\theta$. This results in a loss in tax revenue $dG_1^- = T' \times e \times dT/(1 - T') \times z f d\theta$; since $d\theta = dz/[l(1 + e)]$ by definition of e ,

$$dG_1^- = -\frac{T'}{1 - T'} \frac{e}{1 + e} \times \theta f \times dTdz. \quad (20)$$

Participation effect: Individuals with income greater than z for whom (PC) is active have to be compensated for the increase in taxes. Since preferences are quasilinear in consumption and migration costs constant, an amount $dTdz$ has to be given to them. Let μ_f be a measure generated by the density f on $[\underline{\theta}, \bar{\theta}]$ and Θ^{PC} be the productivity set where (PC) is active. Provided (PC) is only active on non-degenerate intervals, the effect on tax revenue is

$$dG_2^- = -\mu_f([\theta_z, \bar{\theta}] \cap \Theta^{PC}) \times dTdz. \quad (21)$$

At the social optimum, the small tax reform perturbation has no first-order effect. Consequently $dG^+ + dG_1^- + dG_2^- = 0$. Since θ_z has been chosen arbitrarily, the following result is obtained.

¹⁰The simplifying assumptions are made to get the flavour of the result stated in Proposition 4 in the most general case. In addition, optimal tax schemes satisfying all of them are shown to exist in the simulation section.

Proposition 3. *Assume (PC) is not active at isolated points. Then, when preferences are quasilinear in consumption, $e^H(\theta) = e$ and $c(\theta) = \bar{c}$, the Rawlsian optimal marginal tax rates are given by*

$$\frac{T'}{1-T'} = \left(1 + \frac{1}{e}\right) \frac{1-F(\theta)}{\theta f(\theta)} \left(1 - \frac{\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC})}{1-F(\theta)}\right) \text{ for } \theta < \bar{\theta}, \quad (22)$$

and $T' = 0$ for $\theta = \bar{\theta}$, provided there is no bunching at the optimum.

Proof. See A.2 in the Appendix for a formal proof. \square

When $\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC}) = 0$, (22) reduces to the formula derived by Piketty (1997) in closed economy. Three points are worth noting. First, since e^H is constant, it can be rewritten as

$$\frac{T'}{1-T'} = \frac{T'_{cl}}{1-T'_{cl}} \left(1 - \frac{\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC})}{1-F(\theta)}\right), \quad (23)$$

where T'_{cl} is the Rawlsian marginal tax rate the θ -individuals would face in A in the absence of individual mobility. As $\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC}) \leq 1 - F(\theta)$, the marginal tax rates faced by *all* individuals, and not only those of the individuals threatening to emigrate are reduced in the presence of potential mobility¹¹. Second, $\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC}) / (1 - F(\theta))$ is increasing in θ for $\theta \leq \theta^*$. Hence, the closer θ to θ^* , the greater the reduction in marginal tax rates. Third, if the participation constraints are active at any given $\theta < \bar{\theta}$, $T'(z_A(\theta)) = 0$ by Property 4. But, by (22), $T'(z_A(\theta)) = 0$ if and only if $\mu_f([\theta, \bar{\theta}] \cap \Theta^{PC}) = 1 - F(\theta)$, that is if and only if the participation constraints are active for all individuals with productivity greater than θ . Consequently, the participation constraints separate the population into two intervals: they are slack for $\theta < \theta^*$ and binding for $\theta \geq \theta^*$.

4.3.2. The General Case

We extend the previous analysis by relaxing all simplifying assumptions, except the absence of bunching.

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

$$\frac{T'(z_A(\theta))}{1-T'(z_A(\theta))} = A(\theta) B(\theta) C(\theta) \text{ for } \theta < \bar{\theta}, \quad (24)$$

¹¹Hence, the taxes net of the social benefit given to the worst-off individuals $T(z_A(\theta)) - T(z_A(\underline{\theta}))$ are reduced for everyone compared to the closed-economy ones.

where $A(\theta) := \frac{1 + e^M(\theta)}{e^H(\theta)}$, $C(\theta) := \frac{1 - F(\theta)}{\theta f(\theta)}$ and $B(\theta) := B_1(\theta) - B_2(\theta) - B_3(\theta)$ with

$$\begin{aligned} B_1(\theta) &:= \frac{1}{1 - F(\theta)} \int_{\theta}^{\bar{\theta}} \left[1 - \frac{\phi'_{\rho}(V_A(\tau)) u'_x(x_A, z_A; \tau)}{\gamma} \right] \Psi_{\theta\tau} dF(\tau), \\ B_2(\theta) &:= \frac{1}{1 - F(\theta)} \int_{\theta}^{\bar{\theta}} \frac{\pi'(\tau) u'_x(x_A, z_A; \tau)}{\gamma} \Psi_{\theta\tau} d\tau, \quad \pi' \geq 0 \quad (= 0 \text{ if } R(\theta) > 0), \\ B_3(\theta) &:= -\frac{1}{1 - F(\theta)} \frac{\iota(\bar{\theta}) u'_x(x_A, z_A; \bar{\theta})}{\gamma}, \quad \iota(\bar{\theta}) \geq 0 \quad (= 0 \text{ if } R(\bar{\theta}) > 0), \end{aligned}$$

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

In addition,

$$\frac{T'(z_A(\bar{\theta}))}{1 - T'(z_A(\bar{\theta}))} = \frac{A(\bar{\theta})}{\bar{\theta} f(\bar{\theta})} \frac{\iota(\bar{\theta}) u'_x(x_A, z_A; \bar{\theta})}{\gamma} \leq 0 \quad (= 0 \text{ if } R(\bar{\theta}) > 0) \quad (25)$$

and $\gamma > 0$.

Proof. See A.2 in the Appendix. □

Proposition 4 extends Mirrlees's (1971) optimal income tax formula to take the threat of migration into account, using behavioural elasticities as in Saez (2001). (24) reflects the trade-off between efficiency and equity when the government has decided to maintain the national productive capacity to the maximum in preventing its citizens from leaving the country. $A(\theta)$ and $C(\theta)$ are the usual efficiency and demographic factors, respectively. However, the value of $A(\theta)$ is usually not the same whether the individuals can or cannot vote with their feet since it depends on gross income which is endogenous. The factor $B(\theta)$, which combines efficiency and equity, is the only factor which does not write as in Mirrlees's formula, in which the RHS of (24) 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.

Alternatively, $B_1(\theta) - B_2(\theta)$ can be written as

$$B_1(\theta) - B_2(\theta) = \int_{\theta}^{\bar{\theta}} [1 - g(\tau)] \Psi_{\theta\tau} dF(\tau), \quad (26)$$

where $g(\theta) = \left[\frac{\phi'_{\rho}(V_A(\theta))}{\gamma} + \frac{\pi'(\theta)}{\gamma f(\theta)} \right] u'_x(x_A, z_A; \theta)$ is the social marginal weight of the θ -individuals within the population. Higher social priority is thus given to the people threatening to emigrate. Since the optimal marginal tax rate at θ is inversely related to the aggregate social marginal weights of the individuals with greater productivity, we expect individual mobility to decrease the optimal marginal tax rates over a productivity range exceeding that where the participation constraints are binding. We now turn to the different channels captured in Formula (24) to look into this intuition.

As previously, 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 $B_2(\theta)$ and $B_3(\theta)$.

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

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

$$\frac{\pi'(\tau) u'_x(x_A, z_A; \tau)}{\gamma} \times dTdz. \quad (27)$$

The substitution effect combines with the usual *income effect*. Since leisure is a normal good under Assumption 2, the increase in the tax burden paid by all individuals with income greater than z induces them to work more. This allows A 's government to increase the taxes they face. As a result, it is not required to compensate the potentially mobile individuals as high as the increase in taxes they face. We know from Saez (2001) that the magnitude of the uncompensated behavioural response is summarized by $\Psi_{\theta\tau} \geq 1$, which converts the social marginal utility of consumption of the τ -individuals, $u'_x(x_A, z_A; \tau)$, into that of the θ_z -individuals, $u'_x(x_A, z_A; \theta_z)$. Using (27), the social cost of the compensation of the τ -individuals, including income effects, is

$$\frac{\pi'(\tau) u'_x(x_A, z_A; \tau)}{\gamma} \Psi_{\theta\tau} \times dTdz. \quad (28)$$

Now, the overall social cost of compensating the individuals on the upper bound of the population is directly obtained as

$$-\frac{\partial W_{A,\rho}^N / \partial (V_B - c)|_{\bar{\theta}}}{\gamma} \times dTdz. \quad (29)$$

When the participation constraints are active at $\bar{\theta}$, $\partial W_{A,\rho}^N / \partial (V_B - c)|_{\bar{\theta}}$ is equal to $\partial W_{A,\rho}^N / \partial V_A|_{\bar{\theta}}$, which is $-\iota(\bar{\theta})$. Converting (29) into social marginal utility of consumption at θ_z , one gets

$$\frac{\iota(\bar{\theta}) u'_x(x_A, z_A; \theta_z)}{\gamma} \times dT dz. \quad (30)$$

Finally, by (28) and (30), the average social cost of the compensation of all potentially mobile individuals with gross income above z is

$$\begin{aligned} \frac{1}{1 - F(\theta_z)} \left[\int_{\theta_z}^{\bar{\theta}} \frac{\pi'(\tau) u'_x(x_A, z_A; \tau)}{\gamma} \Psi_{\theta_z \tau} d\tau + \frac{\iota(\bar{\theta}) u'_x(x_A, z_A; \theta_z)}{\gamma} \right] \times dT dz \\ = [B_2(\theta_z) + B_3(\theta_z)] \times dT dz. \end{aligned} \quad (31)$$

$B_2(\theta_z)$ is positive as soon as there are individuals with productivity above θ_z for whom the participation constraints are binding. This term goes therefore against progressivity on a range of gross income levels preceding that on which individuals hesitate to leave the country. This is because increasing the marginal tax rates at θ makes the compensation of all more productive individuals threatening to emigrate more expensive in terms of social welfare. In addition, for all $\theta < \theta^* : \pi'(\theta) = 0$, so

$$B_2(\theta) = \frac{1}{1 - F(\theta)} \int_{\theta^*}^{\bar{\theta}} \frac{\pi'(\tau) u'_x(x_A, z_A; \tau)}{\gamma} \Psi_{\theta^* \tau} d\tau. \quad (32)$$

Differentiating, one obtains

$$B'_2(\theta) = \frac{1 - F(\theta^*)}{(1 - F(\theta))^2} f(\theta) B_2(\theta^*), \quad (33)$$

which is strictly positive for $\theta \leq \theta^*$: the closer to θ^* the productivity level at which the small tax reform perturbation takes place, the higher the average compensation required to satisfy the participation constraints. When θ is greater than θ^* , it is not possible to determine the sign of $B_2(\theta)$ in the general case. Since $B_3(\theta)$ is non-negative, it reinforces the decrease in marginal tax rates induced by $B_2(\theta)$.

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

4.4. Citizen and Resident Criteria

Under the National criterion, the whole population is constrained to stay in A . We now relax this constraint to examine whether keeping everybody in the country is not too expensive in terms

of social welfare. For this purpose, we separate Problem 3 into two subproblems to determine the optimal $\hat{\theta}$. In the first subproblem, $\hat{\theta}$ is arbitrarily chosen by A 's government.

Subproblem 1. *Given $\hat{\theta} \in [\underline{\theta}, \bar{\theta}]$, find (x_A, z_A) to maximize $W_{A,\rho}^i(\hat{\theta})$, $i = \{C, R\}$, subject to (FOIC), (SOIC'), (PC), (TR).*

Let $\mathcal{W}_{A,\rho}^i(\hat{\theta})$ be the social value function of this subproblem, $\iota_{\hat{\theta}}^i(\theta)$ the shadow price of (FOIC), and $\pi_{\hat{\theta}}^i(\theta)$ the shadow price of a uniform marginal increase in the reservation utility for all $\theta' \geq \theta$. The solution in $\hat{\theta}$ to Problem 3 is then obtained as:

Subproblem 2. *Find $\hat{\theta}^i \in [\underline{\theta}, \bar{\theta}]$ solution to $\max_{\hat{\theta} \in [\underline{\theta}, \bar{\theta}]} \mathcal{W}_{A,\rho}^i(\hat{\theta})$, $i = \{C, R\}$.*

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

Property 6. *Proposition 4 applies for the Citizen and Resident criteria provided:*

- (i) $\bar{\theta}$ is replaced by $\hat{\theta}^i$ and $1 - F(\theta)$ by $F(\hat{\theta}^i) - F(\theta)$, $i = \{C, R\}$;
- (ii) in $B_1(\theta)$, $\phi'_\rho(V_A)$ is divided by $F(\hat{\theta}^R)$ for the Resident criterion.

Proof. See A.3 in the Appendix. □

In addition, we note that for all $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$, the $\hat{\theta}$ -individuals are indifferent between living in A or B . By Property 1, they thus pay strictly positive taxes at the solution to Subproblem 1.

We are now prepared to examine the allocation of individuals between A and B resulting from the implementation of the Citizen and Resident optimal income tax schedules. Let us assume $\hat{\theta} < \bar{\theta}$. Individuals with productivity above $\hat{\theta}$ are in B . Making them relocate to A has four effects. (i) It increases social welfare because they have utility levels above the average in A 's resident population and $\phi' > 0$. (ii) It increases tax receipts and thus social welfare because strictly positive taxes can be levied on them since $c > 0$. (iii) It requires adjustments to prevent them from imitating less productive individuals. (iv) It brings about a *new upward mimicking behaviour*. Indeed, A 's residents can now have an incentive to mimicking them since they have the most appealing outside options.

The last effect is crucial to understanding the interactions between the incentive-compatibility conditions and the type-dependent participation constraints. In closed economy, individuals have the usual incentive to understate θ to obtain greater social benefit whilst enjoying more leisure¹². When type-dependent participation constraints are taken into account, the individuals may also be tempted to overstate θ in working harder to obtain greater compensation for staying in A . This behaviour reflects *countervailing incentives*. An asymmetry in terms of informational

¹²In the discrete population model of Guesnerie and Seade (1982), a sufficient condition for incentive-compatibility of the tax scheme is that the downward self-selection constraints are binding, which corresponds to a monotonic chain to the left (see also Weymark (1986, 1987)).

constraints between the individuals with productivity below $\hat{\theta}$ and the $\hat{\theta}$ -individuals may therefore arise. Indeed, contrary to the former, the latter can only have the usual incentives. The cost of making the θ -individuals reveal their private information, represented by $\iota_{\theta}^i(\theta) \geq 0$, can thus have a downward jump discontinuity at $\hat{\theta}$. However, making them reveal their private knowledge requires the gap between $V_B(\hat{\theta}) - c(\hat{\theta})$ and $V_A(\theta)$ to be reduced. This increase in $V_A(\theta)$ reduces the social cost of a uniform increase in the reservation utility at θ and above, which is captured by $\pi(\theta)$. This effect stops suddenly when θ tends to $\hat{\theta}$. Consequently, an *upward* jump discontinuity in π_{θ}^i corresponds to the *downward* jump discontinuity in ι_{θ}^i at $\hat{\theta}$. It turns out that these discontinuities have the same magnitude.

Property 7. *At $\hat{\theta}$,*

$$\iota_{\hat{\theta}}^i(\hat{\theta}^-) - \iota_{\hat{\theta}}^i(\hat{\theta}) = \pi_{\hat{\theta}}^i(\hat{\theta}) - \pi_{\hat{\theta}}^i(\hat{\theta}^-) \geq 0 \quad (= 0 \text{ if } (PC) \text{ inactive at } \hat{\theta}), \quad (34)$$

where $\iota_{\hat{\theta}}^i(\hat{\theta}^-) := \lim_{\theta \rightarrow \hat{\theta}^-} \iota_{\theta}^i(\theta)$.

Proof. See (77) in the Appendix. □

We first provide two sufficient conditions for such a discontinuity not to occur.

Property 8. *If $\iota_{\hat{\theta}}^i$ is continuous at $\hat{\theta}$ for all $\hat{\theta} \in (\underline{\theta}, \bar{\theta}]$ and if the $\bar{\theta}$ -individuals pay strictly positive taxes at the second-best National optimum, then $\hat{\theta}^i = \bar{\theta}$, $i = \{C, R\}$.¹³*

Proof. See A.3 in the Appendix. □

Property 9. *If there is a non-empty interval containing $\hat{\theta}$ where (PC) is active at the solution to Subproblem 1 for all $\hat{\theta} \in (\underline{\theta}, \bar{\theta}]$, then $\hat{\theta}^i = \bar{\theta}$, $i = \{C, R\}$.*

Proof. See A.3 in the Appendix. □

These properties hold when the usual downward mimicking behaviour predominates for highly skilled individuals. However, they do not exhaust all possible cases. Indeed, the trade-off between maintaining national capacity to the maximum and sustaining the redistribution programme is more complex when countervailing incentives prevail close to the top. In particular, there are cases where it is optimal to implement an income tax schedule inducing emigration of its highly skilled. A careful inspection of the National solution provides a sufficient condition.

Proposition 5. *Consider the National optimal allocation and the corresponding $(\gamma, \iota, \pi, T, V_A, W_{A,\rho}^N)$.*

Assume $\bar{\theta}$ is an isolated point where (PC) is active and ι has a jump discontinuity. Then

(i) $\hat{\theta}^C < \bar{\theta}$ if

$$\gamma T(z_A) f < \left[\iota(\bar{\theta}^-) - \iota \right] |R'| - \iota(\bar{\theta}^-) V_A'; \quad (35)$$

¹³ $T(z_A(\bar{\theta})) > 0$ cannot be established in the general case. Besides, let $\rho = 0$ (Utilitarianism) and consider quasilinear-in-consumption preferences. Then, the optimal second-best income tax policy in closed economy is the laissez-faire and so $V_A^{cl}(\bar{\theta}) = V_B(\bar{\theta})$. Since $c(\bar{\theta}) > 0$, the National second-best optimal allocation is also the laissez-faire and thus $T(z_A(\bar{\theta})) = 0$.

(ii) $\hat{\theta}^R < \bar{\theta}$ if

$$\gamma T(z_A) f + [\phi_\rho(V_B - c) - W_{A,\rho}^N] f < [\iota(\bar{\theta}^-) - \iota] |R'| - \iota(\bar{\theta}^-) V'_A, \quad (36)$$

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

Proof. See A.3 in the Appendix. \square

Here, it should be kept in mind that the individuals to the very left of $\bar{\theta}$ have utility above their reservation utility because $\bar{\theta}$ is an isolated point where (PC) is active. Since (PC) is active at $\bar{\theta}$, this implies that $V_B - c$ is steeper than V_A at this point, i.e. $R'(\bar{\theta}) < 0$. Individuals to the very left of $\bar{\theta}$ are therefore left with a location rent at the margin.

(35) and (36) reflect the marginal cost-benefit analysis of the presence in A of the most productive individuals. By property 1, the $\bar{\theta}$ -individuals pay positive taxes $T(z_A(\bar{\theta}))$. Since their proportion is represented by $f(\bar{\theta})$ and γ is the unit of count in welfare, these taxes increase social welfare by $\gamma T(z_A(\bar{\theta})) f(\bar{\theta})$. In addition, under the Resident criterion, the $\bar{\theta}$ -individuals have utility above the average in A , which rises social welfare by $[\phi_\rho(V_B(\bar{\theta}) - c(\bar{\theta})) - W_{A,\rho}^N(\bar{\theta})] f(\bar{\theta})$. These positive marginal effects appear on the LHSs.

The RHSs capture the marginal costs and benefits with regard to incentives of the presence in A of the $\bar{\theta}$ -individuals. First, there is a cost due to countervailing incentives. Indeed, individuals to the very left of $\bar{\theta}$ have the possibility to mimic the $\bar{\theta}$ -individuals to benefit from their higher outside options. They can therefore claim an increase in their utility at the margin, equal to $V'_B(\bar{\theta}) - c'(\bar{\theta}) - V'_A(\bar{\theta}) = |R'(\bar{\theta})|$. The shadow price of this behaviour is given by the excess of $\iota(\bar{\theta}^-)$ over $\iota(\bar{\theta})$, which is positive by Property 7. The corresponding marginal social cost is thus

$$[\iota(\bar{\theta}^-) - \iota(\bar{\theta})] |R'(\bar{\theta})| > 0. \quad (37)$$

(37) would vanish if $\bar{\theta}$ were not a non-isolated point where (PC) is active since, in this case, $R'(\bar{\theta}) = 0$.

Hence, because of countervailing incentives, the individuals to the very left of $\bar{\theta}$ have greater utility. They thus are less inclined to mimic less productive individuals. The slope of the indirect utility V'_A at $\bar{\theta}$ required for them to reveal their type truthfully is therefore reduced at the margin. Since $\iota(\bar{\theta}^-)$ is the shadow price of (FOIC), the marginal social benefit of this slackening of the downward incentive compatibility constraints is $\iota(\bar{\theta}^-) V'_A$. Here appears the second impact of countervailing incentives due to the presence in A of the most productive individuals. Finally, the net marginal social cost of countervailing incentives amounts to

$$[\iota(\bar{\theta}^-) - \iota(\bar{\theta})] |R'(\bar{\theta})| - \iota(\bar{\theta}^-) V'_A(\bar{\theta}). \quad (38)$$

The choice of $\hat{\theta}$ by A 's government can thus be regarded as a means of revealing private information. Indeed, if A 's government designs a tax policy such that the individuals with

productivity greater than $\hat{\theta}$ do not receive in A their reservation utility, it *knows* that $\hat{\theta}$ is the maximum productivity in its resident population and consequently that the individuals with productivity greater than $\hat{\theta}$ are in B . Proposition 5 tells us in which cases using this means improves social welfare.

5. NUMERICAL RESULTS

We already know that individual mobility is harmful to progressivity and significantly alters the qualitative properties of the optimal non-linear income tax schedule. It remains to *quantify* the magnitude of the changes with respect to the Mirrleesian closed economy model. In particular, we would like to examine whether potential mobility of a few highly skilled individuals has more than a negligible effect on the optimal policy. For this purpose, we adopt the National criterion and calibrate A 's economy to roughly correspond to the French one.

5.1. Calibration

We employ a truncated lognormal distribution for the lower part of the productivity distribution and complete it with a Pareto tail with density $f(\theta) = K/\theta^{1+a}$. The two parameters of the lognormal distribution are estimated on the French survey data "Budget des familles", year 1995, as portrayed in Figure 4 in Laslier et al. (2003). We get a mean of 0.2398 and a variance of 0.4403¹⁴. Following Hungerbühler et al. (2005), we take $a = 2$, choose K and the boundary between both distributions in such a way that the entire distribution is continuously differentiable. We normalize the productivity levels so that the median individuals have productivity equal to the median income in 1995, i.e. 13 320 euros. The productivity support is the positive real line with an upper bound equal to 15 times the median productivity.

We focus on the case where there are no income effects on labour supply and e^H is constant, with $e^H = 0.2$, as in d'Autume (2000). Preferences are thus given by (15) and (16). For convenience, the government is assumed to be Rawlsian.

Migration costs are the new ingredient of our model. They correspond to all the costs an individual will have to pay because of his choice of migration. Since the model is static, these costs as well as the utility levels should be regarded as expected values. Very few empirical work have studied the individual costs of migration¹⁵. We use constant costs as a benchmark and calibrate them so as to reflect plausible scenarios as regards the proportion of individuals threatening to emigrate: 10%, 5%, 3%, 1%, 0.5% and 0.1%. We obtain migration costs equal to 15 550, 27 900, 40 500, 77 400, 104 300 and 151 900 euros per annum respectively. We also provide simulations for linearly decreasing costs. The highest skilled individuals have migration costs equal to 40 500 and 77 400 euros, values used in the constant case. The costs are then

¹⁴Since our model does not take the family size into account, the population is restricted to single individuals.

¹⁵For instance, the IZA Database for Migration Literature provides 34 matches for "moving costs" (<http://www.iza.org/iza/en/webcontent/links/migration>). These references are mainly theoretical or estimate the macroeconomic costs of migration.

linearly adjusted to obtain threat of migration by 1% and 0.5% respectively.

5.2. Numerical Results

Figures 2–4 and Table 1 in Appendix A.4 contrast the second-best optimal allocations for constant migration costs in the six scenarios described above. For instance, when 3% of the population threaten to emigrate, the social welfare, equal to the redistributive budget under the maximin, is reduced by 5.8%. The individuals paying the maximum average tax have gross income $\hat{z}_A = 70\,834\text{€}/\text{year}$. The optimal average tax rates are decreasing above this level, even if the participation constraints are only active for individuals with gross income above $z_A^* = 98\,779\text{€}/\text{year}$. The range of decrease corresponds to 4.4% of the population. Even if the individual and social utility levels do only slightly vary compared to the benchmark, Figure 2 emphasizes that the changes in the tax schedule are very noticeable, even when the proportion of potentially mobile individuals is very low.

Specifically, even if the average tax rate profile is already single-peaked in closed-economy, the corresponding graphs are far more hump-shaped when the threat of migration goes up (Figure 2.b). The lower bound \hat{z}_A of the range of gross income from which the average tax rate is decreasing (cf. the black circles) is smaller than the gross income z_A^* from which the participation constraints are active (cf. the squares). Interestingly, the smaller the proportion of the population threatening to emigrate, the larger the gap between \hat{z}_A and z_A^* as well as the ratio

$$1 + \frac{1 - F(\theta_{z_A^*})}{1 - F(\theta_{\hat{z}_A})}. \quad (39)$$

The latter is approximately equal to 1.26, 1.34, 1.47, 2, 2.6, 7 for potential emigration by 10%, 5%, 3%, 1%, 0.5%, and 0.1% of the population. In this respect, the threat of migration seems to have a multiplicative power all the stronger as fewer people would like to emigrate. It is really a feature that only simulations may reveal.

Figure 3 and 4 contrast the open and closed optimal allocations from a distributional point of view¹⁶. The highly skilled appear as the real winners, since they pay less taxes (Figure 3.a) and have higher utility (Figure 3.c). The situation of the low-skilled does not worsen as much as one could expect. Individuals with gross income close to \hat{z}_A are actually the real losers in terms of taxes. Nevertheless, they slightly benefit from the openness of the economy in terms of utility. In fact, the decline in marginal tax rates allow them to increase their gross and net income sufficiently to overbalance the resulting loss in leisure (Figure 3.b). Consequently, the deterioration of the middle-skilled workers' situation in terms of taxes does not translate into losses in individual welfare as observed in the first-best. In spite of this rather comforting result, inequality of utilities deepens (Figure 4).

¹⁶The gross income received by the θ -individuals depends on the tax schedule itself. For instance, we have represented the choice of the θ^* -individuals if the economy were closed by empty circles in Figure 2a. As a matter of fact, the variation in taxes in Figure 3.a is not obtained as the difference between taxes paid with and without mobility for a given *gross income* that can be read in Figure 2.c.

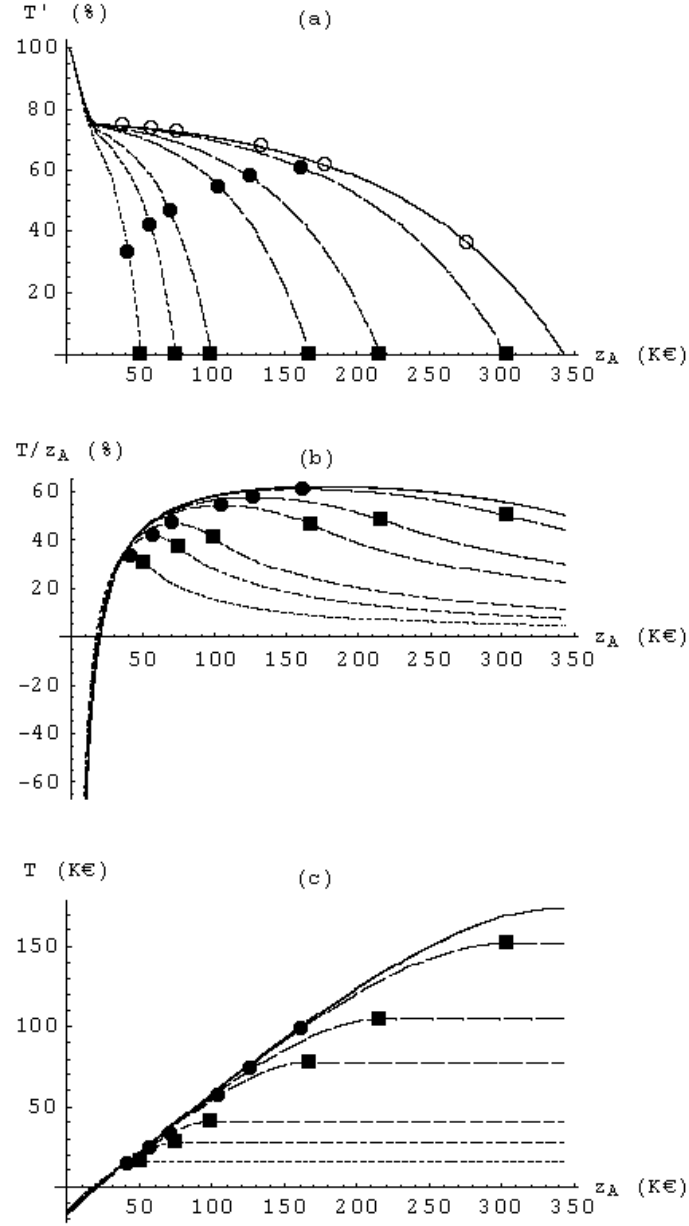


Figure 2: Constant Migration Costs. (a) Optimal Marginal Tax Rates; (b) Optimal Average Tax Rates; (c) Optimal Taxes. The solid line refers to the closed economy benchmark. Otherwise, the less dotted the line, the lower the threat of migration: 10%; 5%; 3%; 1%; 0.5% and 0.1% respectively. Squares correspond to \hat{z}_A ; black circles to z_A^* ; empty circles to the choice the θ^* -individuals would make if the economy were closed.

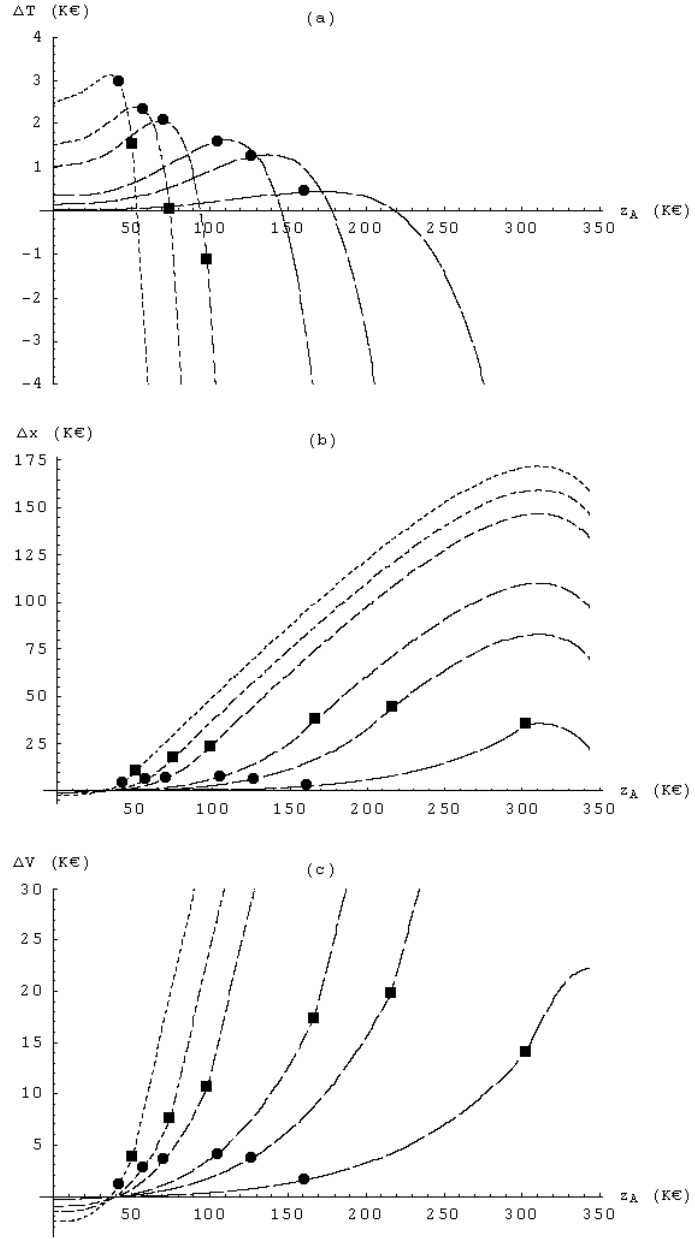


Figure 3: Constant Migration Costs. Increase w.r.t closed economy benchmark: (a) Taxes; (b) Consumption; (c) Utility. The less dotted the line, the lower the threat of migration: 10%; 5%; 3%; 1%; 0.5%; 0.1% respectively. Squares correspond to \hat{z}_A ; circles to z_A^* .

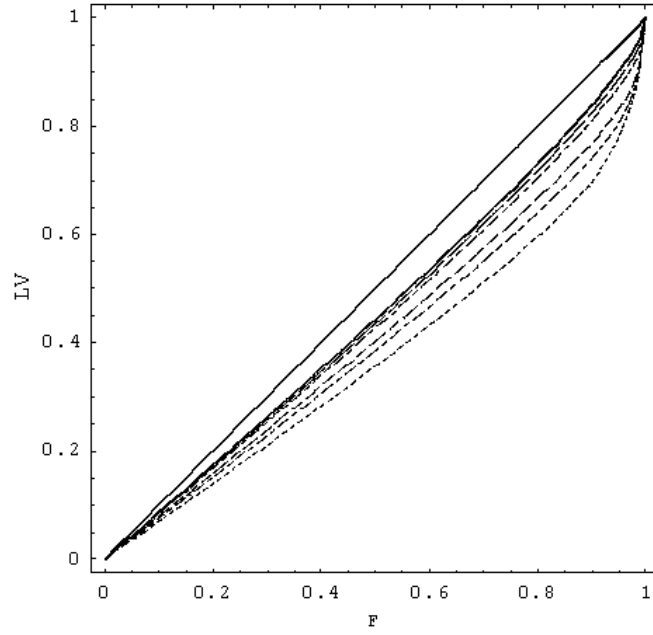


Figure 4: Constant Migration Costs. Lorenz Curves for the Indirect Utility Levels. The less dotted the line, the lower the threat of migration: 10%; 5%; 3%; 1%; 0.5%, 0.1% respectively. The solid line below the 45° -line pertains to the closed economy benchmark.

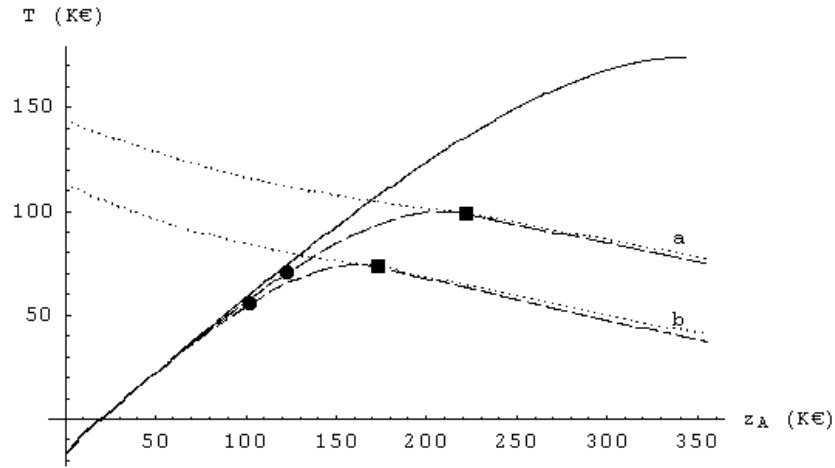


Figure 5: Decreasing Migration Costs. The dotted lines give the migration costs; the dashed lines the optimal tax liabilities. The less dashed the latter, the greater the threat of migration: (a) 0.5%; (b) 1%. Squares correspond to \hat{z}_A ; circles to z_A^* . The solid line pertains to the closed economy benchmark.

Figure 5 provides examples of optimal tax schemes for decreasing migration costs¹⁷. The tax liabilities are hump-shaped so that the middle-skilled have actually to pay greater taxes than the highly-skilled. In contrast to the case where $c(\theta)$ is constant, people on their participation constraints pay lower taxes than their migration costs. In the first-best setting, they would be taxed as high as their migration costs. Therefore, taking incentive compatibility into account restricts the tax levy on the potentially mobile individuals.

6. CONCLUSION

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

In this extended framework, the issue of the optimal allocation of individuals between the home country and abroad is embedded in the optimal income tax problem. Consequently, a new trade-off between maintaining the redistribution programme and preserving national productive capacities adds to the traditional trade-off between equity and efficiency. In the first-best, emigration of highly skilled individuals should always be prevented to maximize social welfare. In the second-best, this may be false because of countervailing incentives.

Key qualitative features of the optimal income tax policy obtained in closed economy do no longer hold. The participation effect does not only favour a decrease in the optimal marginal tax rates; it can also make them strictly negative. Consequently, the optimal average tax rates as well as the optimal tax liabilities can be decreasing. Numerical simulations show that the threat of migration has a significant impact even when the proportion of potentially mobile individuals is very low. They also reveal that if the highly skilled are the real winners, the welfare of the low-skilled is only slightly reduced because quite high taxes can still be levied on the middle-skilled.

Our qualitative and quantitative results convey a curse of the middle-skilled workers: this curse is expressed in terms of utility in the first-best, and in terms of average tax rates and tax liabilities in the second-best.

The backbone of the analysis remains valid even when B 's government implements a non-linear income tax policy, provided this policy is given. Consequently, the material of this paper paves the way for deriving the reaction function of A to B 's tax policy and vice versa. It thus provides the basic ingredients of a symmetric game on redistributive non-linear income taxes, the solution of which is left for further research.

¹⁷Linearity in θ does not imply linearity in z_A , as shown by the dotted lines.

A. APPENDIX

A.1. First-Best

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

$$(\phi'_\rho + \pi') U'_x = \gamma \text{ and } (\phi'_\rho + \pi') U'_l = -\gamma\theta, \quad (40)$$

with

$$\pi' \geq 0, \quad U(x, l) - V_B + c \geq 0, \quad \pi' [U(x, l) - V_B + c] = 0, \quad \forall \theta \in [\underline{\theta}, \bar{\theta}]. \quad (41)$$

Since $\phi'_\rho > 0$, (40) and (41) implies $\gamma > 0$. The following Lemma is admitted since its proof parallels that given in Mirrlees (1974).

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

Step 1: The existence of θ^* is obvious. Indeed, since $V_A^{cl}(\bar{\theta}) < V_B(\bar{\theta}) - c(\bar{\theta})$, the closed-economy solution violates (6); so there are θ such that $\pi' > 0$ at the solution to Problem 1.

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

By (40), $\pi'(\theta) = \gamma/U'_x - \phi'_\rho$, which implies under Assumption 1 and the continuity of T , the continuity of π' . Assume $\theta' := \min\{\theta \in [\theta^*, \bar{\theta}] : \pi'(\theta) = 0\}$ exists. Then, by continuity of π' , there exists $\theta'' > \theta'$ such that $\pi' = 0$ on $[\theta', \theta'']$. By continuity of R , $R(\theta') = 0$. On $[\theta', \theta'']$, $V'_A \leq 0$ by Lemma 1 and $V'_B - c' > 0$ under Assumption 3. Then $R < 0$ for $\theta \in (\theta', \theta'')$, contradicting (PC). Hence, θ' does not exist. \square

Proof of Proposition 2. We denote by $V_A(\theta; \rho)$ and $\pi'(\theta; \rho)$ the values of $V_A(\theta)$ and $\pi'(\theta)$ for a given ρ .

(a) *Citizen criterion.* We proceed by contradiction. Assume $\hat{\theta} < \bar{\theta}$ is optimal. It is possible to give the θ -individuals, with $\theta > \hat{\theta}$, their laissez-faire utility V_B in A . Since $c > 0$ and $\phi'_\rho > 0$, we have $\phi_\rho(V_B - c) < \phi_\rho(V_B)$ and thus

$$\int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A(\cdot; \rho)) dF + \int_{\hat{\theta}}^{\bar{\theta}} \phi_\rho(V_B) dF > \int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A(\cdot; \rho)) dF + \int_{\hat{\theta}}^{\bar{\theta}} \phi_\rho(V_B - c) dF, \quad (42)$$

the RHS of which is $W_{A,\rho}^C(\hat{\theta})$. But the allocation corresponding to the LHS of (42) is also feasible. A contradiction.

(b) *Resident criterion.* We proceed in two steps.

Step 1: $\phi_\rho(V_B(\hat{\theta})) > \phi_\rho(V_A(\underline{\theta}; \rho))$, $\forall \rho$.

Assume $V_B(\hat{\theta}) \leq V_A(\underline{\theta}; \infty)$. Then, $V_B(\theta) \leq V_A(\underline{\theta}; \infty)$, $\forall \theta \leq \hat{\theta}$, because $V'_B > 0$. As $c > 0$ and (41), $\pi'(\theta; \infty) = 0$, $\forall \theta \leq \hat{\theta}$. Hence, by Lemma 1, $V_A(\underline{\theta}; \infty) = V_A(\theta; \infty)$, $\forall \theta \leq \hat{\theta}$. One

thus gets $V_B(\theta) < V_A(\underline{\theta}; \infty)$, $\forall \theta < \hat{\theta}$, which contradicts the Pareto-efficiency of the laissez-faire. Consequently,

$$V_B(\hat{\theta}) > V_A(\underline{\theta}; \infty). \quad (43)$$

In addition, A 's government maximizes $V_A(\underline{\theta}; \rho)$ when $\rho \rightarrow \infty$. Hence,

$$V_A(\underline{\theta}; \infty) \geq V_A(\underline{\theta}; \rho), \quad \forall \rho. \quad (44)$$

By (43) and (44), $V_B(\hat{\theta}) > V_A(\underline{\theta}; \rho)$, $\forall \rho$, which completes the step since $\phi'_\rho > 0$.

Step 2: $\hat{\theta} = \bar{\theta}$ at the optimum.

We proceed by contradiction. Assume $\hat{\theta} < \bar{\theta}$ is optimal. Proposition 1 applies for the given $\hat{\theta}$. Hence, $\phi_\rho(V_A(\underline{\theta}; \rho)) \geq \phi_\rho(V_A(\theta; \rho))$ for all $\underline{\theta} \leq \theta < \theta^*$. By Step 1,

$$\phi_\rho(V_B(\hat{\theta})) \geq \phi_\rho(V_A(\rho; \theta)), \quad \forall \underline{\theta} \leq \theta < \theta^*. \quad (45)$$

In addition, since $\phi'_\rho > 0$ and $c > 0$,

$$\phi_\rho(V_B(\hat{\theta})) > \phi_\rho(V_B(\theta) - c(\theta)), \quad \forall \theta^* < \theta \leq \hat{\theta}. \quad (46)$$

Using (45), (46), and $V_A = V_B - c$, $\forall \theta > \theta^*$, we deduce:

$$\frac{1}{F(\hat{\theta})} \int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A(\rho; \theta)) dF(\theta) < \phi_\rho(V_B(\hat{\theta})), \quad (47)$$

the LHS of which is $W_{A,\rho}^R(\hat{\theta})$. It is always feasible to give the θ -individuals with $\theta > \hat{\theta}$ the utility $V_B(\theta) > V_B(\hat{\theta})$. A contradiction. \square

A.2. Second-Best: National Criterion

Proof of Proposition 3. Since $V_A = \theta l_A - T - v(l_A)$, by 15, the government chooses l_A and V_A to maximize

$$\int_{\underline{\theta}}^{\bar{\theta}} [\theta l_A(\theta) - V_A(\theta) - v(l_A(\theta))] dF(\theta) \quad \text{s.t. (FOIC) and (PC)}.$$

l_A is control variable; V_A state variable with adjoint variable ι . The Hamiltonian and Lagrangian are

$$\begin{aligned} H^N &= \left[\theta l_A - V_A - v(l_A) + \iota \frac{l_A}{\theta} v'(l_A) \right] f, \\ L^N &= H^N + \pi' R. \end{aligned} \quad (48)$$

By Theorem 2 in Seierstad and Sydsaeter (1987, p. 332-335)^{18,19}, necessary conditions are:

$$\partial H^N / \partial l_A = 0 \Leftrightarrow (\theta - v') f + \iota \left(\frac{v'}{\theta} + \frac{l_A}{\theta} v'' \right) = 0, \quad (49)$$

$$\partial L^N / \partial V_A = -\iota'(\theta) \Leftrightarrow \iota'(\theta) = f - \pi', \quad (50)$$

$$\iota(\bar{\theta}) \geq 0 \quad (= 0 \text{ when } R(\bar{\theta}) > 0), \quad (51)$$

$$\iota(\underline{\theta}) \leq 0 \quad (= 0 \text{ when } R(\underline{\theta}) > 0), \quad (52)$$

$$\pi'(\theta) \geq 0, \quad R(\theta) \geq 0, \quad \pi'(\theta) R(\theta) = 0, \quad (53)$$

$$\iota(\theta_j^-) - \iota(\theta_j^+) = \pi(\theta_j^+) - \pi(\theta_j^-) \geq 0 \quad (= 0 \text{ if } R(\theta_j) > 0). \quad (54)$$

Since $v'/\theta = 1 - T'$ and $e^H(\theta) = v'/(l_A v'')$, rearranging (49) yields

$$\frac{T'}{1 - T'} = -\frac{\iota}{\theta f} \left(1 + \frac{1}{e^H(\theta)} \right). \quad (55)$$

Integration of (50) between θ and $\bar{\theta}$ gives

$$\iota(\theta) = \iota(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} \iota'(\theta) d\theta = \iota(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} [f - \pi'] d\theta = \iota(\bar{\theta}) - 1 + F(\theta) + \int_{\theta}^{\bar{\theta}} \pi' d\theta, \quad (56)$$

which is plugged into (55). The proof is completed in two steps.

Step 1: $\int_{\theta}^{\bar{\theta}} \pi' d\theta = \int_{\theta \in [\theta, \bar{\theta}] \cap \Theta^{PC}} f(\theta) d\theta = \mu_f([\theta, \bar{\theta}] \cap \Theta^{PC})$.

Consider any non-empty open interval in $[\underline{\theta}, \bar{\theta}]$ where $\pi' > 0$. By Property 5, $T' = 0$. Hence, for all θ in this interval $\iota(\theta) = 0$ by (55) and thus $\iota'(\theta) = 0$; so by (50), $\pi'(\theta) = f(\theta)$. Since (PC) is assumed not to be active at isolated points, the equality above is obtained.

Step 2: $\iota(\bar{\theta}) = 0$.

If (PC) is inactive at $\bar{\theta}$, $\iota(\bar{\theta}) = 0$ because of (51). Otherwise, (PC) is active at $\bar{\theta}$ and, under assumptions, there is $\varepsilon > 0$ such that (PC) is active on $I = (\bar{\theta} - \varepsilon, \bar{\theta}]$. Since e^H is constant, by Property 5, $T' = 0$ on I . Hence, $\iota(\bar{\theta}) = 0$ by (55). \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 (TR) is taken into account through G' and the boundary conditions $G(\underline{\theta}) = 0$ and $G(\bar{\theta}) = 0$. It is not necessary to take x_A explicitly into account because it is uniquely determined by V_A and z_A . Denote $x_A = h(V_A, z_A; \theta)$; differentiating shows $\partial x_A / \partial V_A = 1/u'_x$ and

¹⁸The Theorem we referred to is applied as follows. Since the adjoint variables are assumed to have a finite number of jump discontinuities and be C^1 elsewhere, the "almost necessary conditions" (p. 335) are in fact necessary. In the Theorem, $q' = \lambda$ is our π' . Hence, q is our π and λ our π' . Consequently, their $\beta^k = \pi(\tau_k^+) - \pi(\tau_k^-)$. We then employ their Eq. (5.37) to get (54).

¹⁹The necessary conditions are often stated for state variables which are fixed at the initial point, which is not the case presently. We have used Seierstad and Sydsaeter (1987, Theorem 5, pp. 185, Eq. 30b) (referred to as (S-S) from now on) to obtain (52). This remark applies to the other proofs in the paper.

$\partial x_A / \partial z_A = s$. The Hamiltonian and Lagrangian are respectively

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

We call $\{\theta_j\}_{j=1}^N$ points where (6) becomes or ceases to be active as well as $\underline{\theta}$ and $\bar{\theta}$. As $\partial u'_\theta / \partial z_A = u''_{\theta z} + s u''_{\theta x} = -u'_x s'_\theta$, and $\partial u'_\theta / \partial V_A = u''_{\theta x} / u'_x$, necessary conditions are:

$$\partial H^N / \partial z_A = 0 \Leftrightarrow \iota u'_x s'_\theta - \gamma(1 - s) f = 0, \quad (57)$$

$$\partial L^N / \partial V_A = -\iota' \Leftrightarrow \iota'(\theta) = -\phi'_\rho(V_A) f - \iota u''_{\theta x} / u'_x - \pi' + \gamma f / u'_x, \quad (58)$$

$$\partial L^N / \partial G = -\gamma' \Leftrightarrow \gamma' = 0, \quad (59)$$

$$(51), (52), (53), (54). \quad (60)$$

$\gamma(\theta)$ is constant, equal to $\gamma > 0$. As $s = 1 - T'$, $T' = \iota u'_x s'_\theta / (\gamma f)$ by (57). In addition, using basic calculus, $[1 + e^M(\theta)] / e^H(\theta) = -\theta s'_\theta / s$. Hence,

$$\frac{T'}{1 - T'} = -\frac{\iota u'_x}{\gamma \theta f} \frac{1 + e^M(\theta)}{e^H(\theta)}. \quad (61)$$

When $\theta = \bar{\theta}$, (61) and (51) yield (25). When $\theta < \bar{\theta}$, (61) can be rewritten as

$$\frac{T'}{1 - T'} = -\frac{\iota u'_x}{\gamma(1 - F(\theta))} \frac{1 + e^M(\theta)}{e^H(\theta)} \frac{1 - F(\theta)}{\theta f(\theta)}, \quad (62)$$

If $(.; \tau)$ means evaluation at $(x_A(\tau), z_A(\tau); \tau)$, integrating (58) between θ and $\bar{\theta}$ yields

$$\iota(\theta) = \iota(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} \left(\phi'_\rho(V_A(\tau)) f(\tau) + \pi'(\tau) - \frac{\gamma f(\tau)}{u'_x(., \tau)} \right) \tilde{\Psi}_{\theta\tau} d\tau, \quad (63)$$

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

$$\Psi_{\theta\tau} := \frac{u'_x(., \theta)}{u'_x(., \tau)} \tilde{\Psi}_{\theta\tau} = \exp \int_{\theta}^{\tau} \left(1 - \frac{e^M(\tau')}{e^H(\tau')} \right) \frac{z'_A(\tau')}{z_A(\tau')} d\tau'. \quad (64)$$

Using (63) and (64),

$$-\frac{\iota(\theta) u'_x(., \theta)}{\gamma} = \int_{\theta}^{\bar{\theta}} \left[1 - \left(\phi'_\rho(V_A(\tau)) + \frac{\pi'(\tau)}{f(\theta)} \right) \frac{u'_x(., \tau)}{\gamma} \right] \Psi_{\theta\tau} dF(\tau) - \frac{\iota(\bar{\theta}) u'_x(., \theta)}{\gamma}, \quad (65)$$

and plug the obtained expression in (62). \square

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

Proof of Property 6. (a) *Citizen criterion.* By definition, $W_{A,\rho}^C(\hat{\theta})$ is maximum when $\hat{\theta} = \hat{\theta}^C$, i.e. when $W_{A,\rho}^C(\hat{\theta}^C)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (PC), (TR). The FOC are the same as (57)–(60), except that $\bar{\theta}$ is replaced by $\hat{\theta}^C$. We then proceed as in the proof of Proposition 4.

(b) *Resident criterion.* By definition, $W_{A,\rho}^R(\hat{\theta})$ is maximum when $\hat{\theta} = \hat{\theta}^R$, i.e. when $W_{A,\rho}^R(\hat{\theta}^R)$ is maximized with respect to (x_A, z_A) subject to (FOIC), (PC), (TR). The FOC are the same as (57)–(60), except that (i) $\bar{\theta}$ is replaced by $\hat{\theta}^R$ and (ii) $\phi'_\rho(V_A)$ is divided by $F(\hat{\theta}^R)$. We then proceed as in the proof of Proposition 4. \square

We introduce the following definition and prove a lemma to establish Properties 8–9 and Proposition 5.

Definition 1. (i) Consider the Citizen criterion and an allocation solution to Subproblem 1. Then,

$$\vartheta^C(\hat{\theta}) = \gamma T(z_A) f + [\phi_\rho(V_A) - \phi_\rho(V_B - c)] f + \iota(\hat{\theta}^-) V'_A + [\iota(\hat{\theta}^-) - \iota] R'. \quad (66)$$

(ii) Consider the Resident criterion and an allocation solution to Subproblem 1. Then,

$$\vartheta^R(\hat{\theta}) = \gamma T(z_A) f + [\phi_\rho(V_A) - \mathcal{W}_{A,\rho}^R] f + \iota(\hat{\theta}^-) V'_A + [\iota(\hat{\theta}^-) - \iota] R', \quad (67)$$

In (66) and (67), all functions are evaluated at $\hat{\theta}$ except otherwise stated.

Lemma 2. For $i = \{C, R\}$, (i) $(\vartheta^i(\hat{\theta}) > 0, \forall \hat{\theta} > \underline{\theta}) \Rightarrow \hat{\theta}^i = \bar{\theta}$; (ii) $\vartheta^i(\bar{\theta}) < 0 \Rightarrow \hat{\theta}^i < \bar{\theta}$.

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

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

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

$$\partial L^i / \partial \zeta_A = 0 \Leftrightarrow \eta + \kappa g'(\zeta_A) = 0, \quad (68)$$

$$\eta' = -\partial L^i / \partial z_A \Leftrightarrow \eta' = \iota u'_x s'_\theta - \gamma(1-s)f, \quad (69)$$

$$\iota' = -\partial L^i / \partial V_A \Leftrightarrow \iota' = -\phi'_\rho f - \iota u''_{\theta x} / u'_x + \gamma f / u'_x - \pi', \quad (70)$$

$$\gamma' = -\partial L^i / \partial G \Leftrightarrow \gamma' = 0, \quad (71)$$

$$\pi' \geq 0, \quad R \geq 0, \quad \pi' R = 0, \quad (72)$$

$$\kappa \geq 0, \quad g(\zeta_A) \geq 0, \quad \kappa g(\zeta_A) = 0, \quad (73)$$

$$\eta(\underline{\theta}) = \eta(\widehat{\theta}) = 0, \quad (74)$$

$$\iota(\underline{\theta}) \leq 0 \quad (= 0 \text{ if } R(\underline{\theta}) > 0), \quad (75)$$

$$\iota(\widehat{\theta}) \geq 0 \quad (= 0 \text{ if } R(\widehat{\theta}) > 0), \quad (76)$$

$$\iota(\theta_j^-) - \iota(\theta_j^+) = \pi(\theta_j^+) - \pi(\theta_j^-) \geq (= 0 \text{ if } R(\theta_j) > 0). \quad (77)$$

η is continuous (see Eq. (75), p. 375, in S-S). We check that $\gamma > 0$. In addition, by continuity of η and (74),

$$\eta(\widehat{\theta}^-) \zeta_A(\widehat{\theta}) = \eta(\widehat{\theta}) \zeta_A(\widehat{\theta}) = 0. \quad (78)$$

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

$$\partial \mathcal{W}_{A,\rho}^C(\widehat{\theta}) / \partial \widehat{\theta} = \frac{\partial}{\partial \widehat{\theta}} \left[\int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A) dF(\theta) \right] - \phi_\rho(V_B(\widehat{\theta}) - c(\widehat{\theta})) f(\widehat{\theta}), \quad (79)$$

$$\partial \mathcal{W}_{A,\rho}^R(\widehat{\theta}) / \partial \widehat{\theta} = \frac{1}{F(\widehat{\theta})} \left[\frac{\partial}{\partial \widehat{\theta}} \int_{\underline{\theta}}^{\widehat{\theta}} \phi_\rho(V_A) dF(\theta) \right] - \frac{f(\widehat{\theta})}{F(\widehat{\theta})} \mathcal{W}_{A,\rho}^R(\widehat{\theta}). \quad (80)$$

Eq. (79), p. 376, in (S-S) gives the value of the square brackets on the RHSs of (79) and (80):

$$H(\widehat{\theta}^-) + [\pi(\widehat{\theta}) - \pi(\widehat{\theta}^-)] R'(\widehat{\theta}). \quad (81)$$

Using the continuity of x_A, z_A, f, V_A , (78), (77) and $T = z_A - x_A$, (79) and (80) reduce to (66) and (67) respectively. Therefore $\vartheta^i(\widehat{\theta})$ are nothing but the variation in social welfare $\partial \mathcal{W}_{A,\rho}^i(\widehat{\theta}) / \partial \widehat{\theta}$ resulting from a small increase in the upper productivity level in A . Consequently,

$$\vartheta^i(\widehat{\theta}) > 0, \quad \forall \widehat{\theta} \geq \underline{\theta} \Rightarrow \widehat{\theta}^i = \bar{\theta}, \quad (82)$$

$$\vartheta^i(\bar{\theta}) < 0 \Rightarrow \widehat{\theta}^i < \bar{\theta}. \quad (83)$$

To complete the proof, note that $\widehat{\theta}^i = \underline{\theta}$ is ruled out since $c > 0$ and the laissez-faire is always implementable. \square

Proof of Property 8. We establish that (i) in Lemma 2 holds under the present assumptions. (a) Since ι is continuous at $\hat{\theta}$, $\iota(\hat{\theta}^-) = \iota(\hat{\theta}) \geq 0$ by (76). Hence, the last terms of (67) and (66) are nil. By (FOIC), $\iota(\hat{\theta}^-) V'_A(\hat{\theta}) \geq 0$. (b) In addition, $\gamma T(z_A(\hat{\theta})) > 0$. Under the Resident criterion, $\phi_\rho(V_A) - \mathcal{W}_{A,\rho}^R \geq 0$. Indeed, as $\phi' > 0$, (FOIC) yields:

$$\phi_\rho(V_A(\hat{\theta})) \geq \frac{1}{F(\hat{\theta})} \int_{\underline{\theta}}^{\hat{\theta}} \phi_\rho(V_A) dF(\theta) \equiv \mathcal{W}_{A,\rho}^R(\hat{\theta}). \quad (84)$$

Under the Citizen criterion, $R(\theta) \geq 0$ for $\theta \leq \hat{\theta}$, and thus $\phi_\rho(V_A(\hat{\theta})) \geq \phi_\rho(V_B(\hat{\theta}) - c(\hat{\theta}))$. Finally, $\vartheta^i(\hat{\theta}) > 0$ for all $\hat{\theta} > \underline{\theta}$. \square

Proof of Property 9. For any $\hat{\theta}$, $R = 0$ on an interval of positive length $[\hat{\theta} - \varepsilon, \hat{\theta}]$. Hence $R' = 0$ at $\hat{\theta}$. We then use the arguments in (b) in the proof of Property 8. \square

Proof of Proposition 5. We use (ii) in Lemma 2 and replace $\mathcal{W}_{A,\rho}^R(\bar{\theta})$ by $W_{A,\rho}^N$. $R' < 0$ because $\bar{\theta}$ is an isolated point where (PC) is active. \square

A.4. Simulations

We use *sufficient* conditions to construct optimal tax schedule. These conditions are equivalent to the necessary ones, provided concavity restrictions are added. By S-S (Theorem 1, p. 317-318), they read:

$$(49) - (54), \quad (85)$$

$$z'_A \geq 0, \quad (86)$$

$$H^N \text{ concave in } l_A \text{ and in } V_A, \quad R \text{ quasiconcave in } V_A, \quad (87)$$

where H^N is defined by (48).

Our strategy is to look for candidate schedules for which ι is continuous and $z'_A \geq 0$, without taking (86) explicitly into account. By (55) and (56), these candidates are such that:

$$\frac{T'}{1-T'} = \left(1 + \frac{1}{e}\right) \frac{1}{\theta f} \left(1 - F(\theta) - \int_{\theta}^{\bar{\theta}} \pi'(\tau) d\tau - \iota(\bar{\theta})\right) \text{ and } \pi' = f - \iota'. \quad (88)$$

If they satisfy all other sufficient conditions, they are then optimal.

We start by noting that conditions (87) always hold. Indeed, quasiconcavity of R is direct. H^N is concave in V_A as $\partial^2 H^N / \partial V_A^2 = 0$. Moreover, differentiating (49) and using (55), one gets:

$$\frac{\partial^2 H}{\partial l_A^2} = -\frac{1}{e} l_A^{1/e-1} f \left[1 - \frac{\iota}{\theta f} \left(1 + \frac{1}{e}\right)\right] = -\frac{1}{e} l_A^{1/e-1} f \frac{T'}{1-T'} \leq 0. \quad (89)$$

Concavity of H in l_A is obtained since $T' = 1 - s < 1$.

We now turn to the computational procedure. We make the guess that θ^* is such that $\pi' > 0$ for all $\theta > \theta^*$ at the optimal solution and look therefore for candidates having this property.

We start by choosing a value for θ^* . We use $V_A = z_A - T(z_A) - v(z_A/\theta)$, (FOIC) and (6) to derive z_A and $T(z_A)$ for $\theta \geq \theta^*$. We compute $T'(z_A) = (dT'/d\theta)/z'_A(\theta)$ for $\theta \geq \theta^*$. By (88), $\iota(\bar{\theta})$ and $\int_{\theta^*}^{\bar{\theta}} \pi'(\tau) d\tau$ are equal to

$$\iota(\bar{\theta}) = -\frac{T'(z_A(\bar{\theta}))}{1 - T'(z_A(\bar{\theta}))} \frac{e}{1 + e} \bar{\theta} f(\bar{\theta}), \quad (90)$$

$$\int_{\theta^*}^{\bar{\theta}} \pi'(\tau) d\tau = 1 - F(\theta^*) - \frac{e}{1 + e} \theta^* f(\theta^*) \frac{T'(z_A(\theta^*))}{1 - T'(z_A(\theta^*))} - \iota(\bar{\theta}). \quad (91)$$

For $\theta^* \leq \theta < \bar{\theta}$, we compute ι , derive ι' and get $\pi' = f - \iota'$. For $\theta < \theta^*$, T' are obtained from (88). We then compute l_A . Since

$$V_A(\theta) = V_A(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} l_A^{1+1/e}(\tau) / \tau d\tau \quad (92)$$

by integration of (FOIC) and $T(\theta l_A) = \theta l_A - v(l_A) - V_A$, we have:

$$\int_{\underline{\theta}}^{\bar{\theta}} T(z_A) f d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \left[\theta l_A - v(l_A) - V_A(\underline{\theta}) - \int_{\underline{\theta}}^{\theta} \frac{l_A^{1+1/e}(\tau)}{\tau} d\tau \right] f d\theta,$$

which leads to

$$V_A(\underline{\theta}) = \int_{\underline{\theta}}^{\bar{\theta}} \left[\theta l_A f - (1 - F(\theta)) \frac{l_A^{1+1/e}}{\theta} - \frac{l_A^{1+1/e}}{1 + 1/e} f \right] d\theta \quad (93)$$

by (TR) and Fubini's theorem. $V_A(\theta)$ is then obtained from (92)–(93). We check that $V_A > V_B - c$ for $\theta < \theta^*$, $V_A(\theta) = V_B(\theta) - c(\theta)$ for $\theta \geq \theta^*$, and $z'_A \geq 0$. If it is the case, the candidate schedule is an optimal one.

References

- Bodin, J. (1578). *Réponse Au Paradoxe de Monsieur de Malestroit*. Paris: Jean du Pays.
- Borjas, G. (1999). The economic analysis of immigration. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Chapter 28, pp. 1697–1760. North-Holland.
- Brito, D. and W. Oakland (1977). Some properties of the optimal income-tax. *International Economic Review* 18, 407–423.
- Chander, P. and L. Wilde (1998). A general characterization of optimal income tax enforcement. *Review of Economic Studies* 65, 165–183.

THREAT BY 0% OF THE POPULATION (BENCHMARK), $W_A = 17\,132\text{€}$						
$F(\theta)$	V_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$	
0.05	17 156€	3 979€	-13 203€	96.0%	-331.8%	
0.50	18 028€	14 037€	- 4 487€	78.8%	- 32.0%	
0.95	27 917€	57 555€	+27 917€	73.4%	+ 48.5%	
THREAT BY 10% OF THE POPULATION						
$W_A^N = 14\,682\text{€}$ ($Loss = -14.3\%$); $\hat{z}_A = 41\,955\text{€}$; $z_A^* = 50\,746\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 12.6\%$.						
$F(\theta)$	V_A	$\Delta V_A\%$	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	14 709€	-14.3%	4 064€	-10 674€	95.7%	-262.6%
0.50	15 738€	-12.7%	14 526€	- 1 820€	74.9%	- 12.5%
0.95	46 970€	+73.4%	75 015€	+15 572€	00.0%	+ 20.7%
THREAT BY 5% OF THE POPULATION						
$W_A^N = 15\,622\text{€}$ ($Loss = - 8.8\%$); $\hat{z}_A = 57\,238\text{€}$; $z_A^* = 74\,952\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 6.7\%$.						
$F(\theta)$	V_A	ΔV_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	15 622€	- 8.8%	4 020€	-11 654€	96.0%	-289.9%
0.50	16 592€	- 8.0%	14 270€	- 2 869€	77.0%	- 20.1%
0.95	34 613€	+27.8%	75 015€	+27 900€	00.0%	+ 37.2%
THREAT BY 3% OF THE POPULATION						
$W_A^N = 16\,131\text{€}$ ($Loss = - 5.8\%$); $\hat{z}_A = 70\,834\text{€}$; $z_A^* = 98\,779\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 4.4\%$.						
$F(\theta)$	V_A	ΔV_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	16 156€	- 5.8	4 003€	-12 179€	96.0%	-304.2%
0.50	17 071€	- 5.3	14 174€	- 3 422€	77.8%	- 24.1%
0.95	29 570€	+ 9.2	64 652€	+29 957€	52.5%	+ 46.3%
THREAT BY 1% OF THE POPULATION						
$W_A^N = 16\,797\text{€}$ ($Loss = - 2.0\%$); $\hat{z}_A = 104\,989\text{€}$; $z_A^* = 167\,006\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 2.0\%$.						
$F(\theta)$	V_A	ΔV_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	16 821€	- 2.0	3 987€	-12 860€	96.1%	-322.6%
0.50	17 707€	- 1.8	14 082€	- 4 130€	78.5%	- 29.3%
0.95	27 636€	+ 1.0	59 413€	+28 691€	68.8%	+ 48.3%
THREAT BY 0.5% OF THE POPULATION						
$W_A^N = 16\,993\text{€}$ ($Loss = - 0.8\%$); $\hat{z}_A = 126\,638\text{€}$; $z_A^* = 261\,228\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 1.3\%$.						
$F(\theta)$	V_A	ΔV_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	17 017€	- 0.8	3 983€	-13 059€	96.1%	-322.0%
0.50	17 896€	- 0.7	14 059€	- 4 337€	78.7%	- 30.8%
0.95	27 366€	+ 1.0	59 439€	+28 279€	71.3%	+ 48.4%
THREAT BY 0.1% OF THE POPULATION						
$W_A^N = 17\,122\text{€}$ ($Loss = - 0.1\%$); $\hat{z}_A = 160\,829\text{€}$; $z_A^* = 302\,894\text{€}$; $1 - F(\theta_{\hat{z}_A}) = 0.7\%$.						
$F(\theta)$	V_A	ΔV_A	z_A	$T(z_A)$	$T'(z_A)$	$T(z_A)/z_A$
0.05	17 146€	- 0.1	3 980€	-13 002€	96.2%	-331.5%
0.50	18 019€	- 0.1	14 041€	- 4 474€	78.8%	- 31.9%
0.95	27 158€	+ 0.3	57 726€	+27 972€	73.0%	+ 48.5%

Note: "Loss" in social welfare w.r.t. benchmark; $\hat{z}_A := z_A$ such that $T(z_A)/z_A$ maximum;
 $z_A^* := \min z_A$ with (PC) active; $(\hat{z}_A, z_A^*) = \text{range of decrease in } T(z_A)/z_A \text{ before (PC) active}$;
 $1 - F(\theta_{\hat{z}_A}) = \% \text{ agents with } T(z_A)/z_A \text{ decreasing}$; $\Delta V_A := \text{change in } V_A \text{ w.r.t. benchmark}$.

Table 1: Optimum Allocations (Maximin, $e=0.2$, constant migration costs)

- d'Autume, A. (2000). L'imposition optimale du revenu: une application au cas français. *Revue Française d'Economie* 3, 3–63.
- DGI (2005). Les délocalisations de contribuables (personnes physiques), Bilan des départs recensés pour l'année 2003. Technical report, MINEFI.
- 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.
- Hindriks, J. (1999). The consequences of labour mobility for redistribution: Tax vs. transfert competition. *Journal of Public Economics* 74, 215–234.
- Huber, B. (1999). Tax competition and tax coordination in an optimal income tax model. *Journal of Public Economics* 71, 441–458.
- Hungerbühler, M., E. Lehman, A. Parmentier, and B. Van der Linden (2006). Optimal redistributive taxation in a search equilibrium model. *Review of Economic Studies* 73, 715–739.
- 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.
- Laslier, J.-F., A. Trannoy, and K. Van Der Straeten (2003). Voting under ignorance of job skills of unemployed: The overtaxation bias. *Journal of Public Economics* 87, 595–626.
- 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. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38(2), 175–208.
- Mirrlees, J. (1974). Notes on welfare economics, information and uncertainty. In M. Balch, M. McFadden, and S. Wu (Eds.), *Essays in economic behavior under uncertainty*, pp. 243–258. North Holland.

- Mirrlees, J. (1982). Migration and optimal income taxes. *Journal of Public Economics* 18, 319–341.
- 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.
- Page, F. and P. Monteiro (2003). Three principles of competitive nonlinear pricing. *Journal of Mathematical Economics* 39, 63–109.
- Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.
- Piaser, G. (2003). Labor mobility and income tax competition. CORE Discussion Paper 2003/6, UCL.
- Piketty, T. (1997). La redistribution fiscale face au chômage. *Revue Française d'Economie* 12, 157–201.
- Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies* 43, 261–267.
- Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies* 68, 205–229.
- Sahota, G. (1968). An economic analysis of internal migration in brazil. *Journal of Political Economy* 76, 218–245.
- Schwartz, A. (1973). Interpreting the effect of distance on migration. *Journal of Political Economy* 81, 1153–1169.
- Seade, J. (1982). On the sign of the optimum marginal income tax. *Review of Economic Studies* 49, 637–643.
- Seade, J. K. (1977). On the shape of optimal tax schedule. *Journal of Public Economics* 7, 203–235.
- Seierstad, A. and K. Sydsaeter (1987). *Optimal Control Theory with Economic Applications*. Amsterdam: North-Holland.
- Simula, L. and A. Trannoy (2006, September, forthcoming). Optimal linear income tax when agents vote with their feet. *FinanzArchiv*.
- Sjaastad, L. (1962). The costs and returns of human migration. *Journal of Political Economy* 70, 80–93.

- Stiglitz, J. (1982, March). Self-selection and pareto efficient taxation. *Journal of Public Economics* 17(2), 213–240.
- Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy* 64, 416–424.
- TRL (2005). The price of off-shore. Technical report, Tax Justice Network.
- Tuomala, M. (1990). *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.
- Weymark, J. (1986). A reduced-form optimal nonlinear income tax problem. *Journal of Public Economics* 30, 199–217.
- Weymark, J. (1987). Comparative static properties of optimal nonlinear income taxes. *Econometrica* 55, 1165–1185.
- Wilson, J. D. (1980). The effect of potential emigration on the optimal linear income tax. *Journal of Public Economics* 14, 339–353.
- Wilson, J. D. (1982). Optimal linear income taxation in the presence of emigration. *Journal of Public Economics* 18, 363–379.