Tax Me If You Can:
Optimal Nonlinear Income Tax between Competing Governments*

Etienne LEHMANN†
CRED (TEPP) Université Panthéon-Assas & CREST

Laurent SIMULA ‡
Uppsala University & Uppsala Center for Fiscal Studies

Alain TRANNOY §
Aix-Marseille Université (Aix-Marseille School of Economics), CNRS & EHESS

April 15,2013
Preliminary Version

Abstract

We investigate how the optimal nonlinear income tax schedule is modified when taxpayers can evade taxation by emigrating. We consider two symmetric countries with Maximin governments. Workers choose their labor supply along the intensive margin. The skill distribution is continuous, and, for each skill level, the distribution of migration cost is also continuous. We show that optimal marginal tax rates are nonnegative at the symmetric Nash equilibrium when the semi-elasticity of migration is decreasing in the skill level. When the semi-elasticity of migration is increasing in the skill level, either optimal marginal tax rates are positive everywhere or they are positive for the lower part of the skill distribution and then negative. Numerical simulations are calibrated using plausible values of the semi-elasticity of migration for top income earners. We show that the shape of optimal tax schedule varies significantly, depending on the profile of the semi-elasticity of migration over the entire skill distribution - a profile over which we lack empirical evidence.

JEL Classification: D82, H21, H87

Keywords: Optimal Income Tax, Income Tax competition, Migration, Labor Mobility, Nash-Equilibrium Tax Schedules

---

*We are grateful to Spencer Bastani, Sören Blomquist, Pierre Boyer, Nathalie Ethchart-Vincent, Laurence Jacquet, Eckhart Janeba, Hubert Kempf, Jean-Baptiste Michau, Olle Folke, Régis Renault, Hakån Selin, Emmanuelle Taugourdeau, Farid Toubal, Bruno Van der Linden and John D. Wilson, as well as participants to seminars at CREST, Ecole Polytechnique, UCFS workshop, the University of Mannheim, the University of Rennes (CREM), the University of Panthéon Assas Paris 2, the ENS-Cachan/Paris-Sud Hotelling seminar, the University of Cergy (THEMA) and the IIES (Stockholm).

†Address: CREST-INSEE, Timbre J360, 15 boulevard Gabriel Péri, 92245, Malakoff Cedex, France. Email:etienne.lehmann@ensae.fr. Etienne Lehmann is also research fellow at IRES-Université Catholique de Louvain, IDEP, IZA and CESifo and junior research fellow of Institut Universitaire de France.

‡Address: Uppsala University and Uppsala Center for Fiscal Studies (UCFS). Box 513, SE-751 20 Uppsala, Sweden. Email: laurent.simula@nek.uu.se. Laurent Simula is also research fellow at IDEP and CESifo, and a research associate at Centre d'Economie de la Sorbonne - Ecole Normale Supérieure de Cachan.

§Address: Aix-Marseille Université (Aix-Marseille School of Economics), CNRS and EHESS. 2, rue de la Charité, 13 002 Marseille, France. Email: alain.trannoy@univmed.fr

I Introduction

In his 1971 seminal article, Mirrlees assumes that migrations are impossible but emphasizes that "since the threat of migration is a major influence on the degree of progression in actual tax systems, this is an assumption one would rather not make"(Mirrlees, 1971, p. 176). This threat of migration is certainly even more topical after four decades of increasing globalization.

This article addresses the design of optimal non-linear income taxes when governments compete on a potentially mobile tax base. The world population consists of individuals both differing in skills and costs of migration. This population is initially perfectly shared between two identical countries. In each country, a benevolent policy-maker aims at redistributing wealth from the more to the less productive individuals. In doing so, the former only knows the joint distribution of skills and migration costs. In particular, it is unable to observe the type of a particular individual. Individual makes choices on two margins. The choice of taxable income operates on the intensive margin, whereas the location decision operates at the extensive margin. An individual decides to move abroad if his/her indirect utility at home is lower than his/her best outside option.[1] The outside option depends on the indirect utility abroad and the individual-specific costs of migration incurred in case of relocation. As emphasized by Borjas (1999), these costs "probably vary among persons [but] the sign of the correlation between costs and wages is ambiguous". For this reason, we do not make any assumption on the relationship between skills and migration costs. We simply consider that there is a distribution of migration costs for each possible skill level.

The model is designed to cast light on the main effects of migrations due to international differences in income taxes. Both countries have the same production function because we do not want individual productivities, and thus pre-tax wages, to depend on the residence country.[2] We characterize the best-response allocations in the two countries, before focusing attention to the symmetric Nash equilibrium tax schedules. In a symmetric Nash equilibrium, migration does not actually take place, but the tax schedules are modified because of the *threat* of migration.

In order to highlight the main economic effects and intuitions, we choose to restrict attention to the case where there is no income effect on the choice of taxable income. Individual preferences over consumption and effort are thus represented by a quasilinear-in-consumption utility function. Because most of the empirical studies give credence to small income effects relative to substitution effects, this case provides a relevant benchmark, which has been extensively used in the literature since the influential work by Diamond (1998).[3] In addition, we concentrate on the situation where each policy-maker maximises the well-being of its worst-off citizens (maximin).[4] Hence, we place ourselves in the situation that would lead, in each country, to the most progressive tax scheme in the absence of mobility (or in the presence of tax coordination), and examine to which extent the latter is modified due to tax competition.

Our main findings can be summarized as follows. We first characterize the best-responses

[1] This is in accordance with 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).

[2] The mobility of highly skilled for tax purposes induces both losses in taxes and in productive capacities in the left countries. It differs from the "brain drain" (Bhagwati and Partington, 1976; Bhagwati, 1976) because its key parameter is not the change in productivity resulting from emigration.

[3] See, e.g., Atkinson (1990); Piketty (1997); d'Autume (2000); Boadway and Pestieau (2007); Saez (2001, 2002).

[4] See Boadway and Jacquet (2008) for a recent study of the optimal tax scheme under the maximin in the absence of individual mobility.

of each policy-maker and obtain a simple formula for the optimal marginal tax rates in the symmetric Nash equilibrium. We interpret this formula using a small tax reform perturbation around the equilibrium. We show that a "migration effect" takes place in addition to the usual closed-economy behavioural responses (see Diamond (1998)). When marginal tax rates are slightly increased on some interval, all individuals above it are facing a lump-sum increase in taxes. This increases out-migration and reduce in-migration. This new effect basically depends on the semi-elasticity of migration, i.e. on the percentage change in the density of taxpayers of a given skill level when their consumption is increased by one unit.[5] We then provide a full characterization of the overall shape of the tax function. First, when the semi-elasticity is decreasing in skills, the tax function is increasing and the top marginal tax rate is strictly positive. This is for example the case when the elasticity of migration is constant. Second, when the semi-elasticity of migration is constant, the tax function is increasing and top marginal tax rates converge to zero. This situation is for example obtained when skills and migration costs are independent. Third, when the semi-elasticity is increasing, the tax function may be increasing, with positive top marginal tax rates, or hump-shaped, with negative top marginal tax rates. A sufficient condition for the hump-shaped pattern is that the semi-elasticity becomes arbitrarily large for top income earners. In that case, progressivity of the optimal tax schedule does not only collapse because of tax competition; the tax liability itself becomes strictly decreasing. This means that there are "middle-skilled" individuals who pay higher taxes than top-income earners. A situation that we can describe as a "curse of the middle-skilled" (Simula and Trannoy, 2010). We then show, through numerical simulations, that the upper part of the tax schedule may be highly sensitive to slight variations in the slope of the semi-elasticity. Our results can be summarized in terms of sufficient statistics: both the semi-elasticity of migration and how it evolves along the skill distribution are required to characterize the optimal tax function, *even at the top*. As far as we know, there are very few empirical studies providing insights into the slope of the semi-elasticity.

There are relatively few articles that consider strategic interaction among governments which can employ fully non-linear income taxes when some individuals are free to choose both their effort and country of residence. As far as we know, Osmundsen (1999) is one of the the first to examine income taxation with type-dependent outside options. This article studies how highly skilled individuals distribute their working time between two countries. Because it directly uses the model Maggi and Rodriguez-Clare (1995), there is no individual trade-off between consumption and effort (as in Mirrlees (1982)). Following Mirrlees (1971), our model takes this trade-off into account. In a recent article, Krause (2008) has examined income taxation and education policy when there exist conflicting incentives for individuals to understate and overstate their productivity. Highly-skilled individuals are better educated and can thus benefit from higher outside options when emigrating. Using quasilinear-in-leisure preferences and a two-type model, different possible regime are identified but no optimal tax scheme is characterized. Moreover, several articles have adopted the viewpoint of tax competition, restricting attention to personalised lump-sum taxes (Leite-Monteiro, 1997), considering a two-type population as in Stiglitz (1982) (Huber, 1999; Hamilton and Pestieau, 2005; Piaser, 2003; Lipatov and Weichenrieder, 2012) or a population with many types (Bierbrauer and Weymark, 2011; Morelli, Yang, and Ye, 2012). Bierbrauer and Weymark (2011) consider that labour is perfectly mobile across countries.

By identifying the key parameters to estimate, our paper also helps clarify some results

[5]The elasticity of migration corresponds to the product of the semi-elasticity and consumption level. In the second best, consumption must be non-decreasing in skills.

obtained in the literature. Brewer, Saez, and Shephard (2010) find that top marginal tax rates should be strictly positive and derive a simple formula to compute them. In contrast, Blumkin, Sadka, and Shem-Tov (2012) find that top marginal tax rates should be zero. This is because the first paper assumes that the elasticity of migration is constant. This implies that the semi-elasticity is decreasing and, thus, that the tax function is increasing, with a positive asymptotic tax rate. The second paper assumes that the skills and migration costs are independent. This implies that the semi-elasticity of migration is constant and, thus, that the asymptotic marginal tax rates are zero. We see that the underlying assumptions on the semi-elasticity of migration and its slope are of critical relevance.

The article is organized as follows. Section 2 sets up the model. Section 3 derives the optimal tax formula for the symmetric Nash equilibrium. Section 4 shows how to sign the optimal marginal tax rates and provides some further characterization of the whole tax function. Section 5 uses numerical simulations to investigate the sensitivity of the tax function to the slope of the semi-elasticity of migration. Section 6 concludes.

II Model

We consider an economy consisting of two symmetric countries, indexed by $i = A, B$. There is a mass 2 of workers. The same technology is available in both countries. It exhibits constant returns to scales. Hence, workers are paid up to their productivity, which is independent of location. Each worker is characterized by three characteristics: her native country $i \in \{A, B\}$, her productivity (or skill) $w \in [w_0, w_1]$, and the migration cost $m \in \mathbb{R}^+$ she supports if she decides to live abroad. Note that w_1 may be either finite or infinite and w_0 is non-negative. In addition, the empirical evidence that some people are immobile is captured by the possibility of infinitely large migration costs.[6] The migration cost corresponds to a loss in utility, due to various material and psychic costs of moving: application fees, transportation of persons and household's goods, forgone earnings, costs of speaking a different language and adapting to another culture, costs of leaving one's family and friends, etc. As emphasized by Borjas (1999), these costs "probably vary among persons [but] the sign of the correlation between costs and wages is ambiguous". For this reason, we do not make any assumption on the relationship between skills and migration costs. We simply consider that there is a distribution of migration costs for each possible skill level. In the next sections, we will see that some apparently innocuous assumptions on this relationship may have significant consequences on the shape of the whole optimal tax profile, even for top income earners. Alternatively, the cost of migration can be regarded as the costs incurred by cross-border commuters, who still reside in their home country but work across the border.

The joint distribution of skills w and migration costs m is initially identical in the two countries. We denote by $f(w)$ the continuously-differentiable skill density, and by $F(w) \equiv \int_{w_0}^{w} f(x)\, dx$ the corresponding cumulative distribution function (CDF). For each skill w, $g(m|w)$ denotes the conditional density of the migration cost and $G(m|w) \equiv \int_0^m g(x|w)\, dx$ the conditional CDF. In each country, the initial joint density of (m, w) is thus $g(m|w)\, f(w)$; note that $G(m|w)\, f(w)$ is the density of individuals of skill w whose migration cost is lower than m.

Following Mirrlees (1971) seminal article, we consider that there is a fundamental distinction between public and private information. The government does not observe individual types (w, m). Moreover, it is constrained to treat native and immigrant workers in the same

[6] Alternatively, we could assume that $m \in [m_0, m_1]$ but this would only complexify the analysis without changing the main insights.

way.[7] Therefore, it can only condition transfers on earnings y through an income tax function $T_i(.)$. It is unable to base the tax on an individual's skill level w, migration cost m, or native country.

II.A Individual Choices

Every worker derives utility from consumption c, and disutility from effort and migration, if any. In the original article by Mirrlees (1971), effort is synonymous of labour supply. Note that effort is a more general concept than working hours, and can encompass choices made by self-workers and entrepreneurs. Let $v(y;w)$ be the disutility of a worker of skill w to obtain pre-tax earnings $y \geq 0$. Let $\mathbb{1}$ be equal to 1 if she decides to migrate, and to zero otherwise. Individual preferences are described by the quasi-linear utility function:

$$c - v(y;w) - \mathbb{1} \times m. \tag{1}$$

The quasi-linearity in consumption implies that there is no income effect on taxable income. Even though there is much less evidence on the magnitude of the income effects in the literature estimating the effect of taxation on reported income than in the labour supply literature, the quasi-linear specification seems to be a reasonable approximation. For example, Gruber and Saez (2002) estimate both income and substitution effects in the case of reported incomes, and find small and insignificant income effects. The cost of migration is additively separable. It is introduced in the model as a monetary loss, which might be due, as previously emphasized, to material or psychological costs. Because of additive separability, two individuals living in the same country and having the same skill level choose the same gross income/consumption bundle, irrespective of their native country. Also, if the two countries implement the same tax schedule, a given individual chooses the same bundle at home and abroad.

The choice of effort corresponds to an intensive margin and the migration choice to an extensive margin.

II.A.1 Intensive Margin

The disutility $v(.;.)$ of effort is a twice continuously differentiable function. It is increasing and convex in effort, thereby in pre-tax earnings y. Moreover, it is decreasing in w because it is easier for a more productive individual to earn a given pre-tax income y. Finally, the marginal cost of increasing pre-tax income is larger for more productive agents. Because indifference curves have equation $c = v(y;w) + u$, this assumption implies that the Spence-Mirrlees strict single-crossing condition holds. In summary:

Assumption 1 *The disutility function $v(.;.)$ satisfies $v'_y > 0 > v'_w$ and $v''_{yy} > 0 > v''_{yw}$.*

Every individual living in country i is liable to an income tax $T_i(.)$, which is solely based on earnings $y \geq 0$, and thus in particular independent of the native country. Therefore, a worker of skill w, who has chosen to work in country i, solves:

$$U_i(w) \equiv \max_y \quad y - T_i(y) - v(y;w). \tag{2}$$

[7] In several countries, highly skilled foreigners are eligible to specific tax cuts for a limited time duration. This is for example the case in Sweden and in Denmark. These exemptions are temporary. In Switzerland, foreign nationals – who take residence there for the first time or after ten years of absence, and who are not employed here –, can benefit from the so-called lump sum taxation regime.

We call $U_i(w)$ the *gross* utility of a worker of skill w in country i. It is the net utility level for a native and the utility level absent migration cost for an immigrant. We call $Y_i(w)$ the solution to programme (2) and $C_i(w) = Y_i(w) - T\left(Y_i(w)\right)$ the consumption level of a worker of skill w in country i.[8] The first-order condition can be written as:

$$1 - T_i'\left(Y_i(w)\right) = v_y'\left(Y_i(w); w\right). \tag{3}$$

Increasing effort to get one extra unit of pre-tax income increases consumption by $1 - T_i'\left(Y_i(w)\right)$ units, but reduces utility by $v_y'\left(Y_i(w); w\right)$ units. Differentiating (3), we obtain the elasticity of gross earnings with respect to the retention rate $1 - T_i'$ and skill level w:

$$\varepsilon_i(w) \equiv \frac{1 - T_i'\left(Y_i(w)\right)}{Y_i(w)} \frac{\partial Y_i(w)}{\partial\left(1 - T_i'\left(Y_i(w)\right)\right)} = \frac{v_y'\left(Y_i(w); w\right)}{Y_i(w)\; v_{yy}''\left(Y_i(w); w\right)}, \tag{4}$$

$$\alpha_i(w) \equiv \frac{w}{Y_i(w)} \frac{\partial Y_i(w)}{\partial w} = -\frac{w\; v_{yw}''\left(Y_i(w); w\right)}{Y_i(w)\; v_{yy}''\left(Y_i(w); w\right)}. \tag{5}$$

II.A.2 Extensive Margin

Migration decisions correspond to a choice along the extensive margin. We start with the migration decisions of individuals born in country A. An individual of type (w, m) gets utility $U_A(w)$ if she stays in A and utility $U_B(w) - m$ if she relocates to B. She therefore emigrates if and only if

$$m < U_B(w) - U_A(w).$$

Hence, among individuals of skill w born in country A, the mass of emigrants is given by $G\left(U_B(w) - U_A(w)\,|w\right) f(w)$ and the mass of agents staying in their native country by $\left(1 - G\left(U_B(w) - U_A(w)\,|w\right)\right) f(w)$. Individuals born in country B behave in a symmetric way. They leave their home country if and only if $m < U_A(w) - U_B(m)$. Hence, among individuals of skill w, the mass of emigrants from B to A is $G\left(U_A(w) - U_B(w)\,|w\right) f(w)$, while the mass of native residents is $\left(1 - G\left(U_A(w) - U_B(w)\,|w\right)\right) f(w)$.

It is important to note that, at a given skill level, migration flows are going in only one direction. Combining the migration decisions made by agents born in the two countries, we see that the mass of residents of skill w in country A depends on the difference in the gross utility levels $\Delta = U_A(w) - U_B(w)$, with:

$$\varphi(\Delta; w) \equiv \begin{cases} \left(1 + G(\Delta|w)\right) f(w) & \text{when } \Delta \geq 0, \\ \left(1 - G(-\Delta|w)\right) f(w) & \text{when } \Delta \leq 0. \end{cases} \tag{6}$$

The function $\varphi(.; w)$ is continuously differentiable, with derivative $\partial\varphi(.; w)/\partial\Delta = g\left(|\Delta|\;\middle|w\right)\; f(w)$. It is increasing in the difference Δ in the gross utility levels. By symmetry, the mass of residents of skill w in country B is given by $\varphi\left(U_B(w) - U_A(w); w\right)$.

All the responses along the extensive margin can be summarized in terms of elasticity concepts. We define the semi-elasticity of migration in country i as:

$$\eta_i\left(\Delta_i(w); w\right) \equiv \frac{\partial\varphi(\Delta_i(w); w)}{\partial C_i(w)} \frac{1}{\varphi(\Delta_i(w); w)} \quad \text{with } \Delta_i(w) = U_i(w) - U_{-i}(w). \tag{7}$$

It corresponds to the percentage change in the density of taxpayers with skill w when their

[8] If (2) admits more than one solution, we make the tie-breaking assumption that individuals choose the one preferred by the government.

consumption $C_i(w)$ is increased at the margin. The elasticity of migration is defined as:

$$\nu_i(\Delta_i(w); w) \equiv C_i(w) \times \eta(\Delta_i(w), w). \tag{8}$$

We will see in the next sections that the semi-elasticity plays a direct part in the characterization of the symmetric Nash equilibrium. In particular, we will see that small changes in this parameter may imply significant variations in the whole tax profile, including at the top of the income distribution. We will also see that the sign marginal tax rates of high-income earners is directly connected with the semi-elasticity and clarify under which testable conditions the latter are positive, equal to zero, or even negative.

II.B Governments

In each country $i = A, B$, a benevolent policy-maker designs the tax system so as to maximize the welfare of the worst-off individuals. We chose a maximin criterion for several reasons. The maximin tax policy is the most redistributive one, as it corresponds to an infinite aversion to income inequality. A first motivation is therefore to explore the domain of potential redistribution in the presence of tax competition. A second motivation is that in an open economy, there is no obvious way of specifying the set of agents whose welfare is to count (Blackorby, Bossert, and Donaldson, 2005). The policy-maker may care for the well-being of the natives, irrespective of their country of residence. Alternatively, it may only account for the well-being of the native taxpayers, or for that of all taxpayers irrespective of native country. As an economist, there is no reason to favour one of these criteria (Mirrlees, 1982). We focus on the maximin because the set of agents whose welfare is accounted for is then independent of the tax policy. So all these criteria are equivalent.[9]

Most countries do not levy income taxes abroad.[10] To make the analysis more transparent and highlight the main forces at stake, we herein consider that taxes are levied according to the residence principle. This implies that the budget constraint faced by country i's government is:

$$\int_{w_0}^{w_1} T_i(Y(w)) \ \varphi(U_i(w) - U_{-i}(w); w) \ dw \geq E \tag{9}$$

where $E \geq 0$ is an exogenous amount of public expenditures to finance.

III Optimal Tax Formula

Following Mirrlees (1971), the standard optimal income tax formula provides the optimal marginal tax rates that should be implemented in a closed economy (e.g., Atkinson and Stiglitz (1980); Diamond (1998); Saez (2001)). From another perspective, these rates can also be seen as those that should be implemented by a supranational organization ("world welfare point of view" in Wilson (1982)) or in the presence of tax cooperation. In this section, we derive the optimal marginal tax rates when policy-makers compete on a common pool of taxpayers. We investigate in which way this formula differs from the standard one. We start with the characterization of the best response allocations, before focusing on the symmetric Nash equilibria. We provide a formal as well as an intuitive derivation based on the analysis

[9] This equivalence holds because, at each skill level, the conditional distribution of migration costs m is unbounded from above. Hence, there are always individuals for whom migration is not a valuable option. See Simula and Trannoy (2011).

[10] US citizens, though, are liable to the US income tax on their world incomes; however, US citizens living abroad benefit from a general tax exclusion of \$92,900 in 2011.

of the effects of a small tax reform perturbation around the equilibrium (Piketty, 1997; Saez, 2001).

III.A Best Responses

Each government is unable to condition taxes on skill levels w, migration costs m and native country. It thus faces a multidimensional screening problem. However, because migration costs enter separably in the individual utility function (1), two individuals with the same skill level make the same intensive choice irrespective of their other personal characteristics. The fact that migration costs and native country are unobservable thus only matters for the migration decision and not for the intensive one. The government's problem thus belongs to the class of multidimensional screening problems with random participation (Rochet and Stole, 2002; Jacquet, Lehmann, and Van der Linden, 2012).

It is easy to extend the standard taxation principle (Hammond, 1979; Guesnerie, 1995) to our economy with tax competition. The main reason is that every agent *in fine* interacts with only one government. It is therefore equivalent to consider that the policy-maker chooses a tax schedule $y \mapsto T_i(y)$ or an incentive-compatible allocation $w \mapsto (C_i(w), Y_i(w))$ that verifies:

$$C_i(w) - v(Y_i(w); w) \geq C_i(w') - v(Y_i(w'); w) \text{ for any } w, w' \text{ in } [w_0, w_1]. \tag{10}$$

These incentive-compatibility constraints ensure that any individual of skill w that chooses to live in country i prefers the bundle $(C_i(w), Y_i(w))$ designed for her, to any bundle $(C_i(w'), Y_i(w'))$ designed for any other skill level w', so the allocation is truthful-telling. Because of the single-crossing condition $v''_{yw} < 0$, these constraints are equivalent to:

$$U_i'(w) = -v_w'(Y_i(w); w), \tag{11}$$

$$Y_i(\cdot) \text{ non-decreasing.} \tag{12}$$

The envelope condition (11) specifies at which rate utility must be increased to induce truth-telling. The second-order condition for incentive-compatibility (12) states that gross income $Y_i(\cdot)$ must be weakly increasing. It in particular implies that $C_i(w)$ is non-decreasing in w. We adopt the so-called 'first-order approach' which do not explicitly account for the monotonicity of $Y_i(\cdot)$ when solving for the optimal schedules and then check, in computations, that the candidate schedules actually satisfy this condition. We also make the usual regularity assumption that $Y_i(\cdot)$ is differentiable.

The best-response allocation of government i to government $-i$ is solution to:

$$\max_{U_i(w), Y_i(w)} \quad U_i(w_0) \quad \text{s.t.} \quad U_i'(w) = -v_w'(Y_i(w); w) \quad \text{and}$$
$$\int_{w_0}^{w_1} (Y_i(w) - v(Y_i(w); w) - U_i(w)) \varphi(U_i(w) - U_{-i}(w); w) \; dw \geq E$$

The social objective is to maximise the utility $U_i(w_0)$ of the worst-off nationals or, equivalently, the utility $U_i(w_0) - m$ of the worst-off immigrants, subject to budget balancedness and incentive compatibility.

In the optimisation problem, it is convenient to choose $Y_i(w)$ as control variable and $U_i(w)$ as state variable. Indeed, for every w, a unique $C_i(w)$ corresponds to the pair $(U_i(w), Y_i(w))$. Instead of looking at the primal problem, we follow Boadway and Jacquet (2008) and use the dual problem to characterize best response allocations. The dual consists in finding an incentive-compatible allocation $(U_i(w), Y_i(w))$ to maximize collected taxes without reducing the utility of the worst-off individuals below a threshold $\underline{U}_i(w_0)$. Formally:

$$\begin{aligned} &\max_{U_i(w),Y_i(w)} \int_{w_0}^{w_1} \left(Y_i\left(w\right) - v\left(Y_i\left(w\right);w\right) - U_i\left(w\right)\right)\ \varphi\left(U_i\left(w\right) - U_{-i}\left(w\right);w\right)\ dw \\ &\text{s.t.}\quad U_i'\left(w\right) = -v_w'\left(Y_i\left(w\right);w\right) \quad \text{and} \quad U_i(w_0) \geq \underline{U}_i(w_0), \end{aligned} \tag{13}$$

in which $\underline{U}_i(w_0)$ and $U_{-i}(.)$ are given. Denoting $q(.)$ the co-state variable, the Hamiltonian is:

$$\mathcal{H}(U_i, Y_i, q; w) \equiv [Y_i - v(Y_i; w) - U_i]\ \varphi(U_i - U_{-i}; w) - q(w)\ v_w'\left(Y_i; w\right).$$

Using Pontryagin's principle, the first-order conditions for a maximum are:

$$1 - v_y'\left(Y_i(w); w\right) = \frac{q\left(w\right)}{\varphi\left(\Delta_i(w); w\right)}\ v_{yw}''\left(Y_i(w); w\right), \tag{14}$$

$$q'(w) = \{1 - [Y_i(w) - v(Y_i(w); w) - U_i(w)]\ \ \eta_i(\Delta_i(w); w)\}\ \ \varphi\left(\Delta_i(w); w\right), \tag{15}$$

$$q(w_1) = 0 \text{ when } w_1 < \infty \text{ and } q(w_1) \to 0 \text{ when } w_1 \to \infty, \tag{16}$$

$$q\left(w_0\right) \leq 0. \tag{17}$$

III.B Symmetric Nash Equilibria

We focus on symmetric Nash equilibria for two reasons. The first reason is intelligibility and tractability. In this case, the gross utility levels of an agent of skill w are the same in A and in B, i.e. $U_A(w) = U_B(w)$. This implies that $\varphi\left(\Delta_i; w\right) = f(w)$ and $\eta\left(\Delta_i; w\right) = g(0|w)$, using (6) and (7). Note that the semi-elasticity of migration $\eta_0(w)$ is then equal to the structural parameter $g(0|w)$. The optimality conditions (14)-(17) can therefore be simplified. For notational convenience, we denote the semi-elasticity of migration at the symmetric Nash equilibrium by $\eta_0(w)$. Note that it corresponds to a structural parameter of the economy. The second reason is that symmetric Nash equilibria appear as insightful benchmarks to investigate the extent to which optimal tax policies are modified in the presence of tax competition. While the potential for free movement of labour constrains what tax schedules are sustainable, nobody actually moves. By investigating symmetric equilibria, we thus illustrate the impact of the *threat* of migration.

Because we from now on focus on symmetric equilibria, we will drop the A and B subscripts, which are no longer necessary. Condition (15) can also be written as:

$$q'(w) = [1 - \eta(0; w)\ T(Y(w))]\ \ \varphi\left(0; w\right) = [1 - \eta_0(w)\ T(Y(w))]\ f(w).$$

Integrating the latter between w and w_1 and using the transversality condition (16), we obtain:

$$q(w) = -\int_{w}^{w_1} [1 - \eta_0(x)\ T(Y(x))]\ \ f(x)\ dx. \tag{18}$$

Dividing (5) by (4) and making use of (3), we get $v_{yw}''(Y(w); w) = (\alpha(w)/\varepsilon(w))\left(1 - T'(Y(w))\right)/w$. We then substitute the latter equation and 18 in (14) to obtain the following characterization.

Proposition 1 *In a symmetric Nash equilibrium, the optimal allocation can be decentralized by a tax function satisfying:*

$$\frac{T'(Y(w))}{1 - T'(Y(w))} = \frac{\alpha(w)}{\varepsilon(w)} \frac{1 - F(w)}{w f(w)} \left\{1 - \mathbb{E}[\eta_0(x) T(Y(x))\,|x \geq w]\right\}, \tag{19}$$

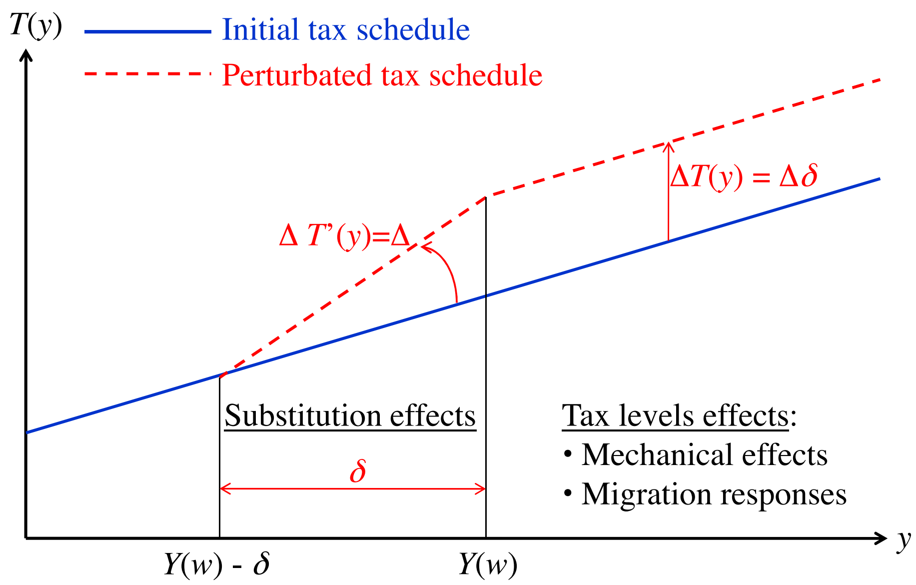

Figure 1: Small Tax Reform Perturbation

with

$$\mathbb{E}[\eta_0(x)T(Y(x))\,|x \geq w] \equiv \frac{1}{1-F(w)} \int_w^{w_1} \eta_0(x)T(Y(x))f(x)dx.$$

The first two factors (elasticity ratio and inverse of the hazard rate divided by w) are identical to the ones in Diamond's (1998) closed-economy formula, which under the maximin reduces to:[11]

$$\frac{T'(Y(w))}{1-T'(Y(w))} = \frac{\alpha(w)}{\varepsilon(w)} \frac{1-F(w)}{wf(w)}.$$

A third factor appears in the present open economy. It captures how the threat of migration affects the tax policy that each non-cooperating government finds optimal to implement. This additional migration factor plays in favour of a reduction of the marginal tax rates faced by rich people, compared to a world where tax policies would be coordinated.

To gain further insights into this new factor, let us consider a symmetric Nash equilibrium and investigate the effects of a small tax reform perturbation in a unilaterally-deviating country: the marginal tax rate $T'(Y(w))$ is uniformly increased by Δ on the interval $[Y(w)-\delta, Y(w)]$ as shown in Figure 1. Hence tax liabilities above $Y(w)$ are uniformly increased by $\Delta\ \delta$.[12] This gives rise to the following effects.

First, everyone with earnings in $[Y_i(w)-\delta, Y_i(w)]$ responds to the rise in the marginal tax rate by a *substitution effect*. Each of them reduces her taxable income by:

$$dY(w) = \frac{Y(w)}{1-T'\left(Y(w)\right)}\ \varepsilon(w)\ \Delta,$$

according to (4). This reduces the tax she pays by:

$$dT\left(Y(w)\right) = \frac{T'\left(Y(w)\right)}{1-T'\left(Y(w)\right)}\ Y(w)\ \varepsilon(w)\ \Delta.$$

[11] In Diamond (1998) $v(y;w) = v(y/w)$. Under this restriction, $\alpha(w) = 1+\epsilon(w)$.

[12] The case where marginal tax rate is uniformly decreased by $\Delta < 0$ is symmetric.

Taxpayers with income in $[Y_i(w)-\delta, Y_i(w)]$ have a skill level within the interval $[w-\delta_w, w]$ of the *skill* distribution. There is a one-to-one relationship between the interval $[Y_i(w)-\delta, Y_i(w)]$ of the *income* distribution and an interval $[w-\delta_w, w]$ of the *skill* distribution. From (5), their widths δ and δ_w are related through:

$$\delta_w = \frac{w}{Y(w)} \frac{1}{\alpha(w)} \delta.$$

The mass of taxpayers whose earnings are in the interval $[Y_i(w)-\delta, Y_i(w)]$ being $\delta_w f(w)$, the total *substitution effect* is equal to

$$dT\left(Y(w)\right)\ \delta_w\ f(w) = \frac{T'\left(Y(w)\right)}{1-T'\left(Y(w)\right)} \frac{\varepsilon(w)}{\alpha(w)} w\ f(w)\ \Delta\ \delta \tag{20}$$

Second, every individual with skill x above w faces a lump-sum increase $\delta\Delta$ in her tax liability. In the absence of migration responses, this *mechanically* increases collected taxes from those x-individuals by $f(x)\ \delta\ \Delta$. This is referred to as the "mechanical" effect in the literature. However, an additional effect takes place in the present open-economy setting. The reason is that the unilateral rise in tax liability reduces the gross utility in the deviating country, compared to its competitor. Consequently, the number of emigrants increases or the number of immigrants decreases. From (7), the number of taxpayers with skill x decreases by $\eta_0(x)\ f(x)\ \Delta\ \delta$, and thus collected taxes are reduced by $\eta_0(x)\ T\left(Y(x)\right)\ f(x)\ \Delta\ \delta$. We define the *tax level* effect $X(w)\ \Delta\ \delta$ as the sum of the mechanical and *migration* effects for all skill levels x above w:[13]

$$X(w)\ \Delta\ \delta = \int_w^{w_1} \left[1-\eta_0(x)\ T\left(Y(x)\right)\right] f(x)\ dx\ \Delta\ \delta. \tag{21}$$

The unilateral deviation we consider cannot induce any first-order effect on the tax revenues of the deviating country. This implies that the substitution (20) must be offset by the tax level (21). We thus obtain the optimal income tax formula (19) of Proposition 1.

An alternative way of writing the formula (19) given in Proposition 1 illuminates the relationship between the marginal and the average optimal tax rates. Using (8) in Equation (19), we obtain:

$$\frac{T'\left(Y\left(w\right)\right)}{1-T'\left(Y\left(w\right)\right)} = \frac{\alpha\left(w\right)}{\varepsilon\left(w\right)} \frac{1-F\left(w\right)}{wf\left(w\right)} \left[1-\mathbb{E}_f\left(\frac{T\left(Y\left(x\right)\right)}{Y\left(x\right)-T\left(Y\left(x\right)\right)}\nu_0\left(x\right)|x\geq w\right)\right]. \tag{22}$$

This alternative way of writing the optimal tax rate formula shows that the new "migration factor" makes the link between the marginal tax rate at a given w and the mean of the average tax liabilities above this w. More precisely, it corresponds to the mean of the average tax rates $\frac{T(Y)}{Y-T(Y)}$ weighted by the semi-elasticity of migration ν_0, for everyone with productivity above w. The reason is that migration choices are basically driven by average tax rates.

IV Signing Optimal Marginal Tax Rates

The overall shape of the tax schedule depends on the sign of the marginal tax rates. In a closed economy, the optimal marginal tax rates are between 0 and 1. This implies that the optimal tax function is not decreasing. Moreover, the marginal tax rates are equal

[13] From (18), the tax level effect $X(w)$ is simply the opposite of the co-state variable $q(w)$. Indeed, under quasilinear-in-consumption preferences, a lump-sum tax of one unit of consumption to all individuals of skill x above w decreases uniformly their utility levels by one unit.

to zero at the bottom if the least productive agents choose to work, there is no bunching, and the policy-maker's aversion to income inequality is finite (Mirrlees, 1971; Ebert, 1992). One of the most notorious results is that the optimal marginal tax rate is equal to zero at the top providing the distribution of skills is bounded from above (Sadka, 1976; Seade, 1977). If the tax function is continuous, the zero-tax-at-the-top result implies that the optimal marginal tax rates must be decreasing on some interval including the richest people. However, this result may be very local. Moreover, even though there is an upper bound of the *actual* distribution of skills, it is difficult for the policy-maker to know this exact value when designing the tax schedule. "The zero rate is therefore practically irrelevant" (Mirrlees, 2006, p. vii). This is why the recent literature usually considers unbounded distributions ($w_1 \to \infty$). Assuming unbounded distributions of skills, as in Mirrlees (1971), Diamond (1998) and Saez (2001) have shown that marginal tax rates may be increasing at the top, and that asymptotic marginal tax rates are usually strictly positive. In the last ten years, the derivation of top marginal tax rates has received considerable attention (see e.g., Piketty and Saez (2012)) and has contributed to the better connection between theory and empirical works.

It is easy to show that competing governments settle non-negative marginal tax rates at the bottom. Moreover, if the distribution of skills is bounded from above, the zero rate result applies.

Proposition 2 *In the symmetric Nash equilibrium,*

i) $T'(Y(w_0)) \geq 0$,

ii) $T'(Y(w_1)) = 0$ *when* $w_1 < \infty$.

Proof *i)* The result is established by contradiction. $q(w_0)$ corresponds to the derivative of the value function of the dual problem (13) with respect to $\underline{U}(w_0)$. Let us assume that $q(w_0) > 0$. Then, increasing $\underline{U}(w_0)$, i.e. social welfare, would also relax the budget constraint. A contradiction. Therefore, $q(w_0) \leq 0$. Because $v''_{yw}(y(w); w) < 0$, (14) implies $v'_y(Y(w_0); w_0) \leq 0$. It then follows from (3) that $T'(Y(w_0)) \geq 0$. *ii)* When $w_1 < \infty$, the transversality condition is $q(w_1) = 0$. This implies that the demographic factor $(1 - F(w)) / (w\ f(w))$ in Proposition 1) is zero and thus that $T'(Y(w_1)) = 0$. □

In spite of its theoretical interest, Proposition 2 is not very informative from an applied perspective. In this section, we try to further characterize the shape of the tax function by looking at the sign of the optimal marginal tax rates over the *whole* income distribution. We are interested in the characterization of the tax schedules that competing governments should implement in a second-best environment where both w and m are non-observable. However, it is instructive, as a first step, to consider the 1.5th-best situation (Jacquet, Lehmann, and Van der Linden, 2012) in which w is public information and m private information. This will allow us to make the connection between the sign of optimal tax rate and the whole profile (both level and slope) of the semi-elasticity of migration.

IV.A A Useful Benchmark: The 1.5th Best

The 1.5th-best allocation is defined as the solution to program (13), without the conditions for incentive-compatibility.[14] The first-order conditions are obtained by setting $q(w) = q'(w) = 0$ in (14) and (15), for $w > w_0$. Denoting the optimal tax function in

[14]The 1.5th best would coincide with the first best in the absence of migration and to the second best in the absence of intensive responses.

the 1.5th best by $T_i^{1.5}(Y(w);w)$, we respectively obtain:

$$1 - v'_y\left(Y_i(w);w\right) = 0, \tag{23}$$

$$T_i^{1.5}(Y_i(w);w) = \frac{1}{\eta_i(\Delta_i(w);w)}. \tag{24}$$

Because w is observable, there is no need to implement distortionary taxes. Indeed, Equations (3) and (23) imply that $\partial T_i^{1.5}\left(Y_i(w);w\right)/\partial y = 0$. The government can therefore modify tax levels without inducing any substitution effect similar to (20). Consequently, the tax level effect (21) vanishes in the 1.5th best. Therefore, at each skill level, the mechanical effect and migration response effect of a higher tax liability on tax revenues cancel out, which leads to (24). The tax liability $T_i^{1.5}(Y_i(w))$ required from the residents with skill $w > w_0$ is equal to the inverse of their semi-elasticity of migration $\eta_i(\Delta_i(w), w)$. The least productive individuals receive a transfer determined by the budget constraint. Therefore, the optimal tax function is discontinuous at $w = w_0$, as illustrated in Figures 2 − 4. Using (8), we can alternatively express the best response of country i's policy-maker using the elasticity of migration instead of the semi-elasticity. We obtain:

$$\frac{T_i^{1.5}(Y_i)}{Y_i - T_i^{1.5}(Y_i)} = \frac{1}{\nu(\Delta_i;w)}.$$

The average tax liability required in each country from the residents with skill w is therefore the inverse of the elasticity of migration. This is the formula derived by Mirrlees (1982) who focuses on extensive responses.

Combining best responses, we easily obtain the following characterization for the symmetric Nash equilibrium. We state it as a proposition because it provides a benchmark to sign second-best optimal marginal tax rates.

Proposition 3 *In the 1.5th best, the symmetric Nash equilibrium allocation can be decentralized by a tax function such that*

$$T^{1.5}(Y(w)) = \frac{1}{\eta_0(w)}. \tag{25}$$

In the 1.5-best setting, the optimal tax liability is increasing in skill when the semi-elasticity of migration $\eta_0(.)$ is decreasing. Symmetrically, the tax liability is decreasing when $\eta_0(.)$ is increasing. Knowing how the semi-elasticity of migration varies with skills is therefore key to determine the profile of the optimal tax schedule. The next subsections will show that the profile of the semi-elasticity of migration will also play an essential part in the second best.

Three natural benchmarks come to mind when thinking about migration. First, the costs of migration may be decreasing in w. This seems to be supported by the empirical evidence that highly skilled are more likely to emigrate than low skilled (Docquier and Marfouk, 2006). This suggests that the semi-elasticity of migration may be increasing in skills. A special case is investigated in Simula and Trannoy (2010) and Simula and Trannoy (2011). These articles assume that there is a unique cost of migration at each skill level. They consider the best response in a given country to a less redistributive tax policy abroad. Hence, there is a threshold skill level below which individuals do not find it profitable to emigrate and above which a no-migration constraint is binding. In other words, the semi-elasticity of migration is zero below this threshold and infinite above. Second, the costs of migration may be independent of w as in Blumkin, Sadka, and Shem-Tov (2012) and Morelli, Yang, and Ye (2012). This makes sense, in particular, if most relocation costs are material (moving costs, flight tickets, etc.).[15] Third, one might want to consider a constant elasticity of migration, as in Brewer, Saez, and Shephard (2010) and Piketty and Saez (2012). In this case, the semi-elasticity must be non-increasing: if everyone receives one extra unit of consumption in country i, then the relative increase in the number of taxpayers becomes smaller for more skilled individuals.

IV.B From the 1.5th Best to the 2nd Best

From Equation (19) of Proposition 1 and (21), we know that the second-best optimal allocation can be decentralized by a tax function satisfying:

$$\frac{T'(Y(w))}{1-T'(Y(w))} = \frac{\alpha(w)}{\varepsilon(w)} \, \frac{X(w)}{wf(w)} \tag{26}$$

Consequently, $T'(Y(w))$ has the same sign as the tax level effect $X(w)$. From (21) and Proposition 3, we see that

$$X(w) = \int_w^{w_1} \left[T^{1.5}(Y(x)) - T\left(Y(x)\right)\right] \eta_0(x) f(x) \; dx. \tag{27}$$

The tax level effect is thus the weighted sum of the difference between optimal tax liabilities in the 1.5th best and in the second best. The weights are given by the product of the semi-elasticity of migration and the skill density. Intuitively, in 1.5th best, the mechanical and migration effects of a change in tax liabilities cancel out. Hence, the 1.5th-best tax schedule defines a *target* for the policy-maker in the second best, where distortions along the intensive margin have also to be minimized. The second-best solution proceeds from the reconciliation of two underlying forces: being as close as possible to the 1.5th-best tax liability to limit the distortions stemming from the migration responses whilst being as flat at possible to mitigate the distortions coming from the intensive margin (cf. Jacquet, Lehmann, and Van der Linden (2012)).

Proposition 2 is thus equivalent to (i) $X(w_0) \geq 0$ and (ii) $X(w_1) = 0$ when $w_1 < \infty$. From (27), the derivative of $X(w)$ is

$$X'(w) = \left[T(Y(w)) - T^{1.5}(Y(w))\right] \; \eta_0(w) \; f(w). \tag{28}$$

Therefore, the tax level effect $X(w)$ is increasing (decreasing) when the tax paid in the second best is larger (lower) than the target $T^{1.5}(Y(w)) = 1/\eta_0(w)$. The following lemma will be useful to further characterize the optimal second-best tax schedule:

Lemma 1 *In the second best, if there exists a skill level $\hat{w} < w_1$ such that:*

i) $T'(Y(\hat{w})) \leq 0$ *and* $T(Y(\hat{w})) > 1/\eta_0(\hat{w})$, *then* $T'(Y(w)) < 0$ *and* $X(w) < X(\hat{w})$ *when* w *goes down, as long as* $T(Y(w)) > 1/\eta_0(w)$.

ii) $T'(Y(\hat{w})) \leq 0$ *and* $T(Y(\hat{w})) < 1/\eta_0(\hat{w})$, *then* $T'(Y(w)) < 0$ *and* $X(w) < X(\hat{w})$ *when* w *goes up, as long as* $T(Y(w)) < 1/\eta_0(w)$.

iii) $T'(Y(\hat{w})) \geq 0$ *and* $T(Y(\hat{w})) > 1/\eta_0(\hat{w})$, *then* $T'(Y(w)) > 0$ *and* $X(w) > X(\hat{w})$ *when* w *goes up, as long as* $T(Y(w)) > 1/\eta_0(w)$.

[15] Morelli, Yang, and Ye (2012) compare a unified nonlinear optimal taxation with the equilibrium taxation that would be chosen by two competing tax authorities if the same economy were divided into two States. In their conclusion, they discuss the possible implications of modifying this independence assumption and consider that allowing for a negative correlation might be more reasonable.

Proof

i) Consider $w^* < \hat{w}$ such that for all $w \in [w^*, \hat{w}]$, one has $T(Y(w)) > 1/\eta_0(w)$. Then, for any $w \in [w^*, \hat{w}]$, $X'(w) > 0$ (using (28)). So, $X(w) < X(\hat{w}) \leq 0$. Equation (26) then implies $T'(Y(w)) < T'(Y(\hat{w})) \leq 0$.

ii) Consider $w^* > \hat{w}$ such that for all $w \in [\hat{w}, w^*]$, one has $T(Y(w)) < 1/\eta_0(w)$. Then, for any $w \in [w^*, \hat{w}]$, $X'(w) < 0$ (using (28)). So, $X(w) < X(\hat{w}) \leq 0$. Equation (26) then implies $T'(Y(w)) < T'(Y(\hat{w})) \leq 0$.

iii) Consider $w^* < \hat{w}$ such that for all $w \in [w^*, \hat{w}]$, one has $T(Y(w)) > 1/\eta_0(w)$. Then, for any $w \in [w^*, \hat{w}]$, $X'(w) > 0$ (using (28)). So, $X(w) > X(\hat{w}) \leq 0$. Equation (26) then implies $T'(Y(w)) > T'(Y(\hat{w})) \geq 0$.

□

IV.B.1 Decreasing Semi-Elasticity of Migration

We first consider the situation when the semi-elasticity of migration is increasing. This for example occurs for a constant elasticity of migration. We show that marginal tax rates are positive in the second best. The proof is based on Figure 2. The dashed curve represents the 1.5th-best tax liability as a function of skills. Because the semi-elasticity is positive and decreasing, the latter is positive and increasing for $w > w_0$. The collected taxes are redistributed to the least productive agents. This implies that the 1.5th-best tax schedule is discontinuous at w_0.

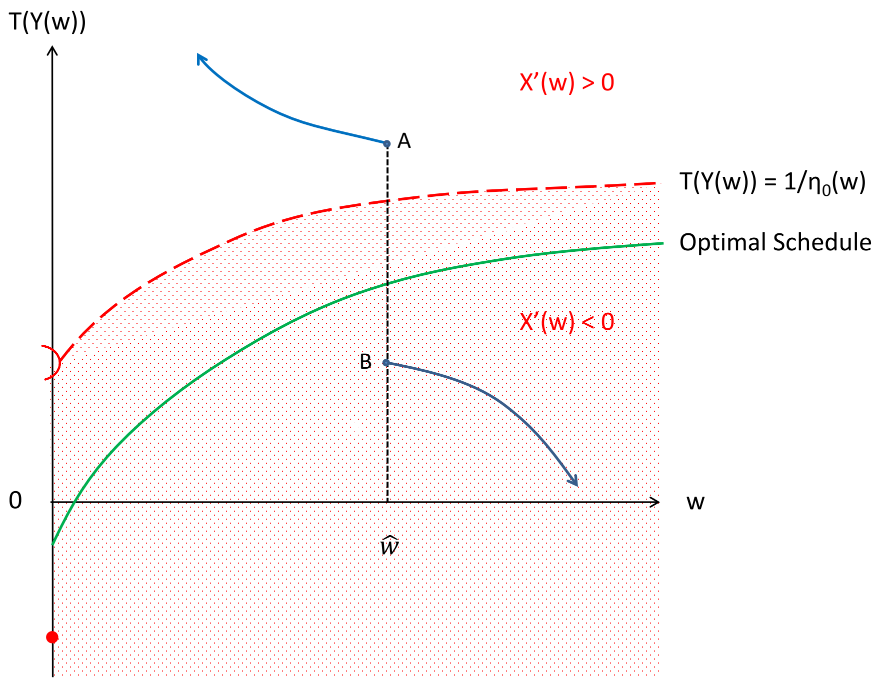

Figure 2: Decreasing Semi-Elasticity of Migration

Assume now by contradiction that there exists a skill level $\hat{w}$ where marginal tax rate is not positive. Then, there are two cases. First, one may have $T(Y(\hat{w})) \geq 1/\eta(\hat{w})$, as in point A. Then, as skill goes down (see the arrow), Lemma 1 *i*) ensures that the tax liability increases and remains above the 1.5th-best tax schedule. Consequently $X(w)$ remains negative, which violates the transversality condition $X(w_0) \geq 0$. Alternatively, one may have $T(Y(\hat{w})) \leq 1/\eta(\hat{w})$, as in point B. Then, as skill goes up (see the arrow), Lemma 1 *ii*) ensures that the tax liability decreases and remains below the 1.5th-best tax schedule. Consequently $X(w)$ remains negative, which violates the transversality condition $\lim_{w \mapsto w_1} X(w) = 0$.[16]

Proposition 4 *If $\eta_0'(\cdot) < 0$, marginal tax rates are positive, i.e. $T'(Y(w)) > 0$, except at w_1.*

Intuitively, the second-best optimal tax policy tries to replicate the 1.5th best, but in a flatter way, to minimize the distortions along the intensive margin. This is illustrated by the solid curve in Figure 2. When the semi-elasticity of migration is decreasing, richer individuals pay higher taxes.

Brewer, Saez, and Shephard (2010) and Piketty and Saez (2012) look at the asymptotic marginal tax rate. They assume that the elasticity of migration is constant ($\nu_0(x) = \nu_0$). From Equation (8), a constant elasticity of migration is a special case of a decreasing semi-elasticity, because $C(w)$ must be increasing in the second best. They also assume that the elasticities $\varepsilon(w)$, $\alpha(w)$ and $\nu(w)$ converge asymptotically to $\varepsilon(\infty)$, $\alpha(\infty)$ and $\nu(\infty)$ respectively. They finally assume that the distribution of skills is Pareto in its upper part, so that $k = (wf(w))/(\alpha(\infty)(1 - F(w)))$. Making skill w tends to infinity in the optimal tax formula (22), we retrieve the optimal asymptotic marginal tax rate of Brewer, Saez, and Shephard (2010) and Piketty and Saez (2012):[17]

$$T'(Y(\infty)) = \frac{1}{1 + k\varepsilon + \nu_0}.$$

The asymptotic marginal tax rate is thus strictly positive. For example, if $k = 1.5$, $\varepsilon = 0.25$ and $\nu_\infty = 0.25$, we obtain $T'(Y(\infty)) = 61.5\%$ instead of 72.7% in the absence of migration responses.

IV.B.2 Constant Semi-Elasticity of Migration

We now consider the case where the semi-elasticity of migration is constant, so $\eta_0(w) \equiv \eta_0$. This is for example the case when the distributions of w and m are independent, as in Blumkin, Sadka, and Shem-Tov (2012) and Morelli, Yang, and Ye (2012). This situation is illustrated in Figure 3.

a) Let us assume that $T(Y(\hat{w})) \geq 1/\eta_0$, as in point A, and $T'(Y(\hat{w})) > 0$. We have $X(\hat{w}) > 0$ by Lemma 1 *i*). Moreover, when w goes down , we must have $X(w) > 0$ for any $w > \hat{w}$. This is incompatible with the transversality condition $\lim_{w \mapsto w_1} X(w) = 0$ (Lemma 1 *iii*)), as shown by the arrow.
b) Let us assume that $T(Y(\hat{w})) \geq 1/\eta_0$, as in point A and $T'(Y(\hat{w})) < 0$. We have $X(\hat{w}) < 0$ by Lemma 1 *i*). Moreover, when w goes down, $X(w)$ increases by Lemma 1 *ii*). So, for $w < \hat{w}$, $X(w)$ remains negative and non-decreasing, which contradicts $X(w_0)$.

[16] A third possible case is: $T(Y(\hat{w})) = 1/\eta(\hat{w})$. Then, because $\eta(\cdot)$ is decreasing, the situation described in A occurs on the left and the situation described in B occurs on the right.

[17] By L'Hôpital's rule, $\lim_{w \mapsto w_1} \frac{T(Y(w))}{Y(w) - T(Y(w))} = \lim_{w \mapsto w_1} \frac{T'(Y(w))}{1 - T'(Y(w))}$.

Because $X(w)$ is increasing above the dashed line and constant along it, the tax liability reaches its maximum at $w = w_0$. This clearly contradicts the maximin social objective.[18]

c) Now let us assume that $T(Y(\hat{w}))$ is just equal to $1/\eta_0$. Because $X(w)$ is constant along the dotted line, it must be that $T(Y(w)) = 1/\eta_0$ for all $w_0 < w \leq w_1$. Contrary to the 1.5th best, the second-best solution cannot involve a discontinuity at w_0. Indeed, if the tax receipts collected on all individuals with $w > w_0$ were given to the least skilled, then the individuals immediately to the right of w_0 would have an incentive to mimic the latter. Hence, $T(Y(w_0)) = 1/\eta_0$. We thus obtain a second-best solution where the tax function is constant for all $w_0 \leq w \leq w_1$. Given that the tax policy is strictly redistributive, this contradicts the maximin objective.

d) Therefore, we are left with only one possibility: $T(Y(\hat{w})) < 1/\eta_0$. Because $X(w)$ is decreasing in that part of the space, $X(\hat{w})$ is strictly positive. Hence, the optimal marginal tax rate is strictly positive for $w < w_1$ with $\lim_{w \mapsto w_1} T'(Y(w)) = 0$.

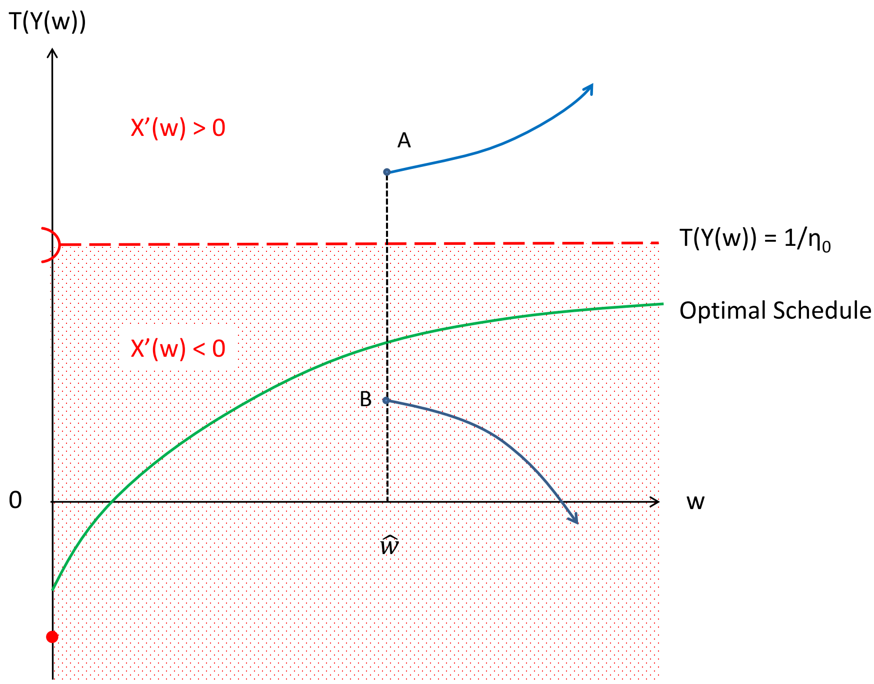

Figure 3: Constant Semi-Elasticity of Migration

Proposition 5 *Assume that $\eta_0'(\cdot) = 0$. Then, $T'(Y(w)) > 0$ for $Y(w) < Y(w_1)$. Moreover,* $\lim_{w \mapsto w_1} T'(Y(w)) = 0$.

IV.B.3 Increasing Semi-Elasticity of Migration

We last investigate the case when the semi-elasticity of migration is increasing in skills. This is illustrated in Figure 4.

Assume by contradiction that there exists a skill level $\hat{w}$ where we both have $T'(Y(\hat{w})) \geq 0$ and $T(Y(\hat{w}) > 1/\eta_0(w)$. Then, when the skill level goes up, Lemma 1 *iii*) ensures that the

[18] For example, a laissez-faire policy would do better.

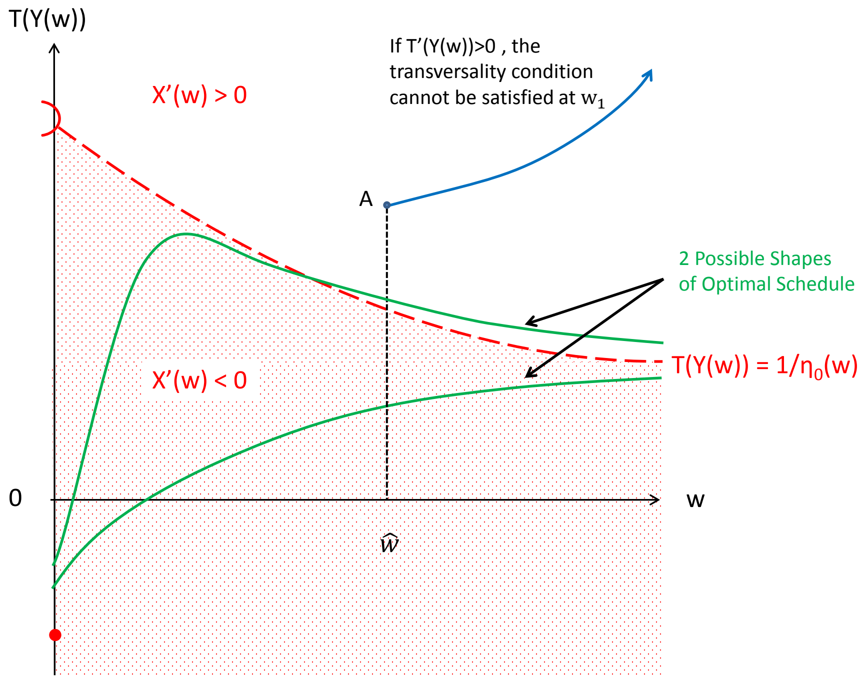

Figure 4: Increasing Semi-Elasticity of Migration

tax liability increases and remains above the 1.5th-best tax schedule. This is shown by the arrow starting from point A in the Figure. We see that $X(w)$ remains positive, which violates the transversality condition $\lim_{w \mapsto w_1} X(w) = 0$. Consequently, only two cases are possible:

a) First, the second-best tax schedule remains always below the 1.5th best one, in which case the tax level effect $X(w)$ is decreasing according to (28). Therefore the tax level effect remains negative everywhere; otherwise the transversality conditions would be violated.

b) Second, the second-best tax schedule is above the 1.5th best tax target for some $\hat{w}$. This implies $X(\hat{w}) > 0$, which is only compatible with $X(\hat{w}) \leq 0$, according to Lemma 1 *iii*). In that case, the tax schedule must be first increasing and then decreasing, approaching $1/\eta_0$ from above. In summary:

Proposition 6 *Assume that $\eta_0'(w) > 0$, then:*

(i) either $T'(Y(w)) > 0$ for every $Y(w) \in (Y(w_0), Y(w_1))$,

(ii) or there exists a threshold $\hat{w} \in (w_0, w_1)$ under which $T'(Y(w)) > 0$ and above which $T'(Y(w)) < 0$.

When we are in the second case, the optimal average tax rate and the optimal tax function are strictly decreasing for the wealthiest part of the population. Therefore, progressivity of the optimal tax schedule does not only collapse because of tax competition; the tax liability itself becomes strictly decreasing. This means that there are "middle-skilled" individuals who pay higher taxes than top-income earners. This situation can be regarded as a "curse of the middle-skilled". For example, this curse occurs when the semi-elasticity tends to infinity. Indeed, the first-and-a-half tax liability will then tend to zero. If the second-best optimal tax function approaches $1/\eta_0(w)$ from below, the tax function must be increasing, reaching

Semi-Elasticity	Elasticity	Shape of Tax Function	Top Tax Rates
Decreasing	–	Increasing	$T'(Y(w_1)) \geq 0$
Constant	Non-Decreasing	Increasing	$T'(Y(w_1)) = 0$
Increasing	Increasing	a) Increasing	$T'(Y(w_1)) \geq 0$
		b) Hump-Shaped	$T'(Y(w_1)) \leq 0$

Table 1: Main Features of the Equilibrium Tax Schedule

its maximum at zero. The budget constraint of the policy-maker must thus be violated. Consequently, we are left with case (ii) in Proposition 6.

Proposition 7 *Assume that* $\eta_0'(w) > 0$ *and* $\lim_{w_1 \to \infty} \eta_0(w) = \infty$*, then there exists a threshold* $\hat{w} \in (w_0, w_1)$ *under which* $T'(Y(w)) > 0$ *and above which* $T'(Y(w)) < 0$.

The main features of the equilibrium tax schedule are summarized in Table 1. It clearly appears that the level of the (semi-)elasticity of migration is *not* a sufficient statistic for the characterization of optimal marginal tax rates, *even* at the top. *Both* the level and the slope of the (semi-)elasticity are required.

V Simulations

This section provides numerical simulations of the equilibrium optimal tax schedule that competing policy-makers should implement. One of our objectives is to emphasize the part played by the slope of the semi-elaticity of migration. In particular, we will show that the marginal tax rates faced by rich individuals may be quite sensitive to the overall shape of the semi-elasticity.

We use the distribution of weakly earnings for singles without children in 2007 (CPS data) to recover the skill distribution $f(w)$. We compute annual earnings Y and then proceed by inversion to find the value of w, assuming a linear tax function $T(Y) = 0.4058Y + 4036$ that approximates the US tax schedule in 2007 (see OECD database). Following Diamond (1998) and Saez (2001), we correct for top coding by extending the obtained estimation with a Pareto distribution of coefficient 1.59. The dis-utility of effort is given by $v(y; w) = (y/w)^{1+1/\epsilon}$. This specification implies a constant elasticity of gross earnings with respect to the retention rate ϵ, as in Diamond (1998) and Saez (2001). In a recent survey, Saez, Slemrod, and Giertz (2012) conclude that "the best available estimates range from 0.12 to 0.4" in the United States. We use a central value, $\epsilon = 0.25$. Public expenditures E are kept at their initial level \$18,157, which corresponds to 33.2% of the total gross earnings of single without children. Our calibration provides a very good approximation of the top of the income distribution as described by Alvaredo, Atkinson, Piketty, and Saez (2013). In the absence of migration responses, we find that the top 0.1%, top 1%, top 5% and top 10% of the population respectively get 6.5%, 18.2%, 34.7% and 45.4% of total income. The corresponding numbers in the World Top Income Database are 8.2%, 18.3%, 33.8% and 45.7%.

The semi-elasticity of migration is a key parameter in our computations. Even though the role of income taxation on migration behaviour has been extensively discussed in the theoretical literature since Tiebout's (1956) seminal contribution, there are still very few studies on the income taxation's effects on migration. Kleven, Landais, and Saez (2010) study tax induced mobility in Europe of football players and find substantial mobility elasticities. More specifically, the mobility of domestic players with respect to domestic tax rate is rather small around 0.15, but the mobility of foreign players is much larger, around 1. Kleven, Landais, Saez, and Schultz (2001) confirm that these results apply to the broader market of highly skilled foreign workers and not only to football players. They find an elasticity above 1 in Denmark. In a given country, the number of foreigners at the stop is relatively small. Hence, these findings would translate into a global elasticity at the top of at most 0.25 for most countries (see Piketty and Saez (2012)).

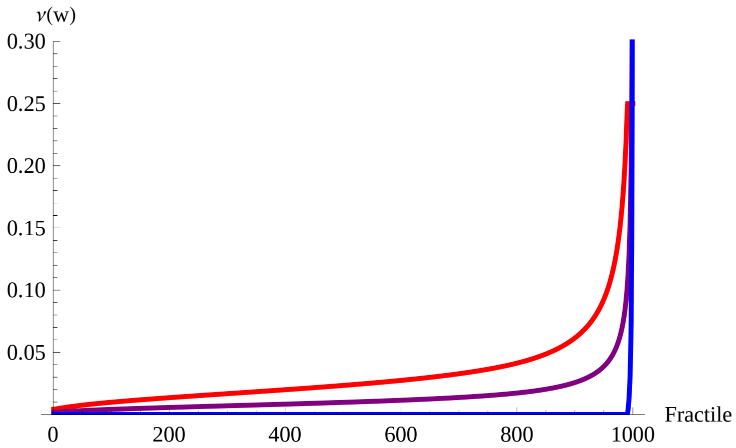

Figure 5: Elasticity of Migration by Fractile of the Actual Earnings Distribution. Case 1 (Red), Case 2 (Purple) and Case 3 (Blue)

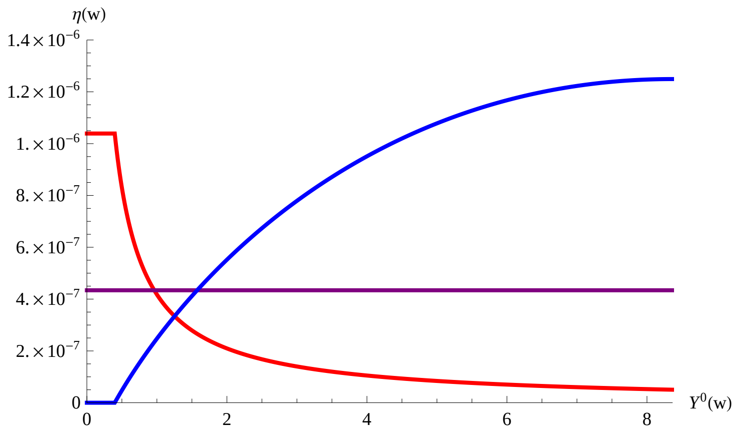

Figure 6: Semi-Elasticity of Migration as a Function of Actual Gross Earnings in Millions of US$. Case 1 (Red), Case 2 (Purple) and Case 3 (Blue)

As far as we know, there are no empirical studies regarding the possible shape of the elasticity or semi-elasticity of migration. We therefore investigate three possible scenarios. In each of them, the average elasticity in the actual economy top 1% of the population is equal to 0.25, as shown in Figure 5, where the population is divided into 1000 fractiles, based on individual earnings $Y^0(w)$ in the actual economy. The average elasticity in the population is much lower: 0.025 in the first one, 0.01 in the second one and 0.003 in the third one. In the first scenario, the semi-elasticity is constant up to the top centile and then decreasing in such a way that the elasticity of migration is constant within the top centile. This is shown in Figure 6. In the second scenario, the semi-elasticity is constant throughout the whole skill distribution. In the third scenario, the semi-elasticity is zero up to the top centile and then increasing.

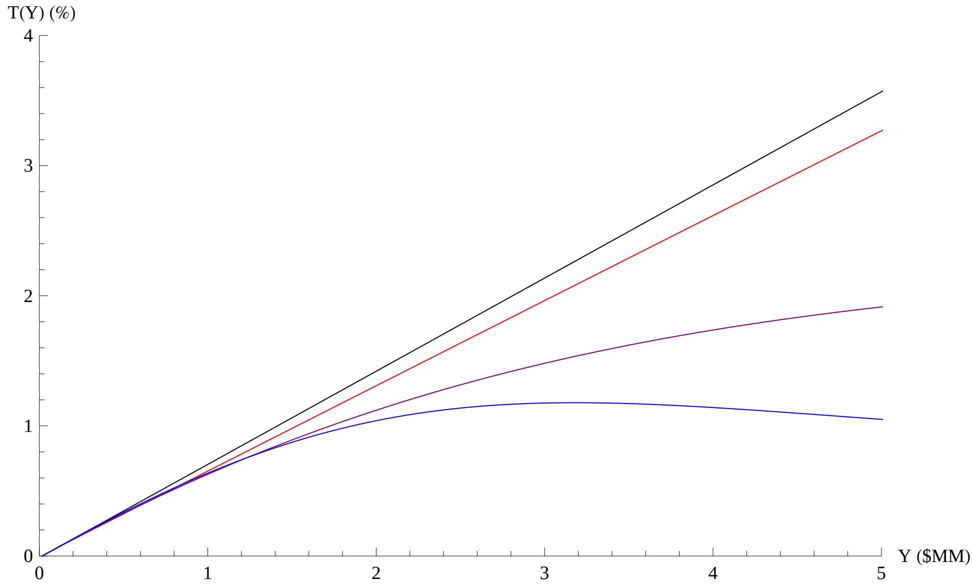

Figure 7: Optimal Tax Liabilities. Autarky (Black), Case 1 (Red), Case 2 (Purple) and Case 3 (Blue).

The optimal equilibrium tax liabilities are shown in Figure 7. The x-axis represents gross earnings and the x-axis the total tax paid, both expressed in millions of US dollars. In addition to the three scenarios presented above, we added the tax liabilities that would be chosen in a closed economy or in the presence of tax coordination (cf. black curve). We observe that the threat of migration implies a non-negligible decrease in the total taxes paid by top income earners. Even though the average elasticity of migration is the same for the top 1% of income earners in the three cases, we observe significant differences due to variations in the shape of the semi-elasticity of migration. In the first case, the tax function is quite flat for high income earners and remains close to the closed-economy benchmark. In the second case, the tax function is more concave for large incomes, but remains increasing. In the third case, the tax function becomes decreasing around $Y = \$3.2$ millions. The richest people are not those paying the largest taxes.

The effect of fiscal competition on tax progressivity is emphasized in Figure 8, which shows the average tax rate. The tax policy is progressive in case 1, but strongly regressive in the two other cases. The average tax rate for rich people ($5 millions of annual earnings) is about 65% in case 1, 39% in case 2 and 21% in case 3.

Figure 9 casts light on the differences in the optimal marginal tax rates. What we see is that differences in the slope of the semi-elasticity of migration may translate into large differences in the top marginal tax rates. Consequently, our numerical results put the stress on the need for empirical studies on the slope of the semi-elasticity of migration, in addition to its level.

VI Conclusion

What is the best redistributive tax policy, in a given country, when individuals have the possibility to exploit their outside options and threaten to move abroad to avoid high tax rates?

Because of the threat of migration, competing policy-makers design tax schedules whose qualitative features may strongly differ from those that would be obtained in a closed econ-

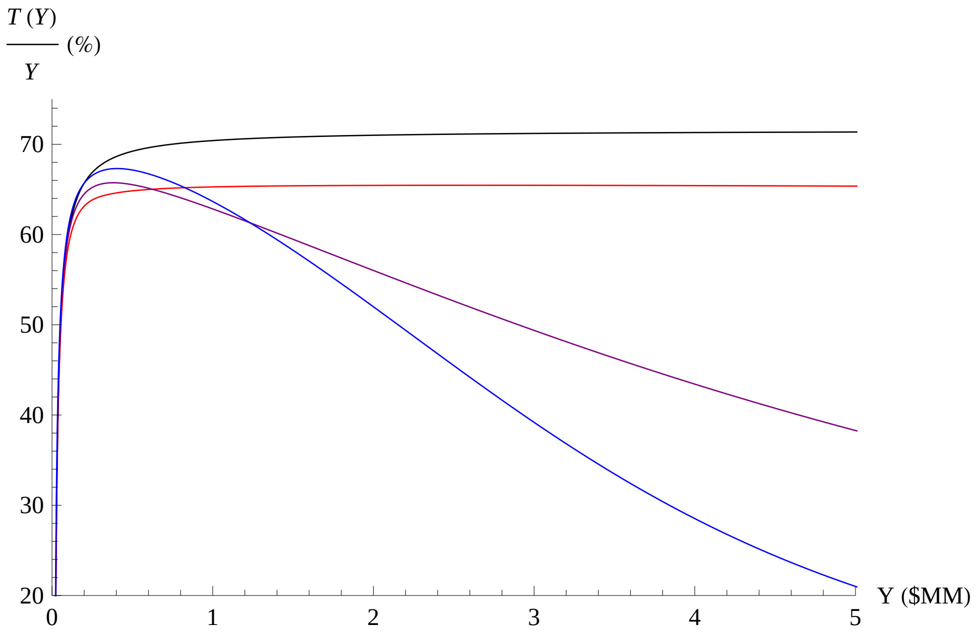

Figure 8: Optimal Average Tax Rates. Autarky (Black), Case 1 (Red), Case 2 (Purple) and Case 3 (Blue).

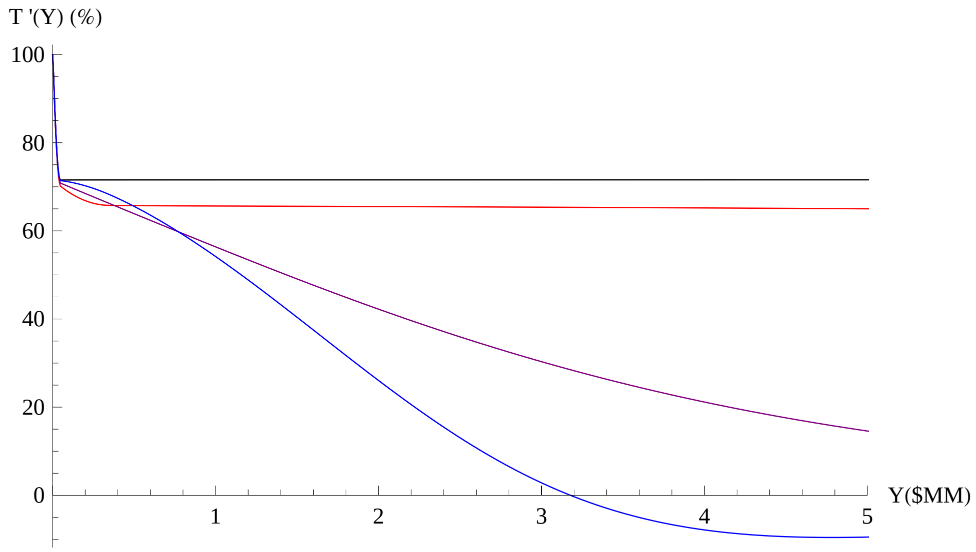

Figure 9: Optimal Marginal Tax Rates. Autarky (Black), Case 1 (Red), Case 2 (Purple) and Case 3 (Blue).

omy. A small tax reform perturbation around the equilibrium has a new migration effect, which 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.

In the presence of tax competition, the elasticity of migration is not a sufficient statistics to summarize the responses along the extensive margin. Both the level and the slope of the (semi-)elasticity are required. Numerical simulations show that the optimal tax function is quite sensitive to variations in the slope of the elasticity. Therefore, there is a need of empirical studies regarding this second statistics.

References

Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez (2013): "The World Top Incomes Database, http://topincomes.g-mond.parisschoolofeconomics.eu/," .

Atkinson, A. (1990): "Public Economics and the Economic Public," *European Economic Review*, 34(2-3), 225–248.

Atkinson, A., and J. Stiglitz (1980): *Lectures on Public Economics*. McGraw-Hill Inc.

Bhagwati, J. (1976): *The Brain Drain and Taxation: Theory and Empirical Analysis*. North-Holland, Amsterdam.

Bhagwati, J., and M. Partington (1976): *Taxing the Brain Drain: A Proposal*. North Holland, Amsterdam.

Bierbrauer, F., B. C., and J. A. Weymark (2011): "Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility," CESifo Working Paper 3329.

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

Blumkin, T., E. Sadka, and Y. Shem-Tov (2012): "International Tax Competition: Zero Tax Rate at the Top Re-established," Discussion Paper 3820, CESifo WP.

Boadway, R., and L. Jacquet (2008): "Optimal Marginal and Average Income Taxation under Maximin," *Journal of Economic Theory*, 143, 425–441.

Boadway, R., and P. Pestieau (2007): "Optimal Income Taxation with Tagging," *Annales d'Economie et de Statistiques*.

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

Brewer, M., E. Saez, and A. Shephard (2010): *Dimensions of Tax Design: The Mirrlees Review*chap. Means-testing and Tax Rates on Earnings, pp. 90–173. Oxford University Press.

d'Autume, A. (2000): "L'imposition optimale du revenu: une application au cas français," *Revue Française d'Economie*, 3, 3–63.

Diamond, P. (1998): "Optimal Income Taxation: An Example with U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88(1), 83–95.

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

Ebert, U. (1992): "A Reexamination of the Optimal Nonlinear Income Tax," *Journal of Public Economics*, 49(1), 47–73.

Gruber, J., and E. Saez (2002): "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 84, 1–32.

Guesnerie, R. (1995): *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, Cambridge.

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.

Hammond, P. J. (1979): "Straightforward Individual Incentive Compatibility in Large Economies," *The Review of Economic Studies*, 46(2), pp. 263–282.

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

Jacquet, L., E. Lehmann, and B. Van der Linden (2012): "Signing Distortions in Optimal Tax and other Adverse Selection Problems with Random Participation," (3766).

Kleven, H., C. Landais, and E. Saez (2010): "Taxation and International Migration of Superstars: Evidence from the European Football Market," NBER Working Papers 16545, NBER, Forthcoming in American Economic Review.

Kleven, H., C. Landais, E. Saez, and E. Schultz (2001): "Taxation and International Migration of Top Earners: Evidence from the Foreigner Tax Scheme in Denmark," Preliminary Working Paper.

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

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

Lipatov, V., and A. Weichenrieder (2012): "Optimal income taxation with tax competition," (1207).

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

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

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

——— (2006): *Welfare, Incentives, and Taxation*. Oxford University Press.

Morelli, M., H. Yang, and L. Ye (2012): "Competitive Nonlinear Taxation and Constitutional Choice," *American Economic Journal: Microeconomics*, 4, American Economic Journal: Microeconomics.

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

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

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

Piketty, T., and E. Saez (2012): "Optimal Labor Income Taxation," Discussion Paper WP No, 18521, NBER.

Rochet, J.-C., and A. Stole, L (2002): "Nonlinear Pricing with Random Participation," *The Review of Economic Studies*, 69(1), pp. 277–311.

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.

——— (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073.

Saez, E., J. Slemrod, and S. Giertz (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50, 350, forthcoming in the Journal of Economic Literature.

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

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

Simula, L., and A. Trannoy (2011): "Shall we keep the highly skilled at home? The optimal income tax perspective," *Social Choice and Welfare*, pp. 1–32.

Sjaastad, L. (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy*, 70, 80–93.

Stiglitz, J. (1982): "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.

Wilson, J. D. (1982): "Optimal income taxation and migration: A world welfare point of view," *Journal of Public Economics*, 18(3), 381 – 397.