

Optimal Income Taxation between Competing Governments

Etienne LEHMANN*

ERMES (TEPP) University Panthéon-Assas and CREST

Laurent SIMULA[†]

Alain TRANNOY[‡]

Uppsala University

AMSE

February 15, 2013

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.

*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: University of Uppsala and Uppsala Center for Fiscal Studies. Box 513, SE-751 20 Uppsala, Sweden. Email: laurent.simula@nek.uu.se. Laurent Simula is also research fellow at IDEP and CESifo.

[‡]Address: Aix-Marseille School of Economics, IDEP, GREQAM and EHESS. 2, rue de la Charité, 13002 Marseille, France. Email: alain.trannoy@univmed.fr

1. 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

¹This 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).

²The 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.

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 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

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

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

⁵The 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.

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 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, Brett, and Weymark, 2011; Morelli, Yang, and Ye, 20012). Bierbrauer, Brett, 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 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.

2. Model

We consider an economy consisting of two symmetric countries, indexed by $i = A, B$. There is a mass 2 of workers. The technology in each country exhibits constant returns to scale, so that workers are paid up to their productivity level. Each worker is characterized by three characteristics: his/her native country $i \in \{A, B\}$, his/her productivity (or skill) $w \in [w_0, w_1]$, and the migration cost $m \in \mathbb{R}^+$ he/she supports if he/she decides to live abroad. Note that w_1 may be either finite or infinite. In addition, the empirical evidence that some people are immobile is captured by the possibility of infinitely large migration costs.⁶ This 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. We will see later on that some assumptions on this relationship, which might look innocuous at first sight, actually have significant consequences for the whole optimal tax profile, even for top income earners. Alternatively, the cost of migration can be regarded as the cost incurred by an individual to settle down his/her source of income abroad (e.g., tax advisor fees).

The joint distribution of skills w and migration cost 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. Therefore, $G(m|w) f(w)$ is the density of individuals of skill w whose migration cost is not larger than m .

2.1. 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 is better suited to the analysis of self-workers' and entrepreneurs' decisions. 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 he/she decides to migrate, and to zero otherwise. Individual

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

preferences are described by the quasi-linear utility function:

$$c - v(y; w) - 1 \times m.$$

The quasi-linearity in consumption implies that there is no income effect on taxable income. Even though there is much less empirical evidence on the magnitude of the income effects in the reported income literature 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 will choose the same gross income/consumption bundle, irrespective of their citizenship.

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

2.1.1. 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. 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 citizenship. Therefore, a worker of skill w , who has chosen to work in country i , solves:

$$U_i(w) \equiv \max_y y - T_i(y) - v(y; w). \quad (1)$$

Note that the choice of y is independent of m . We call $U_i(w)$ the *gross* utility level 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 (1) and $C_i(w) = Y_i(w) - T(Y_i(w))$ the consumption level of a worker of

skill w in country i .⁷ The first-order condition can be written as:

$$1 - T'_i(Y_i(w)) = v'_y(Y_i(w); w). \quad (2)$$

Increasing effort to get one extra unit of pre-tax income increases consumption by $1 - T'_i(Y_i(w))$ units, but reduces utility by $v'_y(Y_i(w); w)$ units. Differentiating (2), we obtain the elasticity of gross earnings with respect to the retention rate $1 - T'_i$ and skill level w :

$$\varepsilon_i(y; w) \equiv \frac{1 - T'_i}{Y} \frac{\partial Y}{\partial (1 - T'_i)} = \frac{v'_y(y; w)}{y v''_{yy}(y; w)}, \quad (3)$$

$$\alpha_i(y; w) \equiv \frac{w}{Y} \frac{\partial Y}{\partial w} = - \frac{w v''_{yw}(y; w)}{y v''_{yy}(y; w)}. \quad (4)$$

2.1.2. 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 he/she stays in A and utility $U_B(w) - m$ if he/she relocates to B . He/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 $G(U_B(w) - U_A(w)|w) f(w)$ and the mass of agents staying in their native country by $(1 - G(U_B(w) - U_A(w)|w)) f(w)$.

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

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 is given by a function φ of the difference $\Delta = U_A(w) - U_B(w)$ in gross utility levels, with $\varphi(\cdot; w)$ defined as:

$$\varphi(\Delta; w) \equiv \underbrace{(1 - G(-\Delta|w)) f(w)}_{\text{Natives}} + \underbrace{G(\Delta|w) f(w)}_{\text{Immigrants}}. \quad (5)$$

Let us consider the agents with skill w . When the utility is larger in B (i.e., $\Delta < 0$), a mass $(1 - G(-\Delta|w)) f(w) > 0$ of natives from country A moves to country B . For natives of country B , a move always corresponds to a decrease in utility. Hence, there

⁷If (1) admits more than one solution, we make the tie-breaking assumption that individuals choose the one preferred by the government.

is no migration from country B to country A , which is captured by $G(\Delta|w) = 0$. Hence, $\varphi(\Delta; w) \equiv (1 - G(-\Delta|w))f(w)$. When the utility is larger in A (i.e., $\Delta > 0$), everyone born in A stays in this country, while a mass $G(\Delta|w)f(w) > 0$ of foreigners moves from country B to country A . Then, $\varphi(\Delta; w) \equiv (1 + G(\Delta|w))f(w)$. When the gross utility levels are the same in both countries ($\Delta = 0$), the mass of residents is given by the initial density $\varphi(0; w) = f(w)$. The function $\varphi(\cdot; w)$ is continuously differentiable, with derivative $\partial\varphi(\cdot; w)/\partial\Delta = g(|\Delta||w)f(w)$. It is increasing in the difference Δ in the gross utility levels. Indeed, a rise in the gross utility in country A , in comparison to that offered in country B , reduces out-migration of natives and increases in-migration of foreigners in country A . By symmetry, the mass of residents of skill w in country B is given by $\varphi(U_B(w) - U_A(w); w) \equiv \varphi(-\Delta; w)$.

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:

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

It corresponds to the percentage change in the density of taxpayers with skill w when their consumption $C_i(w)$ is increased at the margin. The elasticity of migration is defined as:

$$\nu_i(\Delta_i; w) \equiv \frac{C_i(w)}{\varphi(\Delta_i; w)} \frac{\partial\varphi(\Delta_i; w)}{\partial C_i(w)} = C_i(w) \times \eta(\Delta_i, w).$$

We will see in the next sections that the semi-elasticity plays a direct part in the characterisation of a 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 of the top marginal tax rates is directly connected with the semi-elasticity and clarify under which testable conditions the latter are positive, equal to zero, or even negative.

2.2. Tax Policies

In each country $i = A, B$, a benevolent policy-maker designs the tax system so as to maximise the welfare of the worst-off individuals. We chose a maximin criterion for several reasons. A first motivation is based on public policy recommendations. The maximin tax policy is the most redistributive one, as it corresponds to an infinite aversion to income inequality. Therefore, we will obtain the maximum tax rates that can be implemented in the presence of fiscal competition. A second motivation is ethical. In principle, the policy-maker might care for the well-being of the natives, irrespective of their country of residence, of the native taxpayers only, or of the taxpayers irrespective of citizenship. As an economist, there is no reason to favour

one of these criteria (Mirrlees, 1982). We focus on the maximin because, in this case, all the criteria coincide (Simula and Trannoy, 2011). A third reason is technical. Under the maximin, the set of agents whose welfare is to count is independent of the tax policy itself. We therefore avoid the difficulties raised by variable-population criteria (Blackorby, Bossert, and Donaldson, 2005).

Following Mirrlees (1971) seminal article, we consider that there is a fundamental separation 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 way.⁸ Therefore, it can only condition transfers on earnings y , through an income tax function $T_i(\cdot)$, which is assumed to be continuous.

Most countries do not levy income taxes abroad.⁹ 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:

$$E_i \leq \int_{w_0}^{w_1} T_i(Y(w)) \varphi(U_i(w) - U_{-i}(w); w) dw, \quad (6)$$

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

3. 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 of the effects of a small tax reform perturbation around the equilibrium.

3.1. Best Responses

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*

⁸In 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.

⁹US 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.

fine interacts with only one government. It is therefore equivalent to consider that the policy-maker chooses a tax schedule $T_i(Y_i(w))$ or an allocation $(C_i(w), Y_i(w))$.

An allocation is incentive compatible if and only if it satisfies the self-selection constraints:

$$C_i(w) - v(Y_i(w); w) \geq C_i(w') - v(Y_i(w'); w) \text{ for any } w, w' \text{ in } [w_0, w_1]. \quad (7)$$

Because $v''_{yw} < 0$, these constraints are equivalent to:

$$U'_i(w) = -v'_w(Y_i(w); w), \quad (8)$$

$$Y_i(\cdot) \text{ non-decreasing.} \quad (9)$$

The envelope condition (8) specifies at which rate utility must be increased to induce truthtelling. The second-order condition for incentive-compatibility (9) 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 will adopt the so-called ‘first-order approach’ which do not explicitly account for the monotonicity of $Y_i(w)$ when solving for the optimal schedules and then check, in computations, that the candidate schedules actually satisfy this condition.

In a best response, government i ’s optimisation problem is to choose an allocation $(C_i(w), Y_i(w))$, with a non-decreasing $Y_i(\cdot)$, solution to:

$$\begin{aligned} & \max_{U_i(w), Y_i(w)} U_i(w_0) \\ & \text{subject to:} \\ & E_i \leq \int_{w_0}^{w_1} (Y_i(w) - v(Y_i(w); w) - U_i(w)) \varphi(U_i(w) - U_{-i}(w); w) dw \\ & U'_i(w) = -v'_w(Y_i(w); w) \end{aligned}$$

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:

$$\begin{aligned} & \max_{U_i(w), Y_i(w)} \int_{w_0}^{w_1} (Y_i(w) - v(Y_i(w); w) - U_i(w)) \varphi(U_i(w) - U_{-i}(w); w) dw \\ & \text{subject to:} \\ & U_i'(w) = -v'_w(Y_i(w); w) \\ & U_i(w_0) \geq \underline{U}_i(w_0) \end{aligned} \tag{10}$$

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

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

Using Pontryagin's principle, the first-order conditions for a maximum are:¹⁰

$$1 - v'_y(Y_i; w) = \frac{q(w)}{\varphi(\Delta_i; w)} v''_{yw}(Y_i; w), \tag{11}$$

$$q'(w) = \{1 - [Y_i - v(Y_i; w) - U_i] \eta_i(\Delta_i; w)\} \varphi(\Delta_i; w), \tag{12}$$

$$q(w_1) = 0 \text{ when } w_1 < \infty \text{ and } q(w_1) \rightarrow 0 \text{ when } w_1 \rightarrow \infty, \tag{13}$$

$$q(w_0) \leq 0. \tag{14}$$

3.2. Symmetric Nash Equilibria

We focus on symmetric Nash equilibria for at least two reasons. One is intelligibility and tractability. In this case, the indirect utilities 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 $\varphi(\Delta_i; w) = f(w)$ and $\eta(\Delta_i; w) = g(0|w)$. The optimality conditions (11)-(14) 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. A 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.

¹⁰Equation (11) only holds at skill levels where bunching does not occur (i.e., when the constraint $Y'(w) \geq 0$ is not binding). However, this equation is sufficient in the subsequent proofs, as $Y'(w) = 0$ in case of bunching and bunching can only occur at kinks of the tax function for which $T'(Y(w))$ is not defined. The proof of Propositions below are valid under the assumption that $Y(\cdot)$ is continuous everywhere, and differentiable everywhere except on a finite set of skill levels (to allow for bunching). This assumption was made by Guesnerie and Laffont (1984).

Because we from now on focus on symmetric equilibria, we will drop the A and B subscripts, which are no longer necessary. Condition (12) can be written as:

$$q'(w) = [1 - \eta_0(w) \cdot T(Y(w))] \varphi(0; w) = [1 - \eta_0(w) \cdot T(Y(w))] f(w). \quad (15)$$

We integrate (15) between w and w_1 and make use of the transversality condition (13). We obtain:

$$q(w) = - \left[1 - F(w) - \int_w^{w_1} [\eta_0(x) T(Y(x))] f(x) dx \right]. \quad (16)$$

We then substitute the latter in (11), use the definitions of $\varepsilon(y; w)$ and $\alpha(y; w)$, and employ the first-order condition (2) of the individual utility maximisation programme. Rearranging, we obtain the following characterization.

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

$$\frac{T'(Y(w))}{1 - T'(Y(w))} = \frac{\alpha(Y(w); w)}{\varepsilon(Y(w); w)} \frac{1 - F(w)}{wf(w)} \{1 - \mathbb{E}[\eta_0(x) T(Y(x)) | x \geq w]\},$$

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:

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

A third factor appears in the present open economy. This factor captures the fact that the trade-off between equity and efficiency, as described in Diamond's formula, is modified by the threat of migration. We see that the 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 country i : the marginal tax rate $T'_i(Y(w))$ is uniformly increased by Δ on the interval $[Y_i(w) - \delta, Y_i(w)]$ as shown in Figure 1. This gives rise to the following effects. First, everyone located in $[Y_i(w) - \delta, Y_i(w)]$ reduces taxable income, which decreases collected taxes. This is the usual substitution effet. Second, every individual with income above $Y_i(w)$ faces a lump-sum increase $\$ \delta \Delta$ in his/her tax liability. This increases collected taxes from the w -individuals by $\$ \delta \Delta f(w)$. This is referred to as the “mechanical” effect in the literature. However, an additional effect takes place in

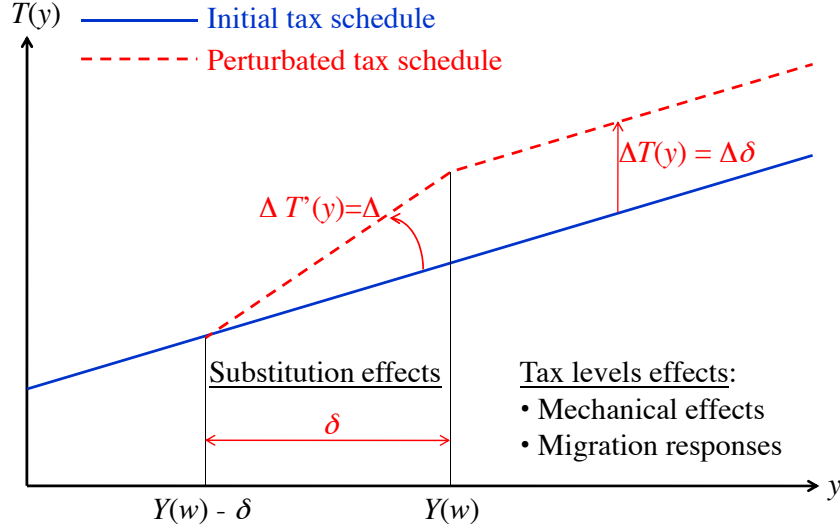


Figure 1: Small Tax Reform Perturbation

the present open-economy setting. The reason is that the rise in tax liability reduces the utility in country i compared to the utility abroad. Consequently, the number of citizens who emigrate is increased and the number of foreigners who immigrate is decreased. This diminishes the number of taxpayers with skill w by $\eta_0(w)f(w)$, and thus collected taxes by $\eta_0(w)f(w)T_i(w)$. Because we considered a unilateral marginal perturbation around a Nash equilibrium, the three effects must sum to zero. We then obtain the optimal marginal tax rates given by Proposition 1.

Migration choices do not directly depend on the marginal tax rates, but on the average tax rates. However, average tax rates are naturally determined by the profile of marginal tax rates. An alternative way of writing the formula given in Proposition 1 illuminates the relationship between the marginal and the average optimal tax rates. Using the definition of $\nu(0; w) \equiv \nu_0(w)$ and $C(w) = Y(w) - T(Y(w))$, we obtain:

$$\frac{T'(Y(w); w)}{1 - T'(Y(w); w)} = \frac{\alpha(Y; w)}{\varepsilon(Y; w)} \frac{1 - F(w)}{wf(w)} \left[1 - \mathbb{E}_f \left(\frac{T(Y(x))}{Y(x) - T(Y(x))} \nu_0(x) \mid x \geq w \right) \right]. \quad (17)$$

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.

4. 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 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 ($w_1 \rightarrow \infty$). 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, in the symmetric Nash equilibrium, competing governments will 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,*

1. $T'(Y(w_0)) \geq 0$,
2. $T'(Y(w_1)) = 0$ when $w_1 < \infty$.

Proof. (1.) The result is established by contradiction. $q(w_0)$ corresponds to the derivative of the value function of the dual problem (10) 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 $v''_{yw}(y(w); w) < 0$, (11) implies $v'_y(Y(w_0); w_0) \leq 0$. It then follows from (2) that $T'(Y(w_0)) \geq 0$. (2.) When $w_1 < \infty$, the transversality condition is $q(w_1) = 0$. This implies that the migration factor (curly bracket in Proposition 1) is zero and thus that $T'(Y(w_1)) = 0$. ■

The results of Proposition 2 are not very informative. 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 in particular make the connection between the (semi)-elasticity of migration and the sign of the optimal marginal tax rates. To this aim, we distinguish three situations: decreasing, constant and increasing semi-elasticities.

4.1. On the Slope of the Semi-Elasticity of Migration

Let us consider that everyone receives one extra euro of net income in country i . This will reduce out-migration and increase in-migration. Altogether, the population of taxpayers of skill w varies by $\partial\varphi(\Delta;w)/\varphi(\Delta;w)\%$. This relative change is the semi-elasticity of migration.

The elasticity of migration is equal to the semi-elasticity multiplied by net income $C(Y(w))$. We have seen that, in the second best, truth-telling requires that $C(Y(w))$ is non-decreasing in w . Therefore, when the semi-elasticity is increasing, the elasticity is also increasing. When the elasticity is constant, the semi-elasticity is non-decreasing. When the semi-elasticity is decreasing, the elasticity can be non-decreasing, constant, or decreasing. So the semi-elasticity may be decreasing while the elasticity is increasing.

Three natural benchmarks come to mind when thinking about migration. First, the costs of migration may be decreasing in w . This seems to be suggested by the empirical evidence that highly skilled are more likely to emigrate than low skilled (Docquier and Marfouk, 2006). This assumption was made by Simula (2010), where the semi-elasticity of migration was then piecewise increasing. Second, the costs of migration may be independent of w as in Blumkin, Sadka, and Shem-Tov (2012) and Morelli, Yang, and Ye (20012). This makes sense, in particular, if most relation costs are material (moving costs, flight tickets, etc.). The second article compares 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. 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 euro of consumption in country i , then the relative increase in the number of taxpayers becomes smaller for more skilled individuals.

4.2. A Useful Benchmark: The First-And-A-Half Best

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 “first-best-and-a-half” situation (Jacquet, Lehmann, and Van der Linden, 2012) in which w is public information and m private information. More precisely, the first-best-and-a-half allocation is defined as the solution to programme (10), without the conditions for incentive compatibility. We will see in the next subsections that the results derived in this informational context are useful to sign optimal marginal tax rates in the second-best environment.

The first-order conditions are obtained by setting $q(w) = 0$ in (11) and (12), for $w > w_0$. Denoting the optimal tax function in the first-and-a-half best by $T_i^{1.5}(Y_i)$, we respectively obtain:

$$1 - v'_y(Y_i; w) = 0, \quad (18)$$

$$T_i^{1.5}(Y_i) = \frac{1}{\eta_i(\Delta_i; w)}. \quad (19)$$

Because w is observable, country i 's policy-maker determines the effort level required from the residents of skill w . This level is implicitly defined by (18) and governs pre-tax earnings. Pre-tax earnings are then taxed according to (19). Hence, given the foreign tax policy, the tax liability $T_i^{1.5}(Y(w))$ required from the residents with skill $w > w_0$ is equal to the inverse of their semi-elasticity of migration $\eta_i(\Delta_i, w)$. The least productive individuals receive a transfer determined by the budget constraint. Therefore, the optimal tax function is discontinuous at $w = w_0$. This is illustrated in Figure 2.

This inverse-elasticity formula has an intuitive interpretation. Let us consider that the tax liability faced by the workers of skill w (with $w > w_0$) is increased by \$1 in country i . In the absence of any migration response, collected taxes are increased by $\varphi(\Delta_i; w)$. However, a migration response takes place in addition to this mechanical effect. Indeed, the tax increase reduces the utility in country i , and thus the value of Δ_i . The resulting behavioral response along the migration margin reduces collected taxes by $\eta_i(\Delta_i; w)\varphi(\Delta_i; w)T_i^{1.5}(Y_i)$. In the best response, the mechanical and behavioral effects must sum to zero, which implies (19).

Using the elasticity of migration $\nu_i(\Delta_i; w) = [Y_i - T_i^{1.5}(Y_i)] \eta_i(\Delta_i; w)$ instead of the semi-elasticity, we can alternatively express the best response of country i 's policy-maker as:

$$\frac{T_i^{1.5}(Y_i)}{Y_i - T_i^{1.5}(Y_i)} = \frac{1}{\nu(\Delta_i; w)}. \quad (20)$$

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 first-and-a-half best, the symmetric Nash equilibrium allocation can be decentralized by a tax function such that*

$$T^{1.5}(Y(w)) = \frac{1}{\eta_0(w)}. \quad (21)$$

We can now make the link between the first-and-a-half best and the second best. From Proposition 1, we know that the second-best optimal allocation can be decentralized by a tax function satisfying

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

with

$$X(w) = \int_w^{w_1} [1 - \eta_0(x)T(Y(x))]f(x)dx.$$

Consequently, $T'(Y(w))$ has the same sign as $X(w)$. Proposition 2 is thus equivalent to (i) $X(w_0) \geq 0$ and (ii) $X(w_1) = 0$ when $w_1 < \infty$. The derivative of $X(w)$ is

$$X'(w) = \eta_0(w)T(Y(w)) - 1.$$

Therefore,

$$X'(w) > 0 \Leftrightarrow T(Y(w)) > \frac{1}{\eta_0(w)}.$$

Note that the right-hand side of this inequality only depends on $g(0|w)$, which is a structural parameter of the economy. We see that $X(w)$ is increasing (decreasing) when the tax paid in the second best is larger (lower) than that paid in the first-and-a-half best $T^{1.5}(Y(w)) = 1/\eta_0(w)$.

4.3. Decreasing Semi-Elasticity of Migration

We first investigate the case when the semi-elasticity of migration is decreasing. This is for example the case when the elasticity is constant.

We know from Proposition 3 that, in the first-and-a-half best, the optimal tax liability is increasing in w . The set of points where $X'(w) = 0$ has equation $T(Y(w)) = 1/\eta_0(w)$ and separates the $(w, T(Y(w)))$ -space in two sets, as shown in Figure 2. The upper contour set corresponds to situations in which $X(w)$ is increasing, because $T(Y(w)) > 1/\eta_0(w)$, and the lower contour set to situations in which $X(w)$ is decreasing.

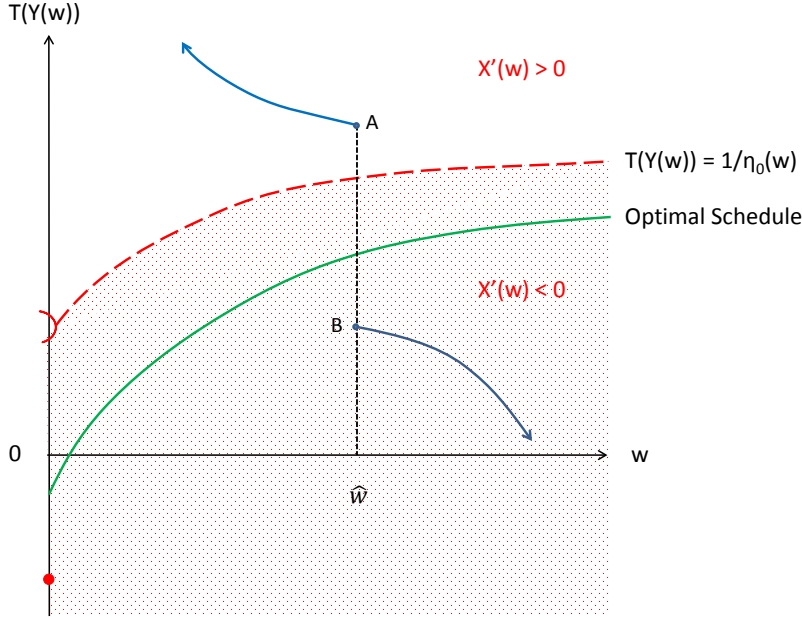


Figure 2: Decreasing Semi-Elasticity of Migration

By close inspection of Figure 2, we can determine the slope of the optimal tax function. Let us assume that there is $\hat{w} > w_0$ such that $T'(Y(\hat{w})) \leq 0$, that is $X(\hat{w}) \leq 0$. If $T(Y(\hat{w})) > 1/\eta_0(\hat{w})$ (like at point A), then $X(w)$ becomes even more negative and increasing when w goes down. This contradicts $X(w_0) \geq 0$. If, on the contrary, $T(Y(\hat{w})) < 1/\eta_0(\hat{w})$ (like at point B), then $X(w)$ becomes even more negative and decreasing when w goes down. This contradicts the transversality condition $X(w_1) = 0$ when $w_1 < \infty$ and $X(w_1) \rightarrow 0$ when $w_1 \rightarrow \infty$. Therefore:

Proposition 4 *If $\eta'_0(w) < 0$, then $T'(Y(w)) > 0$.*

Intuitively, the optimal tax policy tries to replicate the first-and-a-half best, but in a flatter way, to minimize the additional distortions required to satisfy the incentive-compatibility constraints. When the semi-elasticity of migration is decreasing, the former involves richer individuals paying higher taxes.

Brewer, Saez, and Shephard (2010) and Piketty and Saez (2012) look at the asymptotic marginal tax rate, or in an equivalent way at the marginal tax rate for the last bracket of the income tax schedule. They assume that the elasticity of migration is constant ($\nu_0(x) = \nu_0$), which is a special case of the decreasing semi-elasticity. They also assume that the distribution of skills is Pareto in its upper part, with parameter $k = (wf(w))/(\alpha_0(w)(1 - F(w)))$, and that $\varepsilon_i(y; w)$ converges to ε when $w_1 \rightarrow \infty$. By Proposition 4, we know that the optimal asymptotic marginal

tax rate is positive. Moreover, by (17),

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

When $w_1 \rightarrow \infty$, the average tax rate converges to

$$\frac{T(Y(w))}{Y(w) - T(Y(w))} \rightarrow \frac{T'(Y(w))}{1 - T'(Y(w))}.$$

Consequently,

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

Solving for $T'(Y(\infty))$, we obtain the formula derived in Brewer, Saez, and Shephard (2010) and Piketty and Saez (2012):

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

The asymptotic marginal tax rate is thus strictly positive. For example, if $k = 2$, $\varepsilon = 0.15$ and $\nu_0 = 0.1$, we obtain $T'(Y(\infty)) = 71\%$ instead of 77% in the absence of migration responses.

4.4. Constant Semi-Elasticity of Migration

We now consider that the semi-elasticity of migration is constant, with $\eta_0(w) = \eta_0$. This is for example the case when w and m are independent, as in Blumkin, Sadka, and Shem-Tov (2012) and Morelli, Yang, and Ye (20012).

This situation is shown in Figure 3. We do not have an initial condition regarding the value of $T(Y(w))$. However, we know that $X(w_1) = 0$ when $w_1 < \infty$ and $X(w_1) \rightarrow 0$ when $w_1 \rightarrow \infty$. Let us assume that $T(Y(w))$ approaches $1/\eta_0$ from above. Then $T'(Y(w))$, and thus $X(w)$ must be negative on the left of w_1 . Because $X(w)$ is increasing in w in that part of the space, the latter implies that the tax liability is maximum for w_0 , and then decreases, reaching its minimum for top income earners. This clearly contradicts the maximin social objective.¹¹ Now let us assume that $T(Y(w))$ is just equal to $1/\eta_0$ to the left of w_1 . As $X(w)$ is constant along the line $T(Y) = 1/\eta_0$, $T(Y(w))$ is constant for $w_0 < w \leq w_1$. Contrary to the first-and-a-half 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. Consequently, we obtain a second-best solution where the tax function is constant. Given that the tax policy is strictly redistributive, this also contradicts the maximin objective. Therefore, we are left with only one possibility:

¹¹For example, a laissez-faire policy would do better.

$T(Y(w))$ approaches $1/\eta_0$ from below. Because $X(w)$ is decreasing in that part of the space, $X(w)$ is strictly positive below w_1 . Hence, the optimal marginal tax rate is strictly positive for $w < w_1$, equal to 0 at $w_1 < \infty$ and asymptotically equal to 0 when $w_1 \rightarrow \infty$.

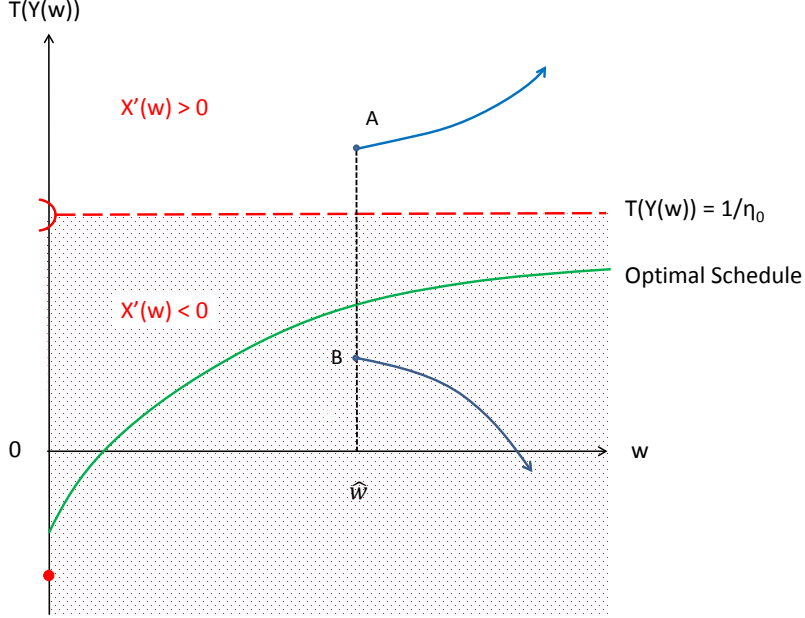


Figure 3: Constant Semi-Elasticity of Migration

Proposition 5 *Assume that $\eta'_0(w) = 0$. Then, $T'(Y(w)) > 0$ for $Y(w) < Y(w_1)$. Moreover, $T'(Y(w_1)) = 0$ if $w_1 < \infty$ and $T'(Y(w_1)) \rightarrow 0$ otherwise.*

4.5. 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.

If $T(Y(w))$ approaches $1/\eta_0$ from above, there must be a skill level $\hat{w} < w_1$ where the tax function crosses the curve $X'(w) = 0$. Otherwise, $T(Y(w))$ would be maximum at w_0 , which would contradict the maximin social objective. Therefore, the equation $X(w) = 0$, i.e. $T'(w) = 0$, has at most one root. This implies that there are two possible patterns for the tax function: (i) the tax function is increasing and approaches w_1 from below or (ii) it is first increasing and then decreasing, approaching $1/\eta_0$ from above. In summary:

Proposition 6 *Assume that $\eta'_0(w) > 0$, then:*

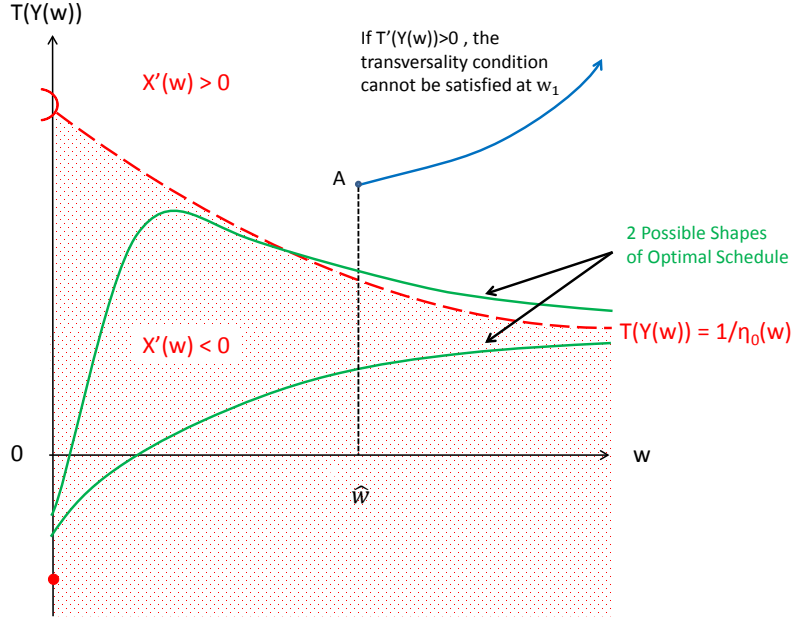


Figure 4: Increasing Semi-Elasticity of Migration

- (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 will occur when skills are not bounded from above and 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 its maximum at zero. The budget constraint of the policy-maker must thus be violated. Consequently, we are left with case (ii) in Proposition 4.

Proposition 7 Assume that $\eta'_0(w) > 0$ and $\lim_{w_1 \rightarrow \infty} \eta_0(w) = \infty$, 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

Semi-Elasticity	Elasticity	Shape of Tax Function	Top Tax Rates
Decreasing	Monotonic	Increasing	$T'(Y(w_1)) > 0$
Constant	Non-Decreasing	Increasing	$T'(Y(w_1)) = 0$ if $w_1 < \infty$ $T'(Y(w_1)) = 0$ if $w_1 \rightarrow \infty$
Increasing	Increasing	a) Increasing	$T'(Y(w_1)) > 0$
		b) Hump-Shaped	$T'(Y(w_1)) < 0$

Table 1: Main Features of the Equilibrium Tax Schedule

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.

5. Simulations

This section provides numerical simulations of the equilibrium optimal tax schedule that competing policy-makers should implement. Our objective is to emphasize the part played by the slope of the semi-elasticity of migration. In particular, we will show that the top marginal tax rates are highly sensitive to the overall shape of the semi-elasticity.

We calibrate the skill distribution $f(w)$ using the CPS 2007 distribution of weakly earnings for singles without kids. It is recovered assuming a tax function $T(Y) = 0.4058Y + 4036$. Following Diamond (1998) and Saez (2001), we correct for top coding by extended the obtained estimation with a Pareto distribution of coefficient 1.5. 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.

The (semi-)elasticity of migration is certainly the 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.

Liebig, Puhani, and Sousa-Poza (2007) study the mobility of highly educated young workers across Swiss Cantons as a response to Canton taxes. They find that a one percentage point higher tax rate increase leads to the outmigration of 33 out

of 1,000 young Swiss college graduates. 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 top 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)). We choose 0.15 in our computations.

There are no empirical studies regarding the possible shape of the (semi-)elasticity of migration. We therefore investigate three possible scenarios, as shown in Figure 5. All of them are calibrated in such a way that the average semi-elasticity of migration is the same. In the first one, the semi-elasticity is increasing. In the second one, it is kept constant. And in the last one, it is decreasing.

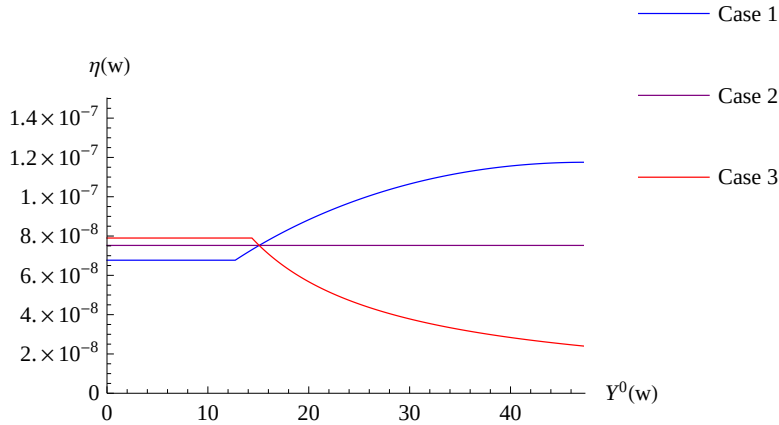


Figure 5: Three Benchmark Cases

We already know about the shape of the tax function (see 1), but we did not know to which extent the tax rates and tax liabilities will differ. The optimal equilibrium tax liabilities are shown in Figure 6. 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 (called “Closed”). We observe that the threat of migration implies a significant decrease in the total taxes paid by top income earners. The decrease is however much smaller when the semi-elasticity is decreasing (Case 3), than when it is constant (Case 2) or increasing

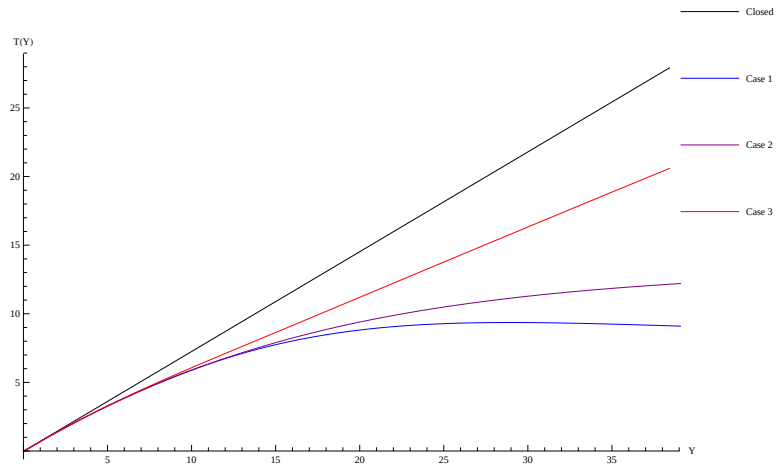


Figure 6: Optimal Tax Liabilities

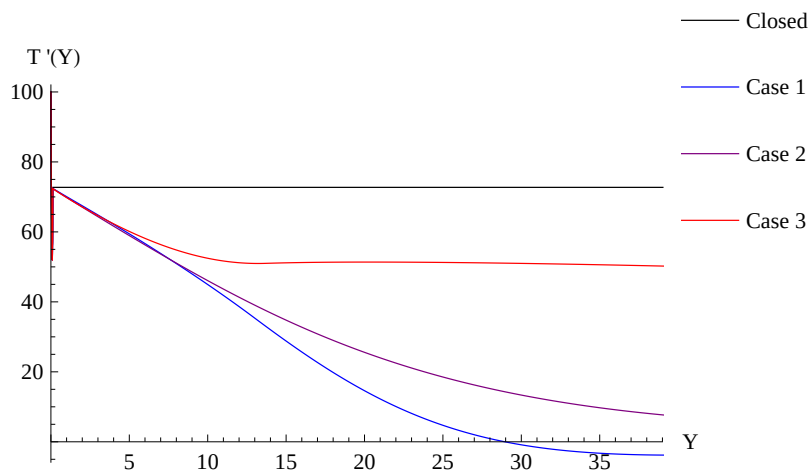


Figure 7: Optimal Marginal Tax Rates

(Case 1). Figure 7 casts light on the differences in the optimal marginal tax rates. What we see is that slight 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.

6. 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 economy. 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 (semi-)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

- 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., C. BRETT, AND J. A. WEYMARK (2011): “Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility,” (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.
- GUESNERIE, R., AND J.-J. LAFFONT (1984): “A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm,” *Journal of Public Economics*, 25, 329–369.

- 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 Across Countries,” *Journal of Public Economics*, 65, 229–244.
- LIEBIG, T., P. PUHANI, AND A. SOUSA-POZA (2007): “Taxation And Internal Migration-Evidence From The Swiss Census Using Community-Level Variation In Income Tax Rates,” *Journal of Regional Science*, 47(4), 807–836.
- 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.
- 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. (2010): “Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties,” *Social Choice and Welfare*, 35, 199–220.
- SIMULA, L., AND A. TRANNOY (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, 381–397.