

The Dark Side of Tax Breaks for Foreigners

Laurent Simula* and Alain Trannoy†

February 3, 2018

Following Mirrlees (1982), the literature has focused on tax competition between Nation States, considering that, in a given country, all tax residents face the same income tax schedule (see, e.g., Blumkin, Sadka, Shem-Tov, 2014; Simula and Trannoy, 2010 and 2012; as well as Lehmann, Simula and Trannoy, 2015). We in contrast allow for the possibility of tax differentiation based on citizenship, in each country, and investigate the impact on well-being from a national as well as from an international perspective. Can it be optimal to rely on such a tax differentiation? Or should it be avoided? To investigate this issue, we use a world consisting of two (non-necessarily symmetric) countries. Individuals differ with respect to three dimensions of heterogeneity: native country, skill and cost of migration. The distribution of types is common knowledge. However, the skill and migration cost of a given individual are private information. In each country, a benevolent policymaker aims at redistributing incomes from the rich to the poor people. Following Mirrlees (1971), the government is unable to observe the skill nor the cost of migration of a given individual. In addition, the latter can only levy taxes on residents. However, contrary to Lehmann, Simula and Trannoy (2014) and the previous literature, it is not constrained to treat natives and foreigners in the same way. Some form of "tagging" (Akerlof, 1978), based on citizenship, is thus allowed. This radically changes the problem each government faces. Indeed, given these circumstances, each policymaker designs two income tax schedules, one for the native residents and one for the resident expatriates. We first characterize the tax schedules of each country in a Nash equilibrium, assuming that skills (but not migration costs) can be observed. This intermediary situation, referred to as the "Tiebout best" in Lehmann, Simula and Trannoy (2014), will give us insights into the second best, in which skills as well as migration costs are private information. We then turn to the characterization of the second best. We contrast the equilibrium schedules and profiles of well-being with those obtained when each country is not allowed to implement tax breaks for foreigners. Numerical illustrations, calibrated using US and Danish data, are then provided, to quantify the reduction in social welfare resulting from the introduction of tax breaks for foreigners. *JEL Codes: D82, H21, H87, F22.*

*University Grenoble Alpes and Grenoble Applied Econ Lab (GAEL). Email: laurent.simula@univ-grenoble-alpes.fr

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS. Email: alain.trannoy@univ-amu.fr

I. INTRODUCTION

The globalization process has not only made the mobility of capital easier. The transmission of ideas, meanings and values across national borders associated with the decrease in transportation costs has also reduced the barriers to international labor mobility. In this context, individuals are more likely to vote with their feet in response to high income taxes. This is, in particular, the case for highly skilled workers. The latter represent about 35% of the OECD immigration stock, for only 11.3% of the world labor force. In this context, a highly skilled is six times more likely to emigrate than a low-skilled agent Docquier and Marfouk (2005). There are still few empirical studies estimating the migration responses to taxation; however, they all suggest that highly skilled are very sensitive to tax differences when choosing where to locate (cf. Liebig et al. (2007), Kleven et al. (2013) and Kleven et al. (2014)). Consequently, the possibility of tax-driven migrations appears as an important policy issue and must be taken into account as a salient constraint when thinking about the design of taxes and benefits affecting households.

In this context, we observe two trends regarding top income tax rates payable to central governments: a tendency to fall down until the financial crisis of 2007-2008 and the introduction, and development in the last few years, of specific tax cuts for highly skilled or rich foreigners. These tax breaks for foreigners exist in all Scandinavian countries, in Germany, in the Netherlands, in Belgium, in France, in the UK, in Switzerland, but also in Australia or in Canada, etc. In addition, they have recently been extended in France (“impatriates” can benefit from it during 8 years, instead of 3 years) or in the UK (to avoid the departure of expatriates after the “Brexit”).

Following Mirrlees (1982), the literature has focused on tax competition between Nation States, considering that, in a given country, all tax residents face the same income tax schedule (see, e.g., Blumkin et al. (2012); Simula and Trannoy (2010); as well as Lehmann et al. (2014)). We, in contrast, allow for the possibility of tax differentiation based on citizenship, in each country, and investigate the impact on well-being from a national as well as from an international perspective. Can it be optimal to rely on such tax differentiation? Or should it be avoided? Our ambition is to determine whether and to which extent this form of tax differentiation is detrimental in terms of social welfare when it is implemented in a variety of countries. If the results of our analysis confirm our preliminary investigations and intuition, our research would cast light on the non-sustainability of tax breaks for highly skilled foreigners and, thus, call for a change of direction regarding ongoing public policy practices.

To investigate this issue, we use a world consisting of two (non-necessarily symmetric) countries. Individuals differ with respect to three dimensions of heterogeneity: native country, skill, and cost of migration. The distribution of types is common knowledge. However, the skill and migration cost of a given individual are private information. In each country, a benevolent policymaker aims at redistributing incomes from the rich to the poor people. Following Mirrlees (1971), the government is unable to observe the skill nor the cost of migration of a given individual. In addition, the latter can only levy taxes on residents. However, contrary to Lehmann, Simula and Trannoy (2014) and the previous literature, it is not constrained to treat natives and foreigners in

the same way. Some form of “tagging” (Akerlof, 1978), based on citizenship, is thus allowed. This radically changes the problem each government faces. Indeed, given these circumstances, each policymaker designs two income tax schedules, one for the native residents and one for the resident expatriates. Given that our focus is on tax-driven migrations, we consider that the production technology is the same in each country and exhibits constant returns to scale. Hence, each individual’s wage is equal to his/her productivity and does not vary in case of migration. We also consider that each agent is constrained to consume and work in the same country. In addition, we consider that, in each country, the policymaker does not account for the well-being of the expatriates in the social welfare function. In other words, expatriates are mainly regarded (by the State) as net taxpayers, i.e., as agents from which taxes can be collected. This assumption can be motivated by the actual deterioration of government balance in a large number of developed countries. Our analysis proceeds in steps. We first characterize the tax schedules of each country in a Nash equilibrium, assuming that skills (but not migration costs) can be observed. This intermediary situation, referred to as the “Tiebout best” in Lehmann et al. (2014), will give us insights into the second best, in which skills as well as migration costs are private information. We then turn to the characterization of the second best. We contrast the equilibrium schedules and profiles of well-being with those obtained when each country is not allowed to implement tax breaks for foreigners.

II. A MODEL ALLOWING FOR TAX BREAKS ON FOREIGNERS

We consider an economy consisting of two countries, indexed by $i = \{A, B\}$. The same constant return to scales technology is available in both countries. 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, w_0 is strictly positive¹, and any agent’s productivity w is unaltered in case of migration. In addition, the empirical evidence that some people are immobile is captured by the possibility of infinitely large migration costs. This in particular implies that there will always be a mass of natives of skills w in each country. 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, and so on. We do not make any restriction on the correlation between skills and migration costs. We simply consider that there is a distribution of migration costs for each possible skill level.

In country $i = \{A, B\}$, we denote by $h_i(w)$ the initial continuous skill density, by $H_i(w) \equiv$

¹This assumption is introduced to avoid corner solutions that would arise if some agents were completely unproductive (i.e., with $w_0 = 0$). Indeed, whatever effort they provide, the latter would never be able to pay positive taxes. To account for the fact that some groups of the population are unable to produce (because, e.g., of inability for example), it would be worthwhile to introduce tagging based on this dimension as emphasized by Akerlof (1978). This, however, is not the issue addressed in the article.

$\int_{w_0}^w h_i(x)dx$ the corresponding cumulative distribution function (CDF) and by N_i the size of the population. For each skill w , $g_i(m|w)$ denotes the conditional density of the migration cost and $G_i(m|w) \equiv \int_0^m g_i(x|w)dx$ the conditional CDF. The initial joint density of (m, w) is thus $g_i(m|w)h_i(w)$ while $G_i(m|w)h_i(w)$ is the initial mass of individuals of skill w with migration costs lower than m .

Following Mirrlees (1971), the government does not observe individual types (w, m) . We refer to this information setting as the “second best”; we will also examine alternative information settings, but with the only aim of illuminating second-best allocations. In addition, each policymaker can only levy taxes on residents. However, contrary to Lehmann, Simula and Trannoy (2014), the latter is not constrained to treat natives and foreigners in the same way. Some form of tagging, based on citizenship, is thus allowed. Given these circumstances, each policymaker designs two income tax schedules, for the native residents on the one hand and for the resident expatriates on the other hand.

II.1. Individual Choices

Every worker derives utility from consumption c , and disutility from effort and migration if any. Effort captures the quantity as well as the intensity of labor supply. The choice of effort corresponds to an intensive margin and the migration choice to an extensive margin. Let $v(y; w)$ be the disutility of a worker of skill w to obtain pre-tax earnings $y \geq 0$ with $v'_y > 0 > v'_w$ and $v''_{yy} > 0 > v''_{yw}$. 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:

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

Note that the Spence-Mirrlees single-crossing condition holds because $v''_{yw} < 0$. The quasi-linearity in consumption implies that there is no income effect on taxable income and appears as a reasonable approximation.

Intensive Margin

We focus on income tax competition while allowing each country to implement tax differentiation between native and foreign residents. Every native living in country i is liable to an income tax $T_i^N(\cdot)$; every foreign resident faces the income tax $T_i^M(\cdot)$, where the superscripts $j = \{N, M\}$ stand for “natives” and “mercenaries” respectively. Hence, an agent of skill w , who has chosen to work in country i , solves:

$$(2) \quad U_i^j(w) \equiv \max_y y - T_i^j(y) - v(y; w).$$

We refer to $U_i^j(w)$ as the corresponding *gross* utility. It is the net utility level for a native and the utility level absent migration cost for a mercenary. We call $Y_i^j(w)$ the solution to program (2) and

$C_i^j(w) = Y_i^j(w) - T_i^j(Y_i^j(w))$ the agent's consumption level.² The first-order condition can be written as:

$$(3) \quad 1 - MTR_i^j(Y_i^j(w)) = v'_y(Y_i^j(w); w),$$

where $MTR_i^j(Y_i^j(w)) \equiv [T_i^j]'$ ($Y_i^j(w)$) denotes the marginal tax rate. Differentiating (3), we obtain the elasticity of gross earnings with respect to the retention rate $1 - MTR_i^j$,

$$(4) \quad \varepsilon_i^j(w) \equiv \frac{1 - MTR_i^j(Y_i^j(w))}{Y_i^j(w)} \frac{dY_i^j(w)}{d(1 - MTR_i^j(Y_i^j(w)))} = \frac{v'_y(Y_i^j(w); w)}{Y_i^j(w)v''_{yy}(Y_i^j(w); w)},$$

and the elasticity of gross earnings with respect to productivity w :

$$(5) \quad \alpha_i^j(w) \equiv \frac{w}{Y_i^j(w)} \frac{dY_i^j(w)}{dw} = -\frac{wv''_{yw}(Y_i^j(w); w)}{Y_i^j(w)v''_{yy}(Y_i^j(w); w)}.$$

The Migration Margin

A native of country i , of type (w, m) , gets utility $U_i^N(w)$ in country i and utility $U_{-i}^M(w) - m$ in country $-i$. She therefore emigrates if and only if $U_{-i}^M(w) - m > U_i^N(w)$, i.e., when $m < U_{-i}^M(w) - U_i^N(w)$. We define:

$$(6) \quad \Delta_i \equiv U_i^N(w) - U_{-i}^M(w)$$

and

$$(7) \quad \Delta_{-i} \equiv U_{-i}^N(w) - U_i^M(w).$$

The number of natives with skill w staying in country i is given by:

$$(8) \quad \varphi_i^N(\Delta_i; w) \equiv \begin{cases} h_i(w)N_i & \text{if } \Delta_i \geq 0 \\ (1 - G_i(-\Delta_i|w))h_i(w)N_i & \text{if } \Delta_i < 0 \end{cases}$$

which is increasing in Δ_i . Because migration costs have support \mathbb{R}_+ , we know that $\varphi_i^N(\Delta_i; w) > 0$ at every w : there is always a positive measure of natives of skill w . The number of foreigners in country i ("tax mercenaries") is obtained as:

$$(9) \quad \varphi_i^M(\Delta_{-i}; w) \equiv \begin{cases} 0 & \text{if } \Delta_{-i} \geq 0 \\ G_{-i}(-\Delta_{-i}|w)h_{-i}(w)N_{-i} & \text{if } \Delta_{-i} < 0 \end{cases}$$

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

This non-negative quantity decreases in Δ_{-i} for $\Delta_{-i} < 0$. Combining natives and foreigners, there are altogether $\varphi_i^N(\Delta_i; w) + \varphi_i^M(\Delta_{-i}; w)$ agents of skill w staying in country i , and a total population of $\int_{w_0}^{w_1} [\varphi_i^N(\Delta_i; w) + \varphi_i^M(\Delta_{-i}; w)] dw$ agents in this country.

It is useful to summarize responses along the migration margin in terms of semi-elasticities. When $\Delta_i < 0$ at a given w , we define the *semi-elasticity of staying* in the native country as:

$$(10) \quad \eta_i^N(\Delta_i; w) \equiv \frac{\partial \varphi_i^N(\Delta_i; w)}{\partial \Delta_i} \frac{1}{\varphi_i^N(\Delta_i; w)} = \frac{g_i(-\Delta_i|w)}{1 - G_i(-\Delta_i|w)}.$$

We otherwise set $\eta_i^N(\Delta_i; w) = 0$. Because of quasi-linearity in consumption, this semi-elasticity corresponds to the percentage change in the density of native taxpayers with skill w when their consumption $C_i^N(w)$ is increased at the margin. When $\Delta_{-i} < 0$ at a given w , we introduce the *semi-elasticity of expatriation* from country $-i$ to country i to summarize extensive responses by tax mercenaries; we denote it by:

$$(11) \quad \eta_i^M(\Delta_{-i}; w) \equiv \frac{\partial \varphi_i^M(\Delta_{-i}; w)}{\partial \Delta_{-i}} \frac{1}{\varphi_i^M(\Delta_{-i}; w)} = -\frac{g_{-i}(-\Delta_{-i}|w)}{G_{-i}(-\Delta_{-i}|w)}.$$

We otherwise set $\eta_i^M(\Delta_{-i}; w) = 0$. At skill w , the semi-elasticity of expatriation corresponds to the percentage change in the density of foreign taxpayers in country i when their consumption $C_i^M(w)$ is increased at the margin.

II.2. Governments

In country $i = \{A, B\}$, a benevolent policymaker designs the tax system to maximize the welfare of the worst-off natives (“maximin”). The maximin tax policy is the most redistributive one given the set of natives, as it corresponds to an infinite aversion to income inequality. The focus on the worst-off is primarily introduced to simplify the technical analysis. Alternatively, we could have considered a weighted sum of the indirect utilities of the natives, irrespective of their country of residence, or of the resident natives only. In the second case, the set of agents whose welfare is to count would be endogenous. This would raise further difficulties.

What is crucial to our setting is the idea that foreigners are regarded by each government as a potential source of tax receipts, which can be extracted through the implementation of a specific tax scheme. Whatever the tax schedule abroad, the assumption that the cost of migration has support \mathbb{R}^+ ensures that a positive measure of *natives* always stay in country i at every skill level w , thus preventing an “empty jurisdiction” issue skill by skill. This moreover corresponds to the empirical observation that there are no holes in the distribution of skills in a given country. The same reasoning does not apply to the set of tax mercenaries. In all generality, this set may either be empty, an interval, or the union of several intervals.

As mentioned earlier, in each country, the tax system consists of two non-linear income tax schedules, T_i^N and T_i^M , offered to the native and foreign residents respectively. The budget con-

straint faced by country i 's government is therefore:

$$(12) \quad \int_{w_0}^{w_1} T_i^N(Y_i^N(w))\varphi_i^N(U_i^N(w) - U_{-i}^M(w); w)dw + \int_{w_0}^{w_1} T_i^M(Y_i^M(w))\varphi_i^M(U_i^M(w) - U_{-i}^N(w); w)dw \geq E,$$

where $E \geq 0$ is an exogenous amount of public expenditures to finance. For simplicity, we concentrate on purely redistributive tax policies, in which case $E = 0$.

III. BEST RESPONSES

III.1. Definition

We start with the characterization of each policymaker's best response. Because a taxpayer interacts with only one policymaker at the same time, it is easy to show that the standard taxation principle holds. Hence, in each country $i = \{A, B\}$, it is equivalent to choose the non-linear income taxes T_i^N and T_i^M , taking individual choices into account, or to directly select an allocation satisfying the incentive-compatible constraints

$$(13) \quad C_i^j(w) - v(Y_i^j(w); w) \geq C_i^j(x) - v(Y_i^j(x); w)$$

for every $(w, x) \in [w_0, w_1]^2$ and $j = \{M, N\}$. Note that, because citizenship is observable, mimicking is impossible between native and foreign residents with a same skill level. Due to the single-crossing condition, the former implies that an allocation is incentive compatible if and only if:

$$(14) \quad [U_i^N]'(w) = -v'_w(Y_i^N(w); w),$$

$$(15) \quad [U_i^M]'(w) = -v'_w(Y_i^M(w); w),$$

$$(16) \quad Y_i^N(w) \text{ non-decreasing,}$$

$$(17) \quad Y_i^M(w) \text{ non-decreasing.}$$

Conditions (14) and (15) are the first-order conditions for incentive compatibility. Because of (14), the indirect utility is increasing in w within the set of native residents. In addition, natives only choose to emigrate abroad if they get larger utility there. Consequently, maximizing the utility of the worst-off natives amounts to setting $U_i^N(w_0)$ to its maximum value. Because of (15), the indirect utility is also increasing in w within the set of "tax mercenaries". The monotonicity constraints (16) and (17) correspond to the second-order conditions for incentive compatibility. We below adopt the so-called "first-order approach" and neglect second-order conditions when deriving qualitative results. These conditions can be checked *ex post* when performing numerical simulations.

Because of the taxation principle, as explained above, choosing a tax policy (T_i^N, T_i^M) is equivalent to determining an incentive-compatible allocation $w \rightarrow (Y_i^N(w), U_i^N(w), Y_i^M(w), U_i^M(w))$. The

best-response allocation of government i is therefore solution to:

$$(18) \quad \max_{\{U_i^j(w), Y_i^j(w)\}_{j=\{M,N\}}} U_i^N(w_0) \text{ s.t. } U_i^j(w) = -v'_w(Y_i^j(w); w), j = \{N, M\} \text{ and} \\ \int_{w_0}^{w_1} (Y_i^N(w) - v(Y_i^N(w); w) - U_i^N(w)) \varphi_i^N(U_i^N(w) - U_{-i}^M(w); w) dw \\ + \int_{w_0}^{w_1} (Y_i^M(w) - v(Y_i^M(w); w) - U_i^M(w)) \varphi_i^M(U_i^M(w) - U_{-i}^N(w); w) dw \geq 0.$$

Optimality Conditions

We use the dual problem to characterize best responses, namely:

$$(19) \quad \max \int_{w_0}^{w_1} (Y_i^N(w) - v(Y_i^N(w); w) - U_i^N(w)) \varphi_i^N(U_i^N(w) - U_{-i}^M(w); w) dw \\ + \int_{w_0}^{w_1} (Y_i^M(w) - v(Y_i^M(w); w) - U_i^M(w)) \varphi_i^M(U_i^M(w) - U_{-i}^N(w); w) dw \\ \text{s.t. } U_i^j(w) = -v'_w(Y_i^j(w); w), j = \{N, M\} \\ \text{and } U_i^N(w_0) \geq \underline{U}_i^N.$$

In a given country, the objective function consists therefore of two separate parts: taxes collected from the National residents, on the one hand, and taxes collected from the Mercenaries, on the other hand. This situation is analogous to that encountered when a duopoly designs non-linear price schedules offered to two distinct set of customers. In a best response, the price schedule offered to a set is independent to that offered to the other set of customers. This is through the best response of the other firm that there is a link between the two price schedules a given firm implements.

When solving for optimization problem (19), we treat $Y_i^N(w)$ and $Y_i^M(w)$ as control variables, while $U_i^N(w)$ and $U_i^M(w)$ are state variables. Denoting the co-state variables associated with the incentive compatibility conditions (14) and (15) by $q_i^N(w)$ and $q_i^M(w)$ respectively, the Hamiltonian corresponding to Problem (19) may be written as:

$$(20) \quad \mathcal{H} = \left(Y_i^N(w) - v(Y_i^N(w); w) - U_i^N(w) \right) \varphi_i^N(U_i^N(w) - U_{-i}^M(w); w) \\ + \left(Y_i^M(w) - v(Y_i^M(w); w) - U_i^M(w) \right) \varphi_i^M(U_i^M(w) - U_{-i}^N(w); w) \\ - q_i^N(w) v'_w(Y_i^N(w); w) - q_i^M(w) v'_w(Y_i^M(w); w)$$

The following Proposition summarizes the associated optimality conditions.

PROPOSITION 1. The optimality conditions are:

$$(21) \quad 0 = (1 - v'_Y(Y_i^N(w); w))\varphi_i^N(U_i^N(w) - U_{-i}^M(w); w) - q_i^N(w)v''_{yw}(Y_i^N(w); w)$$

$$(22) \quad 0 = (1 - v'_Y(Y_i^M(w); w))\varphi_i^M(U_i^M(w) - U_{-i}^N(w); w) - q_i^M(w)v''_{yw}(Y_i^M(w); w)$$

$$(23) \quad [q_i^N]'(w) = \varphi_i^N(U_i^N(w) - U_{-i}^M(w); w) - T_i^N(Y_i^N(w)) \frac{\partial \varphi_i^N(U_i^N(w) - U_{-i}^M(w); w)}{\partial U_i^N}$$

$$(24) \quad [q_i^M]'(w) = \varphi_i^M(U_i^M(w) - U_{-i}^N(w); w) - T_i^M(Y_i^M(w)) \frac{\partial \varphi_i^M(U_i^M(w) - U_{-i}^N(w); w)}{\partial U_i^M}$$

$$(25) \quad q_i^N(w_0) \leq 0$$

$$(26) \quad q_i^N(w_1) = 0 \text{ when } w_1 < \infty \text{ and } q_i^N(w_1) \rightarrow 0 \text{ when } w_1 \rightarrow \infty$$

$$(27) \quad q_i^M(w_0) = 0$$

$$(28) \quad q_i^M(w_1) = 0 \text{ when } w_1 < \infty \text{ and } q_i^M(w_1) \rightarrow 0 \text{ when } w_1 \rightarrow \infty$$

III.2. Best Response Taxation

In a best-response, U_{-i}^N and U_{-i}^M are exogenously given. Consequently, National taxation in country i is determined by equations (21), (23), (25) and (26), while Mercenary taxation depends on equations (22), (24), (27) and (28). These two sets of equations are independent; it is therefore possible to decompose the best-response problem into two independent sub-problems: the first one determines best-response National Taxation and the second one best-response Mercenary Taxation. We now investigate these independent sub-problems in turn.³

National Taxation

We now turn to the characterization of the tax schedule country i offers to its national residents in a best response. Rearranging (21), we see that:

$$(29) \quad 1 - v'_Y(Y_i^N(w); w) = \frac{q_i^N(w)}{\varphi_i^N(\Delta_i; w)} v''_{yw}(Y_i^N(w); w).$$

This expression is always well defined because $\varphi_i^N(\Delta_i; w) > 0$ at every w as explained above. Combining the Euler condition (23) and the transversality condition (26), we obtain:

$$(30) \quad \begin{aligned} q_i^N(w) &= q_i^N(w_1) - \int_w^{w_1} [q_i^N]'(x) dx \\ &= \int_w^{w_1} \left[T_i^N(Y_i^N(x)) \frac{\partial \varphi_i^N(\Delta_i; x)}{\partial U_i^N} - \varphi_i^N(\Delta_i; x) \right] dx. \end{aligned}$$

³This separability into two independent subproblems is not due to the maximin social objective; but to the asymmetric tax treatment of Native and Foreign residents in any given country.

We define:

$$\begin{aligned}
X_i^N(w) &\equiv -q_i^N(w) = \int_w^{w_1} \left[\varphi_i^N(\Delta_i; x) - T_i^N(Y_i^N(x)) \frac{\partial \varphi_i^N(\Delta_i; x)}{\partial U_i^N} \right] dx \\
&= \int_w^{w_1} \left[1 - \frac{T_i^N(Y_i^N(x))}{\varphi_i^N(\Delta_{-i}; x)} \frac{\partial \varphi_i^N(\Delta_i; x)}{\partial U_i^N} \right] \varphi_i^N(\Delta_i; x) dx \\
(31) \quad &= \int_w^{w_1} \left[1 - T_i^N(Y_i^N(x)) \eta_i^N(\Delta_i; x) \right] \varphi_i^N(\Delta_i; x) dx.
\end{aligned}$$

Note that, if $\Delta_i(x) \geq 0$ for any skill $x \geq w$, then $X_i^N(w) = N_i(1 - H_i(w))$.

PROPOSITION 2. In a best response, marginal tax rates faced by National agents staying in country $i = \{A, B\}$ are given by:

$$(32) \quad \frac{MRT_i^N(Y_i^N(w))}{1 - MTR_i^N(Y_i^N(w))} = \frac{\alpha_i^N(w)}{\varepsilon_i^N(w)} \frac{X_i^N(w)}{w \varphi_i^N(\Delta_i(w); w)}.$$

Proof. We first substitute (31) into (29) and recall that marginal tax rates are defined by (3). We then compute the ratio $\alpha_i^N(w)/\varepsilon_i^N(w)$ to obtain:

$$v''_{yw}(Y_i^N(w); w) = - \frac{1 - MTR_i^N(Y_i^N(w))}{w} \frac{\alpha_i^N(w)}{\varepsilon_i^N(w)}.$$

Plugging the latter into the former expression,

$$MTR_i^N(Y_i^N(w)) = \frac{q_i^N(w)}{\varphi_i^N(\Delta_i; w)} v''_{yw}(Y_i^N(w); w) = \frac{X_i^N(w)}{\varphi_i^N(\Delta_i; w)} \frac{1 - MTR_i^N(Y_i^N(w))}{w} \frac{\alpha_i^N(w)}{\varepsilon_i^N(w)}.$$

Rearranging, we obtain (2). □

Mercenary Taxation

We now turn to the tax schedule country i offers to the people it wants to attract as Mercenaries in a best response. Rearranging (22), we see that:

$$(33) \quad [1 - v'_Y(Y_i^M(w); w)] \varphi_i^M(-\Delta_{-i}; w) = q_i^M(w) v''_{yw}(Y_i^M(w); w).$$

Combining the Euler condition (24) and the transversality condition (28), we obtain:

$$\begin{aligned}
q_i^M(w) &= q_i^M(w_1) - \int_w^{w_1} q_i^M(x) dx \\
(34) \quad &= \int_w^{w_1} \left[T_i^M(Y_i^M(x)) \frac{\partial \varphi_i^M(-\Delta_{-i}; x)}{\partial U_i^M} - \varphi_i^M(-\Delta_{-i}; x) \right] dx
\end{aligned}$$

We define:

$$\begin{aligned}
X_i^M(w) &\equiv -q_i^M(w) = \int_w^{w_1} \left[\varphi_i^M(-\Delta_{-i}; x) - T_i^M(Y_i^M(x)) \frac{\partial \varphi_i^M(-\Delta_{-i}; x)}{\partial U_i^M} \right] dx \\
&= \int_w^{w_1} \left[1 - \frac{T_i^M(Y_i^M(x))}{\varphi_i^M(-\Delta_{-i}; x)} \frac{\partial \varphi_i^M(-\Delta_{-i}; x)}{\partial U_i^M} \right] \varphi_i^M(-\Delta_{-i}; x) dx \\
&= \int_w^{w_1} \left[1 + \frac{T_i^M(Y_i^M(x))}{\varphi_i^M(-\Delta_{-i}; x)} \frac{\partial \varphi_i^M(-\Delta_{-i}; x)}{\partial \Delta_{-i}} \right] \varphi_i^M(-\Delta_{-i}; x) dx \\
(35) \quad &= \int_w^{w_1} \left[1 + T_i^M(Y_i^M(x)) \eta_i^M(\Delta_{-i}; x) \right] \varphi_i^M(\Delta_{-i}; x) dx
\end{aligned}$$

We can now substitute (35) into (33) and rearrange along the same lines as in the proof of Proposition 2 to obtain the following result.

PROPOSITION 3. In a best response, marginal tax rates faced in country $i = \{A, B\}$ by foreign Mercenaries are given by:

$$(36) \quad \frac{MTR_i^M(Y_i^M(w))}{1 - MTR_i^M(Y_i^M(w))} = \frac{\alpha_i^M(w)}{\varepsilon_i^M(w)} \frac{X_i^M(w)}{w \varphi_i^M(\Delta_{-i}(w); w)} \text{ if } \varphi_i^M(\Delta_{-i}(w); w) > 0,$$

and $MTR_i^M(Y_i^M(w)) = +\infty$ otherwise.

Given (36), the marginal tax rates $MTR_i^M(Y_i^M(w))$ faced by Mercenaries in country i are only well defined at w for which some agents actually move from country $-i$ to country i (i.e., for which $\varphi_i^M(\Delta_{-i}(w); w) > 0$). At skill levels w for which $\varphi_i^M(\Delta_{-i}(w); w) = 0$ instead, it is not desirable for country i to attract foreigners; without any loss in generality, we therefore set $MTR_i^M(Y_i^M(w)) = +\infty$ to capture this fact.

IV. THE TIEBOUT BEST

To gain further insights into the second-best optimal tax policy, we investigate the situation in which each policymaker observes w but is unable to recover the value of the migration cost m faced by a given individual. We refer to this information setting as the ‘‘Tiebout best’’ in honor of Charles M. Tiebout’s seminal contribution to the analysis of migration in the field of public finance.

IV.1. Best-Response Taxation

In the Tiebout best and for any i and j , the marginal tax rate $MTR_i^j(Y_i^j(w))$ is equal to zero at each w . This implies $X_i^j(w) = 0$. To insist on the fact that the tax schedule directly depends on w , we in this section write $\tilde{T}_i^j(w)$ to denote the tax liability. Using (31) and (35), the following result is obtained.

PROPOSITION 4. In the Tiebout best, and in a best response, tax liabilities in country $i = \{A, B\}$ are given by:

$$(37) \quad \tilde{T}_i^N(w) = \frac{1}{\eta_i^N(\Delta_i; w)} = \frac{1 - G_i(-\Delta_i(w)|w)}{g_i(-\Delta_i(w)|w)} \text{ for every } w > w_0,$$

with an upward jump-discontinuity at $w = w_0$, and

$$(38) \quad \tilde{T}_i^M(w) = -\frac{1}{\eta_i^M(\Delta_{-i}(w); w)} = \frac{G_{-i}(-\Delta_{-i}(w)|w)}{g_{-i}(-\Delta_{-i}(w)|w)} \text{ for every } w > w_0.$$

The tax liabilities paid in country i by the National residents (with $w > w_0$) and the foreign Mercenaries are respectively equal to the inverse of the semi-elasticity of emigration and to the inverse of the semi-elasticity of expatriation. The National residents with $w = w_0$ then receive a lump-sum transfer which equally splits collected taxes among them. There is therefore a discontinuity in the Tiebout-best National tax schedule at w_0 . By contrast, the Tiebout-best Mercenary schedule is continuous and defined for every w .⁴ When policymakers are not allowed to differentiate taxes based on citizenship, as in Lehmann et al. (2014), any agent of skill w staying in country i pays a tax liability equal to the inverse of the average semi-elasticity of migration computed for National citizens *as well as* resident Foreigners.

IV.2. General Features of Any Nash Equilibrium

We first establish that, in any Nash equilibrium, Native and Foreign residents face different tax schedules.

PROPOSITION 5. In the Tiebout best, at any Nash equilibrium, each country has an incentive to offer differentiated tax schedules to National and Foreign residents.

Proof. The proof is by contradiction. Let us suppose that there is a Nash equilibrium in which neither country uses tax differentiation. Hence, $\tilde{T}_A^N(w) = \tilde{T}_A^M(w)$ and $\tilde{T}_B^N(w) = \tilde{T}_B^M(w)$ at any w . Given the social objective in country i , there are at least some w at which $\tilde{T}_i^N(w) < 0$ and some other w at which $\tilde{T}_i^N(w) > 0$, $i = \{A, B\}$. There are therefore four cases to consider.

- (i) $\tilde{T}_i^N(w) < 0$ and $\tilde{T}_{-i}^N(w) < 0$ at a given w (possibly with $\tilde{T}_i^N(w) = \tilde{T}_{-i}^N(w)$). Setting $\tilde{T}_k^M(w) = 0$ is clearly welfare improving in each country, $k = \{i, -i\}$. A contradiction.
- (ii) $\tilde{T}_i^N(w) < 0$ and $0 < \tilde{T}_{-i}^N(w)$ at a given w . Then, the same argument as in Case 1 applies to country i .
- (iii) $0 < \tilde{T}_i^N(w) < \tilde{T}_{-i}^N(w)$ at a given w . Then, there are no foreign taxpayer of skill w in country $-i$. Now let us consider that country $-i$ diminishes $\tilde{T}_{-i}^M(w)$ and set $\tilde{T}_{-i}^M(w) = \tilde{T}_i^N(w) - \epsilon$, with $0 <$

⁴Because $\tilde{T}_{-i}(w_0) > 0$ while $\tilde{T}_i(w)$ is necessarily negative at every w , there is no Mercenary of skill w_0 in either country.

$\epsilon < \tilde{T}_i^N(w)$. This attracts a positive measure of taxpayers, moving from country i to country $-i$, and increases collected taxes. This extra amount of taxes can be redistributed to the worst-off National citizens in a lump-sum manner, which increases $U_{-i}^N(w_0)$. A contradiction.

(iv) $0 < \tilde{T}_i^N(w) = \tilde{T}_{-i}^N(w)$ at a given w . In that case, no taxpayer decides to move abroad. We can apply the same argument as in Case 1 to either country: given $\tilde{T}_{-i}^N(w)$, country i has an incentive to deviate and offer $\tilde{T}_i^M(w) = \tilde{T}_{-i}^N(w) - \epsilon$, with $0 < \epsilon < \tilde{T}_{-i}^N(w)$. Consequently, $\tilde{T}_k^N(w) = \tilde{T}_k^M(w)$ is not a best response, for $k = \{i, -i\}$. \square

We now provide additional features that must be verified by any Nash equilibrium.

LEMMA 1. In the Tiebout best, at any Nash equilibrium, the lowest skilled Natives of country i , with $w = w_0$, remain in country i . In addition, the National tax schedule has an upward jump discontinuity at w_0 , with $T_i^N(w_0) < 0$ while $T_i^N(w) > 0$ for any $w > w_0$.

Proof. The lowest skilled Natives of country i receive a net transfer in country i , i.e., $\tilde{T}_i^N(w_0) < 0$. Mercenary taxation in country $-i$ proceeds from the maximization of tax receipts on the set of Mercenaries. Consequently, $\tilde{T}_{-i}^M(w_0) > 0$. This implies: $U_i^N(w_0) > U_{-i}^M(w_0)$. \square

We call \hat{w}_i the infimum productivity of the set of Mercenaries in country i . The following Lemma establishes that there are Mercenaries at each productivity level, except at w_0 .

LEMMA 2. In the Tiebout best, at any Nash equilibrium, $\hat{w}_i = w_0$.

Proof. Suppose there is a Nash equilibrium for which $\varphi_i^M(w) = 0$ for any $w > w_0$ and $i = \{A, B\}$. For every $w > w_0$, $\tilde{T}_i^N(w) > 0$. Because m has support \mathbb{R}^+ , it is always possible for country i to offer $\tilde{T}_i^M(w) = \tilde{T}_{-i}^N(w) - \epsilon$, with $\epsilon > 0$, and attract additional net taxpayers in country i . This increases collected taxes. A contradiction. \square

LEMMA 3. If η_{-i}^N is non-increasing in w , then φ_i^M is non-decreasing in w .

Proof. φ_i^M is an implicit function of \tilde{T}_{-i}^N , with $\varphi_i^M = \varphi_i^M(\Delta_{-i}(\tilde{T}_{-i}^N))$. By the chain rule, $\partial\varphi_i^M/\partial\tilde{T}_{-i}^N = (\partial\varphi_i^M/\partial\Delta_{-i}) \cdot (\partial\Delta_{-i}/\partial\tilde{T}_{-i}^N)$. We already know that $\partial\varphi_i^M/\partial\Delta_{-i} \leq 0$. In addition, by definition of Δ_{-i} , we have:

$$(39) \quad \Delta_{-i} = Y_{-i}^N(w) - \tilde{T}_{-i}^N(Y_{-i}^N(w)) - v(Y_{-i}^N(w); w) - Y_i^M(w) + \tilde{T}_i^M(Y_i^M(w)) + v(Y_i^M(w); w).$$

Therefore, $\partial\Delta_{-i}/\partial\tilde{T}_{-i}^N = -1$. Hence, $\partial\varphi_i^M/\partial\tilde{T}_{-i}^N \geq 0$. It is then sufficient to note that, by Proposition 4, \tilde{T}_{-i}^N non-decreasing in w is equivalent to η_{-i}^N non-increasing in w . \square

IV.3. Existence and Uniqueness of the Nash Equilibrium

We now show that there is a Nash equilibrium in the Tiebout best, under very weak conditions, and provide sufficient conditions under which it is unique.

PROPOSITION 6. Assume the distribution of the migration costs $g_i(m|w)$ has a finite expectation $\bar{\gamma}(w)$ at any w for $i = \{A, B\}$. Then, there exists a Nash equilibrium in the Tiebout best.

Proof. From Lemma 2, we know that, in any Tiebout-best Nash equilibrium, there are Mercenaries in both countries at any $w > w_0$. Given the definition of φ_i^M provided in Equation (9), a positive fraction of Mercenaries requires $\Delta_i(w) \leq 0$. Because preferences are quasilinear, linear in net income, we have:

$$(40) \quad \Delta_i(w) = \tilde{T}_{-i}^M(w) - \tilde{T}_i^N(w).$$

The existence of a Nash equilibrium boils down to finding $\Delta_i(w) \leq 0$, $i = \{A, B\}$ solution to the latter equation. Using the expressions in Proposition 4, Equation (40) is equivalent, at any w , to $\phi_i(\Delta_i|w) = 1$, with:

$$(41) \quad \phi_i(\Delta_i|w) = 2G_i(-\Delta_i|w) - \Delta_i g_i(-\Delta_i|w).$$

For $\Delta_i = 0$, $\phi_i(\Delta_i|w) = 0$. For $\Delta_i \rightarrow -\infty$, $\phi_i(\Delta_i|w) = 2 + \bar{\gamma}(w)$ given that $g_i(m|w)$ has finite expectation at any w . By continuity of $g_i(m|w)$, there is therefore a solution to $\phi_i(\Delta_i|w) = 1$ at any w , so that the Nash equilibrium exists. \square

Based on the above Proposition, we may provide sufficient conditions under which the Nash equilibrium is actually unique. These conditions employ the concept of inverted star-shapedness applied to the conditional probability density functions $g_i(m|w)$ of the migration costs.

COROLLARY 1. For $i = \{A, B\}$, assume $g_i(m|w)$ (i) has a finite expectation at any w , (ii) is differentiable with respect to w , and (iii) is inverted star-shaped in the sense that

$$g_i(m|w) > m \cdot \frac{\partial g_i(m|w)}{\partial m}$$

at any w . Then, the Nash equilibrium is unique in the Tiebout best.

Proof. Consider $\phi_i(\Delta_i|w)$ defined in (41). Computing its derivative with respect to Δ_i , we obtain: $\partial \phi_i(\Delta_i|w) / \partial \Delta_i = g_i(\Delta_i|w) - \Delta_i \cdot \partial g_i(\Delta_i|w) / \partial \Delta_i$, which is strictly positive because g_i is inverted star-shaped. Consequently, for any given w , there is a unique positive solution in Δ_i to $\phi_i(\Delta_i|w) = 1$. \square

IV.4. *Symmetric Countries: An Example*

We now focus on the situation in which both countries are symmetric. If there is a Nash equilibrium, then it must be such that $\tilde{T}_i^N(w) = \tilde{T}_{-i}^N(w)$, $\tilde{T}_i^M(w) = \tilde{T}_{-i}^M(w)$, $\eta_i^N = \eta_{-i}^N$ and $\eta_i^M = \eta_{-i}^M$ at any w . We therefore in this subsection drop the i and $-i$ subscripts. Without any loss of generality, we normalize the population in each country N to 1.

To gain further insights, we specify the distribution of migration costs. For simplicity, we assume that it is described by an exponential distribution:

$$(42) \quad g(m|w) = \lambda(w) \exp(-m\lambda(w)) \text{ and } G(m|w) = 1 - \exp(-m\lambda(w)),$$

where $\lambda(w) > 0$ is a \mathcal{C}^1 scale function. Then, $g(m|w) > m \cdot \partial g_m(m|w)/\partial m$ at any w is equivalent to: $\lambda(w) > -m\lambda^2(w)$. This inequality is always satisfied given that $m \geq 0$ and $\lambda(w) > 0$. The inverted-shapedness property is thus satisfied.

By definition of the semi-elasticity of emigration given in (10), we obtain: $\eta^N(w) = \lambda(w)$. In order to satisfy the condition under which Lemma 3 applies, we need to verify $\lambda'(w) \leq 0$. For illustrative purposes, we consider $\lambda(w) = 1/(1+w)$. The assumptions of Proposition 6 and Corollary 1 therefore hold: there exists a Nash equilibrium and it is unique. It is easy to show that:

- for any w , $\tilde{T}^N(w) = 1 + w$;
- for $w > w_0$, $\tilde{T}^M(w) = (1+w) \exp(-\Delta(w)/(1+w)) [1 - \exp(\Delta(w)/(1+w))]$;
- $\Delta(w) \simeq -0,44(1+w)$, from which $\tilde{T}^M(w) \simeq 0,56(1+w)$ and $\eta^M(w) \simeq -1,79/(1+w)$;
- the “number” of Mercenaries of skill w is given by $\phi^M(\Delta(w); w) \simeq 0,36 \cdot h(w)$.

The w_0 natives always stay in their home country, where they receive a positive lump-sum transfer $|\tilde{T}^N(w_0)|$. In each country, approximately 36% of the initial population –with the lowest migration costs– move abroad to benefit from the Mercenary tax schedule $\tilde{T}^M(w)$; the rest of the population stays at home and pays $\tilde{T}^N(w)$ in taxes.

Let us now consider that w is uniformly distributed over $[0,01, 10]$. In a given country, collected taxes from National residents amount to 1.8 and collected taxes from Mercenaries to 1.0. Therefore, 2.8 are redistributed to the w_0 Native residents. It is interesting to contrast these results with those obtained in a world in which each country’s policymaker is unable to differentiate taxes between Native and Foreign residents. This situation was examined in Lehmann et al. (2014). Contrary to the present setting, no agent moves in the Tiebout-best Nash equilibrium; each agent with $w > w_0$ pays taxes to her home government, equal to w . Hence, the introduction of tax breaks for foreigners results in social welfare being reduced by about 16% in each country.

IV.5. Asymmetric Countries: An Example

We now turn to asymmetric countries. One of the simplest way to introduce heterogeneity between countries is to use the same setting as in the above example, but assume that $\lambda_A \neq \lambda_B$. All other parameters are identical in both countries. In the numerical simulations, we consider that $\lambda_A(w) = 1/(1+w)$ and $\lambda_B(w) = 2/(1+w)$. Consequently, $\tilde{T}_A^N(w) = 1+w$ and $\tilde{T}_B^N(w) = (1+w)/2$: at a given productivity level w , the agents staying in their home country face a tax liability which is twice larger in A than in B . The assumptions of Proposition 6 and Corollary 1 hold: there exists a Nash equilibrium and it is unique.

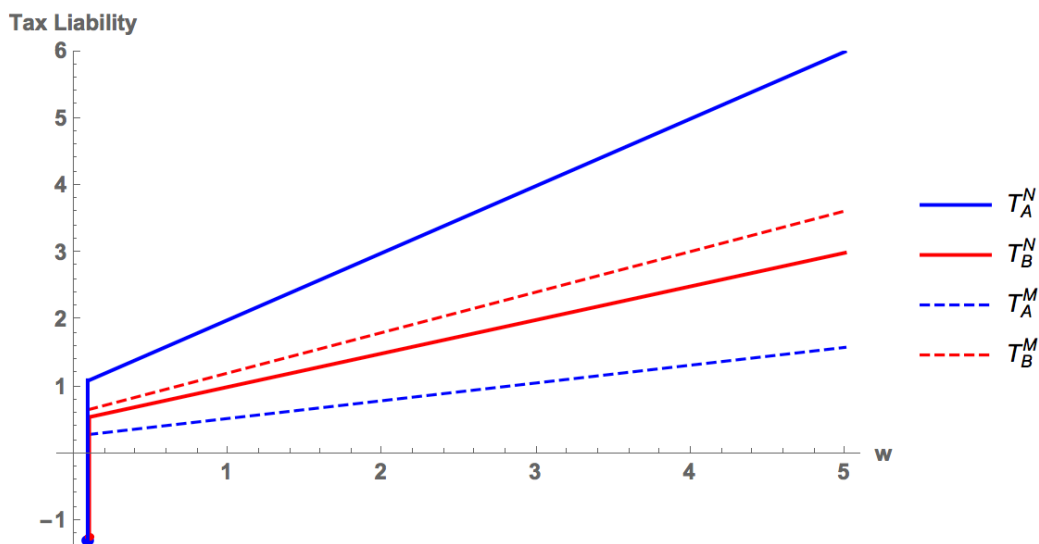


FIGURE I: EXAMPLE OF NASH EQUILIBRIUM IN THE TIEBOUT BEST (ASYMMETRIC COUNTRIES)

For $w > w_0$, we have: $\tilde{T}_A^M(w) = 0.5(1+w) \exp(-2\Delta_B(w)/(1+w)) [1 - \exp(2\Delta_B(w)/(1+w))]$ and $\tilde{T}_B^M(w) = (1+w) \exp(-\Delta_A(w)/(1+w)) [1 - \exp(\Delta_A(w)/(1+w))]$. We then compute $\tilde{T}_i^M(w) - \tilde{T}_{-i}^N(w)$ and equate it to $\Delta_{-i}(w)$ for $i = \{A, B\}$. We obtain two equations that we solve in $\Delta_A(w)$ and $\Delta_B(w)$. We check that $\varphi_i^M(w)$ is positive for any $w > 0$ before computing $\tilde{T}_A^M(w)$ and $\tilde{T}_B^M(w)$. The numerical results are shown in Figure I for $w \in [0.01, 10]$. We now further assume that $h(w)$ is uniformly distributed over $[0, 10]$ and normalize the size of the initial population in each country to $N_A = N_B = 1$. The partition of the population between National Residents and Mercenaries is then as follows:

Countries	A	B	Total
National Residents	0.673	0.625	1.298
Mercenaries	0.375	0.327	0.702
Total	1.048	0.952	2

V. THE SECOND BEST

We now investigate the second best in which both skills and migration costs are private information. To save on notations, we from now on denote the skill densities of taxpayers and the semi-elasticities in the Nash equilibrium by $f_{i,j}^* \equiv \varphi_i^j$ and $\eta_{i,j}^*$ respectively, $j = \{N, M\}$.

V.1. National Taxation

The Tiebout-best tax schedule provides insights into the second-best solution, where both skills and migration costs are private information. Using (37), Equation (31) can be rewritten as:

$$(43) \quad X_i^N(w) = \int_w^{w_1} \left[\tilde{T}_i^N(w) - T_i^N(Y_i^N(x)) \right] \eta_{i,N}^*(x) f_{i,N}^*(x) dx$$

Hence, for National residents, $X_i^N(w)$ is a weighted sum of the difference between the Tiebout-best tax liabilities and second-best tax liabilities for all skill levels x above w . The weights are given by the product of the semi-elasticity of emigration (from country i to country $-i$) and the skill density of country i 's Native population, i.e. by the mass of pivotal Native individuals of skill w , who are indifferent between migrating or not.

In the Tiebout best, the tax liability schedule consists of a net transfer to the w_0 -National citizens, with an upward discontinuity at that point: all other National citizens pay strictly positive taxes. The Tiebout-best tax schedule defines a *target* for the policymaker in the second best, where distortions along the intensive margin have also to be minimized. The second-best solution thus proceeds from the reconciliation of three underlying forces: *i*) maximizing the welfare of the worst-off National citizens; *ii*) being as close as possible to the Tiebout-best tax liability to limit the distortions stemming from the migration responses; *iii*) being as flat as possible to mitigate the distortions coming from the intensive margin. In the second-best, these three goals cannot be pursued independently because of the incentive-compatibility constraints (14).

PROPOSITION 7. In a Nash equilibrium:

- i) if $\eta_{i,N}^{*'}(\cdot) = 0$, then $MTR_i^N(Y_i^N(w)) > 0$ and $T_i^N(Y_i^N(w)) < 1/\eta_{i,N}^*$ for all $w \in (w_0, w_1)$;
- ii) if $\eta_{i,N}^{*'}(\cdot) < 0$, then $MTR_i^N(Y_i^N(w)) > 0$ for all $w \in (w_0, w_1)$;
- iii) if $\eta_{i,N}^{*'}(\cdot) > 0$, then either
 - (a) $MTR_i^N(Y_i^N(w)) \geq 0$ for all $w \in (w_0, w_1)$;
 - (b) or there exists a threshold $\check{w} \in [w_0, w_1)$ such that $MTR_i^N(Y_i^N(w)) \geq 0$ for all $w \in (w_0, \check{w})$ and $MTR_i^N(Y_i^N(w)) < 0$ for all $w \in (\check{w}, w_1)$.
- iv) if $\eta_{i,N}^{*'}(\cdot) > 0$ and $\lim_{w \rightarrow \infty} \eta_{i,N}^*(w) = +\infty$, then there exists a threshold $\check{w} \in (w_0, w_1)$ such that $MTR_i^N(Y_i^N(w)) \geq 0$ for all $w \in (w_0, \check{w})$ and $MTR_i^N(Y_i^N(w)) < 0$ for all $w \in (\check{w}, w_1)$.

Proof. Because of the separability of the social planner's problem into two independent sub-problems, this proof is a direct adaptation of Proposition 3 in Lehmann et al. (2014). \square

This proposition casts light on the part played by the slope of the semi-elasticity of emigration. It considers the three natural benchmarks that come to mind when thinking about it. First, the costs of migration may be independent of w as in Blumkin et al. (2012) and Morelli et al. (2012), implying a constant semi-elasticity in a symmetric equilibrium. This makes sense, in particular,

if most relocation costs are material (moving costs, flight tickets, etc.).⁵ Second, one might want to consider a constant elasticity of emigration, as in Brewer et al. (2010) and Piketty and Saez (2012). In this case, the semi-elasticity must be decreasing: 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. Third, 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 emigration may be increasing in skills. A special case is investigated in Simula and Trannoy (2010 and 2011), with a semi-elasticity equal to zero up to a threshold, and infinite above it.

V.2. Mercenary Taxation

We now turn to the characterization of the second-best Mercenary tax schedule. We first establish that the least productive Nationals from country i never have an incentive to move abroad and settle down in country $-i$. This, in particular, implies that there is no Nash equilibrium with Mercenaries of skill w_0 .

LEMMA 4. In any Nash equilibrium, the lowest skilled Natives of country i , with $w = w_0$, remain in country i .

Proof. The lowest skilled Natives of country i receive a net transfer in country i , i.e., $T_i^N(Y_i^N(w_0)) < 0$. Mercenary taxation in country $-i$ proceeds from the maximization of tax receipts on the set of Mercenaries. Consequently, $T_{-i}^M(Y_{-i}^M(w_0)) > 0$. Consequently, $U_i^N(w_0) > U_{-i}^M(w_0)$. \square

We now show that, if the National tax schedule offered in country $-i$ increases in skills, the proportion of Mercenaries in country i increases in w .

LEMMA 5. If T_{-i}^N is non-decreasing in w , φ_i^M is non-decreasing in w .

Proof. φ_i^M is an implicit function of T_{-i}^N , with $\varphi_i^M = \varphi_i^M(\Delta_{-i}(T_{-i}^N))$. By the chain rule, $\partial\varphi_i^M/\partial T_{-i}^N = (\partial\varphi_i^M/\partial\Delta_{-i}) \cdot (\partial\Delta_{-i}/\partial T_{-i}^N)$. We already know that $\partial\varphi_i^M/\partial\Delta_{-i} \leq 0$. In addition, by definition of Δ_i , we have:

$$(44) \quad \Delta_{-i} = Y_{-i}^N(w) - T_{-i}^N(Y_{-i}^N(w)) - v(Y_{-i}^N(w); w) - Y_i^M(w) + T_i^M(Y_i^M(w)) + v(Y_i^M(w); w).$$

Therefore, $\partial\Delta_{-i}/\partial T_{-i}^N = -1$. Hence, $\partial\varphi_i^M/\partial T_{-i}^N \geq 0$. \square

This Lemma implies that, if $\varphi_i^M > 0$ at a given w' , then $\varphi_i^M > 0$ at any $w > w'$. As a result, in any Nash equilibrium, there is a threshold \hat{w} below which there is no Mercenary and above which there are Mercenaries at every w . Moreover, the greater \hat{w} the larger the proportion of Mercenaries. Combining both Lemmas with Proposition 3, we obtain:

⁵Morelli et al. (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.

PROPOSITION 8. If T_{-i}^N is non-decreasing in w , there is a \hat{w} such that:

- (i) for $w < \hat{w}$, $f_{i,M}^*(w) = 0$;
- (ii) for $w > \hat{w}$, $f_{i,M}^*(w) > 0$, $[f_{i,M}^*]'(w) > 0$, $\eta_{i,M}^* > 0$ and

$$(45) \quad \frac{MTR_i^M(Y_i^M(w))}{1 - MTR_i^M(Y_i^M(w))} = \frac{\alpha_i^M(w)}{\varepsilon_i^M(w)} \frac{X_i^M(w)}{w f_{i,M}^*(w)}$$

with

$$(46) \quad X_i^M(w) = \int_w^{w_1} [\tilde{T}_i^M(Y_i^N(x)) - T_i^M(x)] \eta_{i,M}^*(w) f_{i,M}^*(w) dx$$

By Proposition 7, we know that T_{-i}^N is non-decreasing in w if $\eta_{i,N}^{*'}(\cdot) = 0$ or $\eta_{i,N}^{*'}(\cdot) < 0$. Consequently, Proposition 8 holds in both cases.

V.3. Symmetric Countries

We focus in this subsection on the situation in which both countries are symmetric. We therefore can drop the i and $-i$ subscripts.

LEMMA 6. Consider both countries are symmetric. The following conditions must be verified in any Nash equilibrium:

- (i) $X^M(w) = \int_w^{w_1} [T^M(Y^M(w)) - \tilde{T}^M(w)] \eta_M^*(w) f_M^*(w)$;
- (ii) $X^M(w_0) = 0 \Leftrightarrow \int_{w_0}^{w_1} T^M(Y^M(w)) \eta_M^*(w) f_M^*(w) = \int_{w_0}^{w_1} \tilde{T}^M(w) \eta_M^*(w) f_M^*(w)$;
- (iii) $X^M(w_1) = 0$ when $w_1 < \infty$ and $X^M(w_1) \rightarrow 0$ when $w_1 \rightarrow \infty$;
- (iv) $dX^M(w)/dw = [\tilde{T}^M(w) - T^M(Y^M(w))] \eta_M^*(w) f_M^*(w)$.

Proof. Condition (i) is a definition. Condition (ii) follows from (i), and the transversality condition $q^M(w_0) = 0$, and $X^M(w) = -q^M(w)$. Condition (iii) follows from $X^M(w) = -q^M(w)$ and the transversality condition $q^M(w_1) = 0$ when $w_1 < \infty$ and $q^M(w_1) \rightarrow 0$ when $w_1 \rightarrow \infty$. Condition (iv) is obtained by differentiation of $X^M(w)$ defined in Condition (i). \square

Note that Condition (ii) implies that T^M either coincide with \tilde{T}^M or intersects it at least once. We look at situations in which η^N is increasing, constant or decreasing, and η^M is increasing, constant or decreasing. There are therefore 9 possible combinations. By definition of the semi-elasticities η^N and η^M , we have:

$$(47) \quad \eta^N(\Delta; w) = -\eta^M(\Delta; w) \frac{G(-\Delta|w)}{1 - G(-\Delta|w)}$$

Let α be defined as: $\alpha(w) = \frac{G}{1-G}$. Then, η'_n has the same sign as: $\eta'_M \alpha - \eta_M \alpha'$, with

$$\alpha' = \frac{G'(1-G) + GG'}{(1-G)^2} = \frac{G'}{(1-G)^2}$$

and

$$G' = \frac{\partial G}{\partial \Delta} \frac{d\Delta}{dw} + \frac{\partial G}{\partial w} = g\Delta' + G'_w.$$

Combining the above results, we obtain:

$$(1-G)^2 \eta'_n = (1-G)^2 \eta'_M \alpha - \eta_M (g\Delta' + G'_w)$$

where $G'_w \geq 0$.

Case 1: $G'_w \leq 0$. If the marginal tax rate for mercenaries is lower than the corresponding marginal tax rate for nationals (at a given w), then $\Delta' < 0$. We obtain:

- If $\eta'_n = 0$, then $\eta'_M < 0$;
- If $\eta'_n < 0$ then $\eta'_M < 0$;
- If $\eta'_n > 0$, then $\eta'_M \geq 0$.

Hence, only 5 cases are a priori possible within the 9 possible ones.

Case 2: $G'_w > 0$. Then: $(1-G)^2 \eta'_n + \eta_M (g\Delta' + G'_w) = (1-G)^2 \eta'_M \alpha$, so that 9 subcases are possible.

Note that $G'_w(m|w) = \int_0^m \frac{\partial g(x;w)}{\partial w} dx$. The assumption that this quantity increases with w for a given m implies that there are less agents with large migration costs when w goes up. This assumption seems to be in accordance with recent empirical evidence suggesting that top productive agents are highly responsive to differences in taxation when choosing where to locate. We therefore ignore Case 1 and focus on Case 2 below.

$G'_w > 0$ with η^N and η^M constant

We now consider that (i) $G'_w > 0$, (ii) the semi-elasticity of expatriation η^M is constant, (iii) the semi-elasticity of migration η^N is constant (and different from 0) and (iv) $w_1 \rightarrow \infty$.

- By definition of the tax liability effect $X^N(w)$, we have:

$$(48) \quad \frac{X^N(w)}{\eta^N} = \int_w^{w_1} [\tilde{T}^N(x) - T^N(Y(x))] \varphi^N(\Delta(x); x) dx = \int_w^{w_1} \left[\frac{1}{\eta^N} - T^N(Y(x)) \right] \varphi^N(\Delta(x); x) dx.$$

The transversality condition (26) is equivalent to:

$$(49) \quad \frac{X^N(w_1)}{\eta^N} \rightarrow 0 \text{ when } w_1 \rightarrow \infty.$$

By (48), the weighted mean of $1/\eta^N - T^N(Y(w))$, with weights given by $\varphi^N(\Delta(w); w)$, converges to 0 when w approaches ∞ . Moreover, by Proposition 7, $T^N(Y^N(w))$ is increasing in w in any Nash equilibrium, implying that $1/\eta^N - T^N(Y^N(w))$ decreases in w . Applying the contra-positive of Cesaro means theorem (for weighted averages), it follows that $1/\eta^N - T^N(Y^N(w))$ tends to 0 when w approaches ∞ , i.e.,

$$(50) \quad T^N(Y^N(w)) \rightarrow \tilde{T}^N(w) \text{ when } w \rightarrow \infty.$$

- Because countries are symmetric, we know that in any Nash equilibrium, $0 < \tilde{T}^M(w) < \tilde{T}^N(w)$ for all $w \in (w_0, w_1)$. Moreover, $T^N(Y^N(w))$ is continuous and increasing, from $T^N(Y^N(w)) < 0$ to $\tilde{T}^N(\infty)$. Hence, \tilde{T}^M intersects T^N only once, from above.
- \tilde{T}^M is desirable in the second-best. It is furthermore feasible because it is a non-personalized lump-sum transfer; given quasilinear preferences, the same quantity is added to the left- and the right-hand sides of the incentive-compatibility constraints. Hence, $T^M(Y(w)) = \tilde{T}^M(w)$ for all $w \in (w_0, w_1)$.
- Let w^* be the unique intersection of $\tilde{T}^M(w)$ and $T^N(Y(w))$. In each country, there are no Mercenaries for $w < w^*$; on the contrary, there are Mercenaries at each w for any $w > w^*$. For $w \in (w_0, w_1)$, $T^N(Y(w)) < \tilde{T}^M(w)$. Because $T^N - T^M = T^N - \tilde{T}^M < \tilde{T}^N - \tilde{T}^M$ for $w \in (w^*, \infty)$, there are fewer Mercenaries at each $w > w^*$ than in the Tiebout-best.

The following Proposition summarizes the results.

PROPOSITION 9. Consider a world with symmetric countries. Assume (i) $G'_w > 0$, (ii) the semi-elasticity of expatriation η^M is constant, (iii) the semi-elasticity of stayers η^N is constant (and different from 0) and (iv) $w_1 \rightarrow \infty$. Then, the unique Nash equilibrium is such that:

- $T^M(Y^M(w)) = \tilde{T}^M(w) = 1/\eta^M$;
- $T^N(Y^N(w))$ is increasing in w .

The Nash equilibrium tax schedules are illustrated in Figure II. Interpretation: Swiss lump-sum tax (forfait fiscal).

$G'_w > 0$ with η^N constant and η^M decreasing

We know that \tilde{T}^M and $T^N(Y^N(w))$ are increasing, starting at w_0 from a positive and negative value respectively. The only way for the two schedules not to intersect is that \tilde{T}^M tends to $T^N(Y^N(w))$ when w increases. In other words, $\eta(w)$ should tend to $-\eta(w)$ when $w \rightarrow \infty$. By (47), this is only possible if $G(-\Delta(w)|w)$ tends to $1/2$ when $w \rightarrow \infty$. We assume that this is not the case, so that \tilde{T}^M and $T^N(Y^N(w))$ intersect at least once. This assumption is reasonable in the second best because it corresponds to restrictions on the Tiebout-best schedules (which are then exogenous). We further

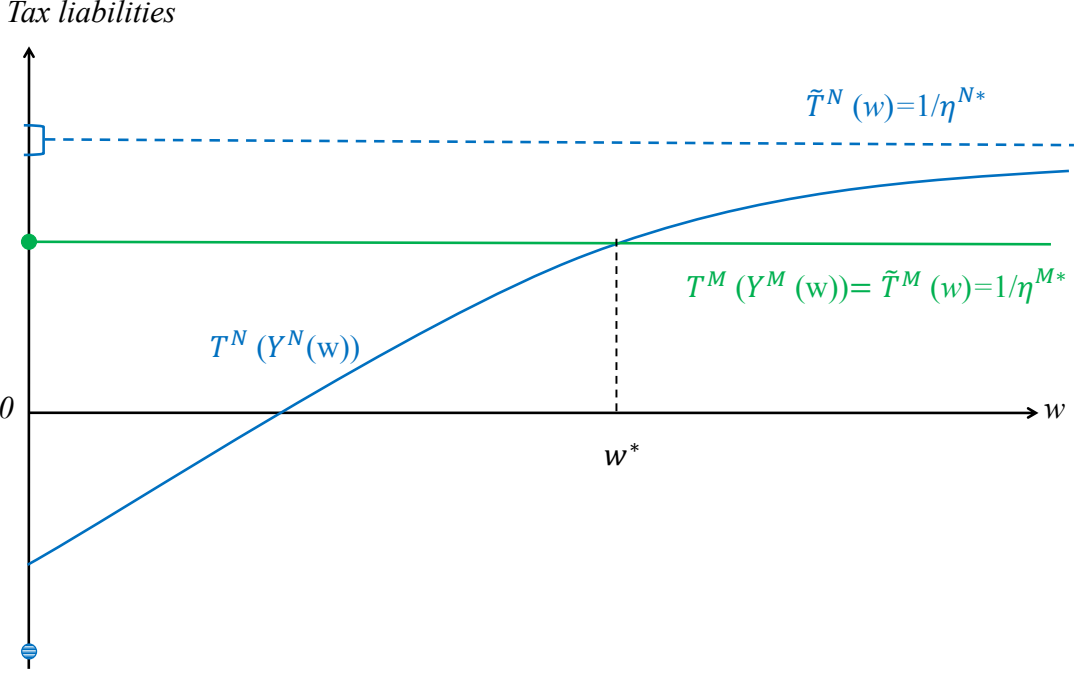


FIGURE II: NASH EQUILIBRIUM FOR SYMMETRIC COUNTRIES, $G'_w > 0$ AND CONSTANT SEMI-ELASTICITIES

know that $T^N \rightarrow \tilde{T}^N$ when $w \rightarrow \infty$. Consequently, $\eta^M(w)$ converges to $-\gamma\eta^N(w)$ with the scalar $0 < \gamma < 1$.

T^M must end below T^N and start above it. Consequently, there is an odd number of intersections (1 or larger). Proposition 8 applies. Therefore, there is a unique w^* from which there are Mercenaries. This implies that we cannot have 'sinusoidal' patterns and that the intersection is unique.

A vérifier: condition sur la moyenne pondérée de T^M et de \tilde{T}^M . A priori, il n'y a pas de raison pour que ça ne soit pas vérifié. On n'a pas de restrictions là-dessus.

Remarque: T^M tend vers \tilde{T}^M quand $w \rightarrow \infty$.

PROPOSITION 10. Consider a world with symmetric countries. Assume (i) $G'_w > 0$, (ii) the semi-elasticity of expatriation η^M is constant, (iii) the semi-elasticity of staying η^N is decreasing, (iv) $w_1 \rightarrow \infty$ and (v) $G(-\Delta(w)|w) < 1/2$ when $w \rightarrow \infty$. Then, the unique Nash equilibrium is such that:

- (i) $T^M(Y^M(w))$ is increasing;
- (ii) $T^M(Y^M(w_0)) = 0$ and $T^M(Y^M(w_1)) \rightarrow 0$ when $w_1 \rightarrow \infty$;
- (ii) $T^M(Y^M(w))$ has an inflection point;
- (iv) T^M intersects T^N for a productivity level w^* smaller than \bar{w} the intersection between T^M and \tilde{T}^M ;

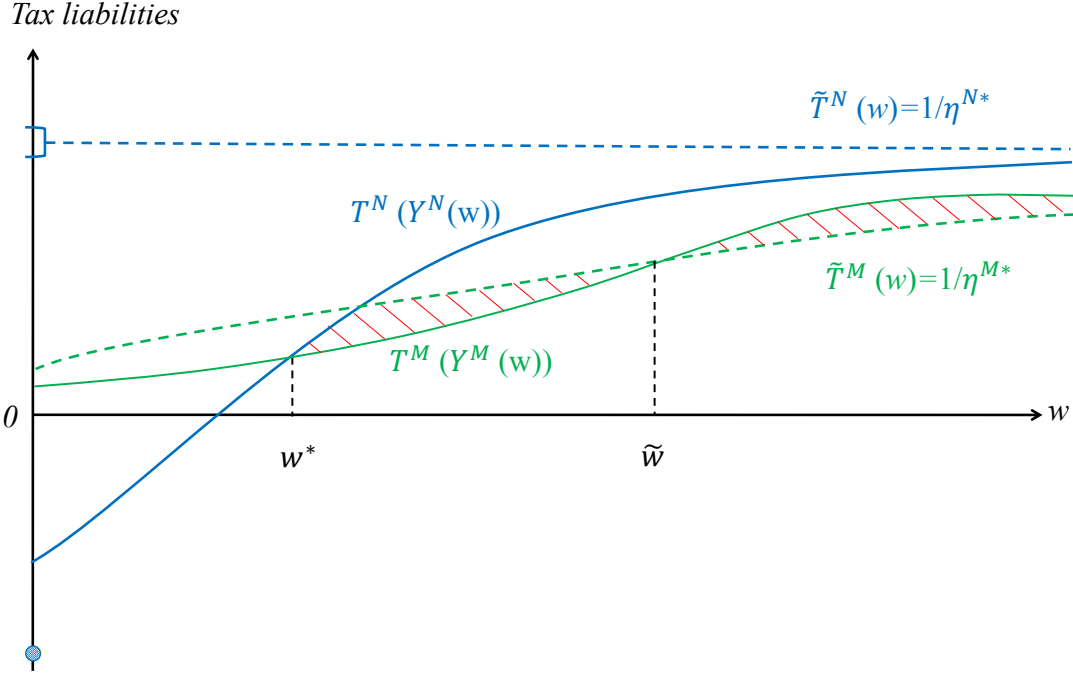


FIGURE III: NASH EQUILIBRIUM FOR SYMMETRIC COUNTRIES, $G'_w > 0$, η^N CONSTANT, AND η^M DECREASING

(v) T^M intersects T^N from below.

Proof. (i) T^M croissant: adaptation de QJE. (ii) Lemme 6. (iii) Conséquence de (i) et (ii). (iv) Si coupe après, il y a contradiction. La condition de moyenne ne peut pas être vérifiée. (v) Conséquence de (iv). Si coupe avant, il ne peut que couper par le bas. Coupe conséquence de (iv). \square

Voir si on ne peut pas montrer que T^M tendrait vers \tilde{T}^M en l'infini. Pas sûr pour l'instant.

$G'_w > 0$ with η^N constant and η^M increasing

On a décidé d'éliminer le cas où $\eta^M \rightarrow \infty$.

VI. Numerical Illustration

This section numerically implements the equilibrium optimal tax formula.

For simplicity, we consider that the world consists of two symmetric countries. The distribution of the skill levels is based on the CPS data (2007) extended by a Pareto tail, so that the top 1% of the population gets 18% of total income, as in the US. The disutility 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). We choose $\epsilon = 0.25$, which is a reasonable value based on the survey by Saez et al. (2012).

Even though the potential impact of income taxation on migration choices has been extensively discussed in the theoretical literature, there are still few empirical studies estimating the migration responses to taxation. A first set of studies consider the determinants of migration across US states (see Barro and Sala-i Martin (1992); Barro and Sala-I-Martin (1991), Ganong and Shoag (2013) and Suarez Serrato and Zidar (2013)). They find that per capita income has a positive effect on net migration rates into a state. This conclusion is entirely compatible with an explanation based on tax differences between US states, but may also be due to other differences (e.g., in productivities, housing rents, amenities or public goods). Strong structural assumptions are therefore required to disentangle the pure tax component. A second set of studies focuses exclusively on migration responses to taxation. Liebig et al. (2007) use differences across Swiss cantons and compute migration elasticities for different subpopulations, in particular for different groups in terms of education. Young and Varner (2011) use a millionaire tax specific to New Jersey. Because the salience of this millionaire tax is limited, their estimates of the causal effect of taxation on migration are not statistically significant, except for extremely specific subpopulations. Still, their results suggest that the elasticity of migration is increasing in the upper part of the income distribution. Only two studies are devoted to the estimation of migration elasticities between countries. Kleven et al. (2013) examine tax-induced mobility of football players in Europe 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 et al. (2014) confirm that these large estimates 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 however relatively small. Hence, these findings would translate into a global elasticity at the top of about 0.25 (see Piketty and Saez (2012)).

Our model pertains to international migrations and based on our survey of the empirical literature, we believe that the best we can do is to use an average elasticity of 0.25 for the national residents of the top 1% of the income distribution; and of 1 for the mercenaries.

We are presently working on the simulations, the results of which will be obtained within the next few weeks.

REFERENCES

- George A. Akerlof. The economics of “tagging” as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review*, 68(1):8–19, 1978.
- Robert J. Barro and Xavier Sala-I-Martin. Convergence across states and regions. *Brookings Papers on Economic Activity*, 1991(1):107–182, 1991.
- Robert J. Barro and Xavier Sala-i Martin. Convergence. *Journal of Political Economy*, 100(2):223–51, April 1992.

- Tomer Blumkin, Efraim Sadka, and Yotam Shem-Tov. International tax competition: Zero tax rate at the top re-established. Technical Report 3820, CESifo WP, May 2012.
- Mike Brewer, Emmanuel Saez, and Andrew Shephard. Means-testing and tax rates on earnings. In J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba, editors, *Dimensions of Tax Design: the Mirrlees Review*, chapter 3, pages 90–173. Oxford University Press, 2010.
- Peter A. Diamond. Optimal income taxation: An example with u-shaped pattern of optimal marginal tax rates. *American Economic Review*, 88(1):83–95, 1998.
- F. Docquier and A. Marfouk. International migration by education attainment, 1990-2000. In *International Migration, Remittances and the Brain Drain*, pages 151–200. World Bank, 2005.
- Frédéric Docquier and Abdeslam Marfouk. International migration by educational attainment (1990-2000) - release 1.1. In Calglar Ozden and Maurice Schiff, editors, *International Migration, Remittances and Development*, chapter 5. Palgrave Macmillan, New-York, 2006.
- Peter Ganong and Daniel Shoag. Why has regional income convergence in the u.s. declined? Technical report, HKS WP RWP12-028, 2013.
- Henrik J. Kleven, Camille Landais, and Emmanuel Saez. Taxation and international migration of superstars: Evidence from the european football market. *American Economic Review*, 103:1892–1924, 2013.
- Henrik J. Kleven, Camille Landais, Emmanuel Saez, and Esben A. Schultz. Migration and wage effects of taxing top earners: Evidence from the foreigners’ tax scheme in denmark. *Quarterly Journal of Economics*, 129:333–378, 2014.
- Etienne Lehmann, Laurent Simula, and Alain Trannoy. Tax me if you can! optimal nonlinear income tax between competing governments. *Quarterly Journal of Economics*, 129(4):1995–2030, 2014.
- Thomas Liebig, Patrick A. Puhani, and Alfonso Sousa-Poza. 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, 2007.
- James A. Mirrlees. Migration and optimal income taxes. *Journal of Public Economics*, 18:319–341, 1982.
- Massimo Morelli, Huanxing Yang, and Lixin Ye. Competitive nonlinear taxation and constitutional choice. *American Economic Journal: Microeconomics*, 4(1):142–75, 2012.
- Thomas Piketty and Emmanuel Saez. Optimal labor income taxation. Technical Report WP No. 18521, NBER, Forthcoming in *Handbook of Public Economics*, vol. 5., December 2012.

- Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229, 2001.
- Emmanuel Saez, Joel Slemrod, and Seth H. Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50, March 2012.
- Laurent Simula and Alain Trannoy. Optimal income tax under the threat of migration by top-income earners. *Journal of Public Economics*, 94:163–173, July 2010.
- Laurent Simula and Alain Trannoy. Shall we keep the highly skilled at home? the optimal income tax perspective. *Social Choice and Welfare*, pages 1–32, 2011.
- Juan C. Suarez Serrato and Owen Zidar. Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms. Technical report, UC Berkeley and Stanford U., 2013.
- Cristobal Young and Charles Varner. Millionaire migration and state taxation of top incomes: Evidence from a natural experiment. *National Tax Journal*, 64:255–284, 2011.