

International tax competition: zero tax rate at the top re-established

Tomer Blumkin · Efraim Sadka ·
Yotam Shem-Tov

Published online: 9 September 2014
© Springer Science+Business Media New York 2014

Abstract In this paper, we extend the zero tax at the top result obtained in the closed economy case with bounded skill distributions for the case of unbounded skill distributions in the presence of international labor mobility and tax competition. We show that in the equilibrium for the tax competition game, the optimal marginal income tax rate converges to zero as the income level tends to infinity. We further show in simulations that the zero-marginal tax result is not a local property: over a substantial range at the higher end of the income distribution, the optimal tax is approximately given by a

Electronic supplementary material The online version of this article (doi:[10.1007/s10797-014-9335-y](https://doi.org/10.1007/s10797-014-9335-y)) contains supplementary material, which is available to authorized users.

Helpful comments from Soren Blomquist, Etienne Lehmann, Laurent Simula, two anonymous referees and seminar participants at CesIfo ESP and PSE conferences and Uppsala University are gratefully acknowledged.

T. Blumkin
Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel
e-mail: tomerblu@bgumail.bgu.ac.il

T. Blumkin · E. Sadka
CesIfo, Munich, Germany

T. Blumkin · E. Sadka
IZA, Bonn, Germany

E. Sadka (✉)
The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel
e-mail: sadka@post.tau.ac.il

Y. Shem-Tov
Department of Economics, UC Berkeley, Berkeley, CA, USA
e-mail: shemtov@berkeley.edu

lump-sum tax set at its *Laffer* rate. We further show that the range in which the optimal marginal tax is approximately set to zero is widening as migration costs decrease.

Keywords Tax competition · Migration · Zero-marginal tax at the top

JEL Classification D6 · H2 · H5

1 Introduction

The general setting examined in [Mirrlees \(1971\)](#) seminal paper provides fairly limited insights regarding the desirable properties of the optimal nonlinear tax schedule, except that its marginal rate is nonnegative. When confining attention to bounded skill distributions, [Sadka \(1976\)](#) and [Seade \(1977\)](#) were nonetheless able to demonstrate that the marginal tax rate levied on the individual with the highest skill level is optimally set to zero (known henceforth as the zero tax at the top result). This led [Sadka \(1976\)](#) to conclude that, assuming that the tax schedule is continuous in income, the marginal tax rate cannot be monotonically increasing with respect to income, that is, the marginal tax rate must decline at high-income levels.

The zero tax at the top property and its implied marginal tax regression within the range of high-income levels is deemed highly controversial. *Prima-facie*, it stands in sharp contrast to observed patterns of progressive income tax schedules in most OECD countries, where, as is often the case, the statutory marginal tax rates increase with respect to income (usually in a piecewise linear fashion). However, the optimal tax analyzed in the literature refers to the overall tax-transfer system and not merely to the statutory tax system. Indeed, with this broader look at the tax system, the marginal progression may dampen considerably in practice. In most countries, where the bulk of welfare (transfer) programs is means-tested (e.g., guaranteed subsistence income level programs), the effective (implicit) marginal tax rates at the bottom of the income distribution are fairly high and may well exceed the statutory marginal income tax rates at the high end of the income distribution.

Early simulations conducted by [Tuomala \(1990\)](#) demonstrate the desirability of setting declining marginal tax rates for a broad range of high incomes. More recently, [Mankiw et al. \(2009\)](#), using [as in [Tuomala \(1990\)](#)] log-normal skill distributions, provide numerical simulations based on the empirical wage distributions from Current Population Survey Data showing modest and decreasing marginal tax rates in the upper part of the skill distribution. Comparing theory and practice, they further argue that, as recommended by theory, based on OECD data, marginal tax rates levied on top earners have decreased over the last three decades. In contrast, in two recent papers, [Diamond \(1998\)](#) and [Saez \(2001\)](#) challenged the zero tax at the top property. [Diamond \(1998\)](#), examining a quasi-linear utility (in consumption) and [Saez \(2001\)](#), extending the analysis to incorporate income effects, have demonstrated that with unbounded distributions,¹ the zero tax at the top fails to hold under reasonable parametric assump-

¹ The justification provided by [Diamond](#) and [Saez](#) for using unbounded skill distributions (although actual distributions are clearly both bounded and with a finite population) derives from an assumption that the

tions regarding the skill distribution, namely top skill levels are distributed according to a Pareto distribution. It is worth noting, though, that whereas the zero tax at the top result fails to hold under the assumption of a Pareto distribution, one can show that with income effects in place marginal tax rates may be declining even for a Pareto distribution of wages [see [Dahan and Strawczynski \(2000\)](#)]. [Diamond and Saez \(2011\)](#) conclude that the zero tax at the top result is of little policy relevance and, in any case, is only a local property that applies to the very top earner, even when bounded distributions are allowed for.

In this paper, we revisit the zero tax at the top result. We deviate from the standard Mirrleesian setting and the bulk of the subsequent literature, which focus on the closed economy case, by allowing for labor migration. Incorporating this extensive margin consideration into the standard framework, we examine the role played by international tax competition in shaping the optimal tax-and-transfer system. We show that even when the skill distribution is assumed to be unbounded, the optimal asymptotic marginal tax rate is zero when migration is taken into account. We further demonstrate by simulations that there exists an income threshold above which the optimal marginal tax rate is approximately set to zero. Notably, the “lump-sum” tax over this range of high-income levels is set at its *Laffer* rate, maximizing tax revenues against the back-drop of the disincentive associated with migration. Thus, the zero tax at the top result is far from being a local property, which applies only to the top-earning individual.

Our paper contributes to a relatively recent strand in the optimal tax literature examining the properties of the optimal *nonlinear* labor income tax system in the presence of mobile labor. Some papers cast the problem in a partial-equilibrium setting, examining the effect of migration on the properties of the optimal nonlinear tax schedule of the state, taking the other states’ tax schedules as exogenous outside options [see [Wilson \(2006\)](#), [Krause \(2009\)](#) and [Simula and Trannoy \(2010\)](#)]. [Hamilton and Pestieau \(2005\)](#) consider a general equilibrium setting where some of the workers are (perfectly) mobile. They focus on the case of small-open economies, where each country ignores the effect of its redistributive policy on international migration; hence, governments are not strategic competitors. Other papers ([Piaser 2007](#); [Bierbrauer et al. 2012](#); [Morelli and Yang 2012](#); [Lehmann et al. 2013](#)) consider a general equilibrium setting and explicitly model the strategic interaction across tax authorities. [Piaser \(2007\)](#) considers a setting with two identical countries and two skill levels, and demonstrates that when migration costs are sufficiently small, the income tax schedule entails no distortions, in contrast to the case of autarky (no migration). [Bierbrauer et al. \(2012\)](#) consider a setting with two governments and a finite number of types. They illustrate a “race to the bottom” argument in the case of tax competition and perfect mobility of the workforce, by showing that there do not exist equilibria in which either the lowest skilled receive transfers or that there exists skill types paying taxes in one country whose utility is higher than the average utility in the other country. [Morelli and Yang](#)

Footnote 1 continued

government does not know the exact realization of the earnings distribution when setting the tax policy. Alternatively, it is assumed that individual skill levels are randomly drawn from some underlying distribution, which is known to the government. The government is hence assumed to maximize expected social welfare subject to a revenue constraint, which holds in expectation.

(2012) consider an extension of [Piaser \(2007\)](#) with three types of workers. They focus on the constitutional choice, within a federation comprised of two states, between a unified (centralized) tax system, where an identical tax system for both states is set by the central authority and an independent (decentralized) tax system, where the tax schedule is independently determined by each state, taking into account that citizens can migrate from one state to the other (a tax competition setting).² [Lehmann et al. \(2013\)](#) examine a setting where both the skill distribution and the distribution of mobility costs are continuous. They show that the properties of the tax schedule generally depend on the correlation between the skill level and the semi-elasticity of migration (the percentage change in the population of a given skill level in response to a unit increase in consumption), and in particular, demonstrate that the optimal marginal tax rates may well be negative at the higher end of the skill distribution. In contrast to [Lehmann et al. \(2013\)](#) that allow for systematic correlations between the skill levels and the costs of migration and examine the general properties of the optimal tax schedule in relation to these correlations, we choose to set focus on what seems to be a natural benchmark for analysis, a setting where migration costs are identically and independently distributed across skill levels, and re-examine the challenged controversial zero tax at the top property. We provide sufficient empirically plausible conditions for the asymptotic marginal tax rate to converge to a finite limit and claim that this limit has to be zero. We further demonstrate that this property is non-local and holds over an entire range of income levels. Finally, we provide an intuitive explanation for the mechanism at work.

In a recent survey on optimal income taxation, [Piketty and Saez \(2012\)](#) extend the asymptotic marginal tax rate formulae provided in [Diamond \(1998\)](#) and [Saez \(2001\)](#) to the case with mobile workforce. They argue that whereas the asymptotic optimal marginal tax rate is substantially reduced in the presence of migration in comparison with the autarky benchmark, the marginal tax rate is still bounded away from zero under empirically reasonable parametric assumptions. The difference between their prediction and our result derives from our assumption that migration costs are independently and identically distributed across skill levels. As will be shown below, this assumption translates into an empirically plausible assumption that migration elasticity is increasing with respect to skill level (rather than being constant, as assumed by [Piketty and Saez](#)). In the limit where skill level grows without bound, this implies an infinite elasticity of migration asymptotically, which implies that the marginal tax rate will converge to zero.

The structure of the paper is as follows: in the next section, we describe the analytical framework. We describe the tax-transfer choice of each national government in Sect. 3. Section 4 characterizes the optimal tax-transfer policy in the international tax competition equilibrium. Some useful numerical simulations are offered in Sect. 5. Section 6 concludes.

² As a robustness check for the three-type setting, [Morelli and Yang \(2012\)](#) also examine an extension of their model to the case of a continuum of types (skill levels). For tractability, they analyze a particular quasi-linear (in leisure) functional form of utility and further assume that skill levels are uniformly distributed over a bounded support.

2 The model

We consider a global economy which is comprised of two identical countries ($i = 1, 2$). World population is normalized to a measure of 2. Each country produces a single consumption good employing labor inputs with different skill levels. We follow [Mirrlees \(1971\)](#) by assuming that the production technology exhibits constant returns to scale and perfect substitutability across skill levels.

We assume that individuals differ in two attributes: (i) innate productive ability (skill level) and (ii) mobility costs (between the two countries). The individual skill level is denoted by θ and is distributed according to the cumulative distribution function $F(\theta)$ with strictly positive densities, $f(\theta)$, over the support $[\underline{\theta}, \infty)$. We follow [Mirrlees \(1971\)](#) by assuming that skill levels are private information unobserved by the government. Notice that we are allowing for general skill distributions and, in particular, allow for an unbounded support.

Turning next to mobility costs, we assume that in the absence of any differences between the two countries (in terms of the fiscal policy implemented by the respective local governments), the world population of each skill group is equally divided between the two countries (population in each country would be given by a unit measure of “permanent residents” in such a case). The mobility cost, measured in consumption terms, incurred by a resident of country i who migrates to the other country, is denoted by m . In order to render our analysis, tractable m is assumed to be distributed uniformly over the support $[0, \delta]$ for each skill level θ [as in [Piaser \(2007\)](#) and [Morelli and Yang \(2012\)](#)].³

Individuals share the same preferences. Following [Atkinson \(1990\)](#) and [Diamond \(1998\)](#), we simplify by assuming that preferences are represented by some quasi-linear utility function of the form:

$$U(c, l, d) = c - h(l) - d \cdot m, \quad (1)$$

where c denotes consumption (gross of migration costs), l denotes labor, d is an indicator function assuming the value of one if the individual migrates and zero otherwise, and $h(\cdot)$ is strictly increasing and strictly convex. To ensure interior solutions throughout INADA, conditions are assumed to hold.

For later purposes, as is common in the optimal tax literature, we reformulate the utility function (gross of migration costs) and represent it as a function of gross income (y), net income (c), and the individual skill level (θ):

$$V(\theta, c, y) \equiv c - h(y/\theta). \quad (2)$$

Hence, utility (net of migration costs) is given by

³ Our main result is robust to the precise specification of the distribution of migration costs, provided that migration costs distribute identically and independently across skill levels (IID). The latter seems to be a natural benchmark to examine. We invoke the assumption of uniform distribution for tractability and in order to remain in line with previous studies. Moreover, as will be shown below (see the discussion in Sect. 5) and consistent with existing empirical evidence suggesting that individuals with a higher skill level are more likely to migrate [see, e.g., [Simula and Trannoy \(2010\)](#)], our IID assumption implies that the migration elasticity is increasing with respect to the skill level.

$$U(\theta, m, c, y, d) \equiv V(\theta, c, y) - d \cdot m. \tag{2'}$$

3 The government problem

We turn next to formulate the government problem. For concreteness, we will focus on country $i = 1$ that takes as given the fiscal policy (tax-and-transfer system) implemented by country $i = 2$. We will then solve for the symmetric *Nash* equilibrium of the fiscal competition game formed between the two countries. We first introduce some useful notation. Denote by $V(\theta)$ the utility level (gross of migration costs) derived by an individual of skill level θ in country 1. Further denote by $c(\theta)$ and $y(\theta)$, correspondingly, the net income and gross income chosen by an individual of skill level θ in country 1.

By virtue of our quasi-linear specification, an individual who incurs mobility cost m will migrate from country $i = 2$ if, and only if, the following condition holds

$$V(\theta) - m \geq V_2(\theta), \tag{3}$$

where $V_2(\theta)$ denotes the utility level derived by the migrating individual in the source country, $i = 2$.

Denote by $m^*(\theta) \equiv V(\theta) - V_2(\theta)$, the cost of migration incurred by an individual of type θ who is just indifferent between staying in country 2, or, migrating to country 1. Thus, any individual of type θ who incurs a cost of migration lower than or equal to the above threshold will migrate to country 1. Recalling our assumption that migration cost is distributed uniformly over the support $[0, \delta]$ in both countries, and that world population is normalized to a measure of 2 (hence, in the absence of differences between the two countries, each is populated by a measure of 1), it follows, by symmetry, that the term $\frac{f(\theta) \cdot m^*(\theta)}{\delta}$ represents the extent of migration of individuals of skill level θ between the two countries. If the term is positive, there is migration from country 2 to country 1, and vice versa.⁴

Clearly, a more generous policy of the government in country $i = 1$ toward individuals of any given skill level will attract a higher migration rate of that skill level, *ceteris paribus*, and vice versa. In a symmetric equilibrium, no migration will take place ($m^* = 0$).

It is worth emphasizing that in our model paying lower taxes serves as the only incentive to migrate, as we choose to set focus on the fiscal competition issue. We do this to simplify our exposition without discounting the importance of other factors affecting migration decisions, which are addressed by the voluminous literature on the economics of migration.

The government in country $i = 1$ is seeking to maximize a *Rawlsian* welfare function, namely the well-being of the least well-off individual

$$W \equiv V(\theta), \tag{4}$$

⁴ We assume that $|m^*(\theta)| < \delta$, namely that only a fraction of the population (for each skill level) is migrating.

by choosing a thrice-continuously differentiable tax schedule, $T(y)$, subject to the revenue constraint (assuming with no loss of generality no revenue needs for the government):

$$\int_{\underline{\theta}}^{\infty} T(t) \cdot \eta(t) \cdot dt = 0, \quad (5)$$

where $\eta(\theta)$ is the measure of individuals of type θ in country 1, which, by virtue of the earlier derivations, is given by

$$\eta(\theta) = f(\theta) \cdot (1 + (V(\theta) - V_2(\theta)) / \delta), \quad (6)$$

and subject to the self-selection constraint [given by the individual envelope condition, see [Salanie \(2003\)](#) for details]:

$$V'(\theta) = \frac{y(\theta) \cdot h'(y(\theta)/\theta)}{\theta^2}. \quad (7)$$

Several remarks are in order. First, notice that our choice of an extremely egalitarian social welfare function (focusing on the least well-off individual) is made in order to demonstrate our key argument in its sharpest relief. We demonstrate that even under the strongest case for redistribution, which lends itself to setting high marginal tax rates on top-earning individuals, the restraining effect of labor migration on the government redistributive ability manifests itself in the form of moderate marginal tax rates levied on workers with the highest skill levels (converging to zero as the skill level diverges). Note further that unlike in the standard optimal tax problem, the population in our setting is endogenously determined rather than being fixed. The standard case of no migration is obtained for the special limiting case where $\delta \rightarrow \infty$, in which case $\eta(\theta) \equiv f(\theta)$. The fact that the population is endogenously determined raises conceptual difficulties when defining the proper social objective, namely determining who are the agents whose welfare is to count [see the discussion in [Simula and Trannoy \(2012\)](#)]. We assume that the government aims to maximize the utility derived by the least well-off individuals among the “permanent residents” population, namely the initial population residing in the country under a *laissez-faire* regime [we thus invoke the “citizen criterion” suggested by [Simula and Trannoy \(2012\)](#)]. Finally, notice that each government is taking the tax policy of the other country as given, namely country 1 takes as given $V_2(\theta)$, the utility derived by an individual of type θ in country 2, when choosing its tax policy. We will look for a symmetric *Nash* equilibrium for the fiscal competition game between the two countries. Note that symmetry implies that in equilibrium the same tax schedule will be implemented by both countries, that is, $\eta(\theta) = f(\theta)$, which immediately follows from condition (6) when symmetry is imposed.

For later purposes, it would be useful to re-formulate the revenue constraint in (5), employing the condition in (2) and the identity $T \equiv y - c$, to obtain the following expression:

$$\int_{\underline{\theta}}^{\infty} (y(t) - V(t) - h(y(t)/t)) \cdot \eta(t) \cdot dt = 0. \quad (8)$$

4 Characterization of the optimal policy in equilibrium

We let $\hat{c}(V_2(\theta))$, $\hat{y}(V_2(\theta))$, $\hat{c}_2(V(\theta))$, and $\hat{y}_2(V(\theta))$ denote the optimal solution for the government problem in country 1 (and in country 2, respectively) as a function of the utility derived by individuals in country 2 (correspondingly, in country 1). A symmetric equilibrium for the game between the two countries is given by the implicit solution to the following equation:

$$V(\theta) \equiv V[\theta, \hat{c}(V_2(\theta)), \hat{y}(V_2(\theta))] = V_2(\theta). \tag{9}$$

We solve the constrained optimization program as an optimal control problem employing *Pontryagin’s* maximum principle. We choose $y(\theta)$ as the control variable and $V(\theta)$ as the state variable. Formulating the *Hamiltonian* and employing the first-order conditions, the *transversality* conditions, and the symmetry property yield following some re-arrangements (see Appendix 1 in Electronic supplementary material for details), the following expression that characterizes the optimal marginal tax rate:

$$\frac{T'(y)}{1 - T'(y)} = \left(\int_{\theta}^{\infty} \left(1 - \frac{(y - V - h(y/t))}{\delta} \right) \cdot f(t) dt \right) \cdot \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1}{f(\theta) \cdot \theta}, \tag{10}$$

where y is the gross level of income associated with the skill level θ , and ε_y denotes the labor supply elasticity.⁵

It is straightforward to verify that in the limiting case of no migration (an autarky), where $\delta \rightarrow \infty$, the condition in (10) reduces to the standard condition found in the literature [see [Diamond \(1998\)](#) and [Salanie \(2003\)](#), *inter alia*]:

$$\frac{T'(y)}{1 - T'(y)} = \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1 - F(\theta)}{f(\theta) \cdot \theta} \tag{11}$$

The additional component, given by the first term of the expression on the right-hand side of condition (10), captures the restraining effect of migration on the extent of redistribution, reflected by reduced marginal tax rates relative to the autarky setting with no migration in place.

The expression in (10) is quite complex, as the income distribution is endogenously determined and is affected by the (optimally designed) nonlinear labor income tax schedule. Whereas the differential equation given in (10) does not admit for a general closed-form solution, one can still provide some partial characterization of its prop-

⁵ We follow the standard approach in the literature, where the equilibrium is characterized by the necessary first-order conditions of the government optimization problem and the sufficient second-order conditions are assumed to hold throughout.

erties, summarized in the following proposition (the proof is relegated to Appendix 2 in Electronic supplementary material):⁶

Proposition (i) $T'(y) \geq 0$ for all y ; (ii) $\lim_{y \rightarrow \infty} T'(y) = 0$ and (iii) $\lim_{y \rightarrow \infty} T(y) = \delta$.

The proposition indicates that the asymptotic marginal tax rate is zero. In simulations, we demonstrate that the zero-marginal tax result is far from being a local property by showing that there exists a whole range of income levels at the higher end of the income distribution for which the optimum tax is approximately given by a lump sum. Our results stand in sharp contrast to previous results in the optimal tax literature: the zero tax at the top property (Sadka 1976; Seade 1977) only applies to the top earner and is restricted to bounded income distributions. The subsequent literature (Diamond 1998; Saez 2001), focusing on unbounded distributions, suggests that a zero-marginal tax cannot be part of the optimal schedule for empirically reasonable income distributions. The economic rationale behind the result we derive is straightforward. In the absence of migration, the incentive constraint faced by the local government is related to the intensive margin, that is, the tax schedule is designed in a way that ensures no mimicking by the high skilled (attempting to mimic their low-skilled counterparts in order to reduce their tax liability). With migration in place, an extensive margin comes into play, as the local government attempts to attract high-skill migrants from the neighboring country (or, alternatively, to mitigate the incentive of high-skill residents to migrate to the other country). For the very high-skill population (the top range of the income distribution), the extensive margin becomes highly manifest and hence the dominant one. In this range of incomes, the government is approximately levying a lump-sum tax given by δ . The lump-sum tax is set at its *Laffer* rate, namely set at the rate that maximizes total tax revenues (taking into account the disincentive effect on migration). Notice that over the range in which the tax is set to be a lump sum (local), no mimicking incentives are trivially satisfied. Thus, the only focus is indeed on migration (extensive margin) incentives.

To see why the *Laffer* rate is indeed given by δ , denote by $V(\theta)$ the utility (gross of migration costs) derived by a typical individual of skill level θ in country 1 in the absence of taxes. Further denote by T the lump-sum tax set by country 1. By virtue of quasi-linearity (and the nature of a lump-sum tax that induces no substitution effect), the net-of-tax utility (gross of migration costs) derived by a θ -type individual in country 1 is given by $V(\theta) - T$. Denoting by $V_2(\theta)$ the utility (gross of migration costs) derived by a typical individual of skill level θ in country 2, by virtue of the earlier derivations [condition (6)], the measure of individuals of type θ in country 1 is given by

$$\eta(\theta) = f(\theta) \cdot (1 + (V(\theta) - T - V_2(\theta)) / \delta). \quad (12)$$

⁶ In the proof of the proposition, we invoke an additional technical assumption that $\lim_{\theta \rightarrow \infty} [1 + 1/\varepsilon_y] \cdot \frac{1 - F(\theta)}{f(\theta) \cdot \theta} < \infty$. This assumption is supported by standard assumptions in the literature (e.g., assuming an

Footnote 6 continued
isoelastic disutility from labor and that the skill distribution is approximated by a *Pareto* distribution above a certain income level, as in Diamond (1998), among others).

The total tax revenues associated with a lump-sum tax, T , raised from individuals of skill level θ , denoted by $TR(\theta, T)$, are then given by

$$TR(\theta, T) = T \cdot f(\theta) \cdot (1 + (V(\theta) - T - V_2(\theta)) / \delta). \tag{13}$$

Differentiating the expression in (13) with respect to T and equating to zero yield

$$\begin{aligned} \partial TR(\theta, T) / \partial T &= f(\theta) \cdot (1 + [(V(\theta) - T - V_2(\theta)) / \delta]) - f(\theta) \cdot T / \delta = 0 \\ \Leftrightarrow T &= [\delta + (V(\theta) - V_2(\theta))] / 2. \end{aligned} \tag{14}$$

By virtue of (14), it follows that $\partial^2 TR / \partial T^2 = -\frac{2f(\theta)}{\delta} < 0$; hence, the second-order condition is satisfied. Furthermore, it follows from (14) that $\frac{\partial T}{\partial V_2(\theta)} = -\frac{1}{2} > -1$. In words, a unit increase in the tax level set by country 2 [implying a unit decrease in $V_2(\theta)$] induces an optimal increase of half of a unit in the tax level set by country 1 in response. Invoking the symmetry property, and applying with respect to the best response function of country 2, implies that a unit increase in the tax level set by country 1 will induce an increase of half of a unit in the tax level set by country 2 in response. This ensures that the equilibrium is stable (in the standard myopic sense). Finally, imposing the (symmetric) equilibrium condition $[V(\theta) - T = V_2(\theta)]$ yields upon re-arrangement $T = \delta$.

5 Numerical simulations

As we were unable to obtain a closed-form solution for the differential equation in condition (10), we resort to numerical solution of the optimal tax schedule. This will enable us to examine the effect of migration on the tax schedule and specifically on the marginal tax rate that the higher skilled individuals are faced with. The key purpose of the simulations was to demonstrate that the zero-marginal tax asymptotic result we obtain analytically is far from being a local property, but rather holds over a substantial range of incomes at the higher end of the income distribution.

We make several parametric assumptions. We first assume that the skill level is distributed according to a *Pareto* distribution, namely $f(\theta) = \frac{\alpha\theta^\alpha}{\theta^{\alpha+1}}$, $\theta > \underline{\theta}$, which implies that $\frac{1-F(\theta)}{f(\theta)\cdot\theta} = 1/\alpha$. We further assume, as is common in the literature [see [Diamond \(1998\)](#) and [Salanie \(2003\)](#), among others], that the disutility from labor takes an isoelastic functional form, namely $h(l) = \frac{l^{1+\varepsilon_y}}{1+\varepsilon_y}$; hence, the pre-tax income elasticity, ε_y , is constant.

The optimal tax schedule is the solution to differential equation given in (10) with two boundary conditions:

$$\lim_{\theta \rightarrow \infty} T'[y(\theta)] = 0 \quad \text{and} \quad T'[y(\underline{\theta})] = B / (1 + B), \tag{15}$$

where $y(\theta)$ is given by the implicit solution to $\theta [1 - T'(y)] = h'(\frac{y}{\theta})$ and $B = \left[\frac{1+\varepsilon_y}{\varepsilon_y} \right] \cdot \frac{1}{\alpha}$.

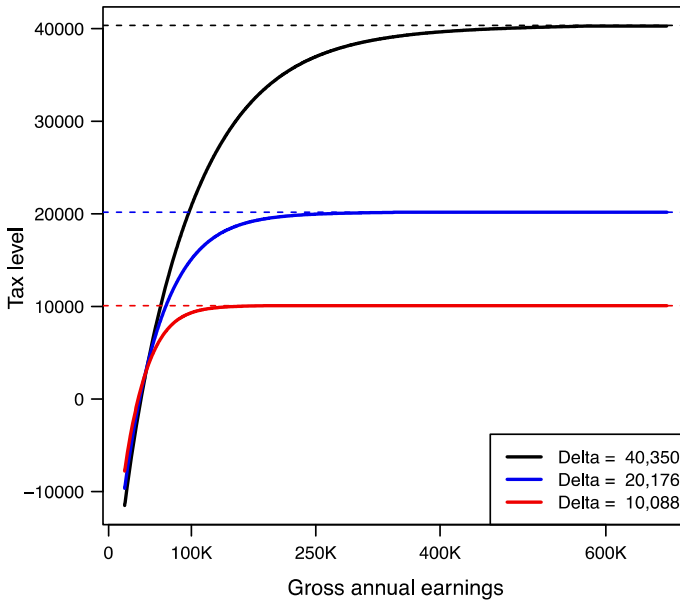


Fig. 1 Effect of migration on the optimal tax schedule

Notice that the first boundary condition follows from the fact that by virtue of the proposition, the asymptotic marginal tax is zero. Further notice that the second boundary condition is obtained by substituting into (10), employing our assumption that the total tax revenues raised by the government are zero.

We turn next to specify our calibration assumptions. We follow [Piketty and Saez \(2013\)](#) by assuming that the pre-tax income elasticity is $\varepsilon_y = 0.2$. We follow [Saez \(2001\)](#) by assuming that the shape parameter of the underlying skill distribution is given by $\alpha = 2$. Finally, we calibrate $\underline{\theta}$, the scale parameter of the *Pareto* skill distribution, based on the observed annual earnings data (total population, full-time wage and salary 16 years, and older workers) from the US Bureau of Labor Statistics (2013), assuming [as in [Saez \(2002\)](#)] a constant 40% tax rate in place.⁷ We use the 10th percentile observed income to calibrate the lower-bound skill level. We turn next to investigate the effect of migration on the optimal schedule.

Figures 1 and 2 below depict the optimal tax and marginal tax schedules, respectively, for different costs of migration, measured by the parameter δ . We present the optimal schedules for three different levels of δ , given by the observed median income (40,352\$), 50% of the observed median income (20,176\$) and 25% of the observed median income (10,088\$).

Notice crucially that whereas migration is a one-time cost, there is a lifetime benefit associated with a reduced tax liability. One may argue therefore that our parameter

⁷ Our qualitative results are robust to changing the calibrating assumption about the marginal tax rate in place. Allowing the government to raise positive funds (rather than assuming, as we do, that taxation is purely redistributive) does not change our qualitative results either.

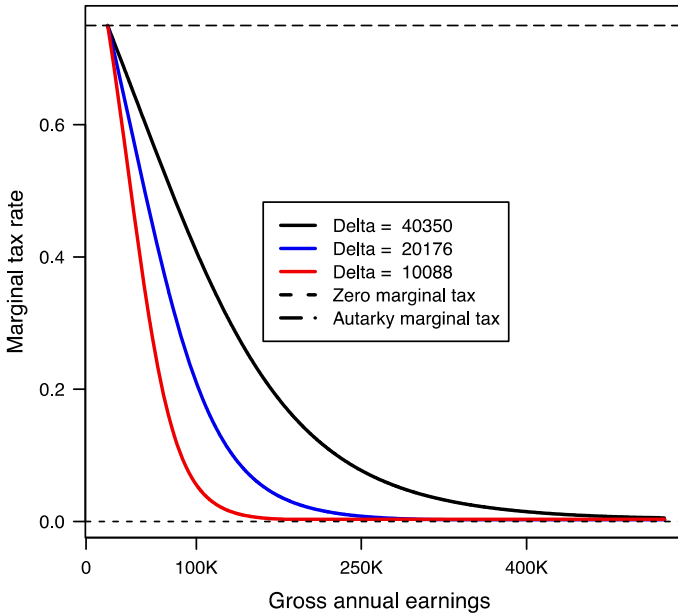


Fig. 2 Effect of migration on the optimal marginal tax rates

choices for the cost of migration are too small. However, our calibrations make use of annual income distribution data, so that migration costs would be better interpreted as yearly flows rather than as a one-time expenditure.

As can be readily observed from both figures, there exists an income threshold above which the tax schedule becomes approximately flat, namely individuals are faced with a zero-marginal tax rate, where the lump-sum tax in this income range is given by δ . The asymptotic tax levels and marginal tax rates are represented by the dashed curves in both figures. In addition, as a reference for comparison, we also represent in Fig. 2 the marginal tax schedule under autarky, which is given by a flat curve due to our assumption that skill levels distribute according to a *Pareto* distribution [this property immediately follows from condition (11), recalling our assumption of constant pre-tax income elasticity]. Moreover, the interval of incomes over which individuals are faced with a zero-marginal tax rate is expanding as the costs of migration decrease [in the limiting cost-less migration case, which is not presented in the figures, the optimal schedule converges to the *laissez-faire* zero tax schedule reflecting an extreme race to the bottom, in which, trivially, all individuals are faced with a zero-marginal tax rate]. Finally, the figure indicates that the marginal tax rates are declining over the entire range of incomes [in fact, we prove this property in Appendix 3 (Electronic supplementary material) for the case where skill levels are distributed according to a *Pareto* distribution and pre-tax income elasticity is constant].

The fact that the range in which individuals are subject to a zero-marginal tax rate is expanding as migration costs decrease is an extension of Piaser (2007), who demonstrates that the patterns of the binding self-selection constraints crucially hinge on the level of migration costs. In a two-type model, Piaser (2007) demonstrates that

when the costs of migration are sufficiently small, both self-selection constraints will be non-binding in the optimal solution. In contrast, when migration costs are large enough, including the limiting case of autarky, the self-selection constraint of the high-skill individual will bind (the standard case in the literature). This implies that for sufficiently small migration costs, there will be zero tax at the bottom as well as at the top (a property holding for large migration costs as well). The (lump-sum) tax levied on the high-skill residents in this case is set (optimally) at the *Laffer* level, namely the tax is set so as to maximize the total revenues raised from the high-skill population. Piaser thus demonstrates that the range in which individuals are faced with a zero-marginal tax rate is expanding as migration costs decrease. To see the intuition for this result, recall that an egalitarian government seeking to redistribute wealth from the high-skill toward the low-skill residents is essentially faced with two challenges. The first one is the standard one on the intensive margin (which applies in the case of an autarky, as well) and derives from the mimicking threat of high-skill individuals. The second one on the extensive margin (which applies only when tax competition takes place) derives from the migration threat of high-skill residents. With large enough migration costs, the impact of the extensive margin consideration (the potential threat of a massive migration of the high skill) is relatively small; hence, the standard result (as in the case of autarky) applies. When migration costs are small enough, the migration incentive comes into play. Although the government can increase the tax burden shifted on the high-skill residents without inducing the latter to mimic, the reduction in the tax base due to the ensued migration is large enough to offset the gain from increasing the tax rate. We find in the continuum case similar patterns. We demonstrate that the zero-marginal tax at the top property holds over an interval of skill levels, rather than being a local property holding for the top-earning individual only (as is the standard argument in the literature with no migration and a bounded skill distribution) and further show that this interval expands, as migration costs decrease.

One may be tempted to attribute the zero tax property at the top obtained in our setting to a “Ramsey” type of argument (an inverse-elasticity feature). The idea is that levying a small marginal tax on the top-earning folks (whose skill level is unbounded) would result in an *unbounded* benefit from migration to the other jurisdiction (compared with the bounded cost of migration). This implies infinite migration elasticity and justifies the zero-marginal tax at the top. It is important to bear in mind, however, that our zero tax property is not confined to the top-earning individual. To illustrate the argument, it would be instructive to return to the case of bounded skill distribution in which it is well known that for the top-earning individual, the marginal tax is zero (and clearly the migration elasticity is not infinite). The reason for this property is not driven by *Ramsey* reasoning but rather by the existence of an extensive margin associated with migration option. The easiest way to capture this is to look at the finite type version. In the absence of migration considerations, local incentive compatibility constraints are binding (type j is indifferent between choosing his bundle and mimicking type $j - 1$ for any j). When the marginal tax rate on type $j - 1$ is zero, one can slightly shift his bundle downwards along the indifference curve (in the gross income–net income plane) by imposing a strictly positive marginal tax rate, thereby creating a slack in the incentive compatibility (IC) constraint of j (without violating the revenue constraint of the government). This allows the government to extract some more funds from the

types $j, j + 1, j + 2$, and above and use these extra resources to enhance redistribution, without violating any of the IC constraints of types j and above. This is the standard argument that warrants imposing a positive marginal tax rate on any individual barring the top-earning one. Clearly, the justification for introducing marginal distortions derives from the nature of binding IC constraints. Now, consider the case of migration. Suppose that the incentive constraint of j is not binding. It would be tempting to conclude that by increasing the tax on j (and higher types) in a lump-sum fashion, one can obtain extra funds and enhance redistribution. However, when migration is an option, the increase in the tax burden on j (and above) would induce some of the j -type (and higher skill types) population to emigrate. With small enough migration costs, a large fraction of the j (and above) population will leave. This will imply that the increase in the tax burden will result in a deficit rather than a fiscal surplus. Thus, the IC constraint for j and above may well be non-binding. There is no justification to set a positive marginal tax on these guys, hence (not just on the top-earning guy whom nobody attempts to mimic). The optimal tax rate will be set so as to maximize the surplus from these types against the backdrop of migration and will imply setting the tax at the *Laffer* rate, given by δ in our setting.

Reinterpreting our results using elasticity terms, notice that the migration elasticity for any given skill level is defined as follows: (% change in population)/(% change in net income). Formally, denoting by $e(\theta)$ the migration elasticity of individuals of skill level θ , the elasticity formula is given by

$$e(\theta) \stackrel{\text{def}}{=} \frac{\partial \eta(\theta)}{\partial c(\theta)} \cdot \frac{c(\theta)}{\eta(\theta)}, \quad (16)$$

where $\eta(\theta)$ and $c(\theta)$ denote, respectively, the measure of individuals of type θ and their corresponding level of net income (consumption). Substituting for $\eta(\theta)$ from (6) into (16), employing the symmetric equilibrium property [$f(\theta) = \eta(\theta)$] and re-arranging yields

$$e(\theta) = c(\theta)/\delta. \quad (17)$$

Notice that the elasticity given in (17) is decreasing with respect to migration costs (δ) and increasing with respect to net income [$c(\theta)$]. Thus, consistent with empirical data [see, e.g., the discussion in [Simula and Trannoy \(2010\)](#)], high-skill workers are indeed more likely to migrate. In particular, when migration costs are small enough, the migration elasticity of individuals with skill levels at the higher end of the skill distribution is sufficiently large, so that their corresponding incentive compatibility constraints are not binding in the optimal solution and the optimal tax is set at the *Laffer* rate, implying a zero-marginal tax rate. Moreover, with unbounded skill distribution, the elasticity diverges to infinity (for any finite level of migration costs). Thus, the zero-marginal tax property holds asymptotically for any finite level of migration costs.

An alternative way to illustrate the relationship between the costs of migration and the desirability of levying a zero-marginal tax rate is given in Fig. 3 below. The figure depicts for different costs of migration the fraction of the population (at the high end of the income distribution) faced with a marginal tax rate lower than or equal

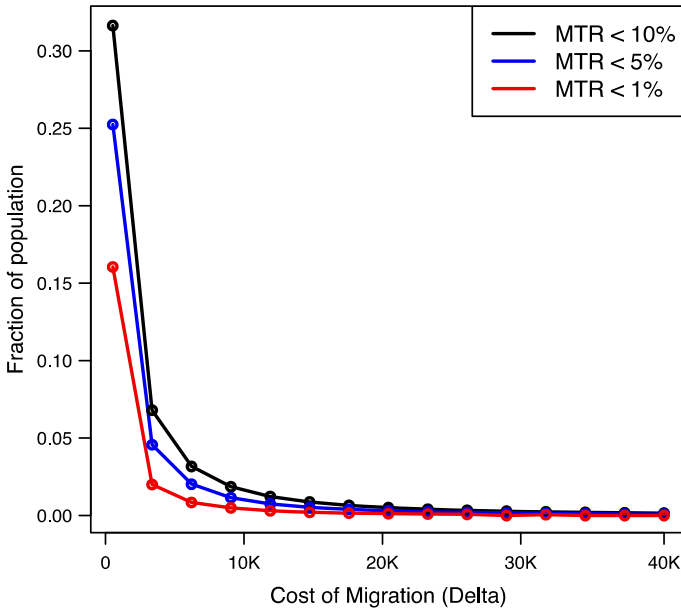


Fig. 3 Relationship between zero-marginal tax and migration

to 1%, 5%, and 10%, respectively. As can be observed, reduced migration costs imply that a higher fraction of the population is faced with (an almost) zero-marginal tax rate (less than 1%) and a fairly moderate marginal tax rate (less than 5 or 10%, respectively).⁸

In order to illustrate the substantial impact migration and tax competition bear on the optimal marginal tax rates levied on top earners, suppose that the parameter δ is equal to the observed median annual income, given by 40,352\$. In this case, the top 0.87% of the population (individuals who earn an income level exceeding 441,394\$) is faced with a marginal tax rate lower than 1% and the top 1.75% of the population (individuals who earn an income level exceeding 287,361\$) is faced with a marginal tax rate lower than 5%.

Moving further down the income distribution, it turns out that the top 2.54% of the population (individuals who earn an income level exceeding 227,897\$) is faced with a marginal tax rate lower than 10 percent, the top 4% of the population (individuals who earn an income level exceeding 168,345\$) is faced with a marginal tax rate lower than 20%, whereas the top 4.88% of the population (individuals who earn an income level exceeding 148,323\$) is faced with a marginal tax rate lower than 25%! Notably, the elasticity of migration evaluated at the threshold income level above which the marginal tax rate is lower than 20% (25%), given, respectively, by 168,345\$ and 148,323\$, is correspondingly given by 3.387 and 2.948. For comparison, [Kleven](#)

⁸ As indicated by the simulations presented in Fig. 1 and as proved analytically for *Pareto* distributions in Appendix 3 (Electronic supplementary material), the marginal tax rate is decreasing with respect to income. Clearly, the population subject to the lowest marginal tax rates is at the higher end of the income distribution.

et al. (2013), using Danish data, estimated the elasticity of migration (in response to tax incentives) at the high end of the income distribution to be in the range of 1.5. It is important to emphasize that the tax incentives examined in Kleven et al. (2013) were temporary (a reduction from a regular 59% tax rate to 25% rendered over a period of 3 years to immigrants with annual income levels exceeding 103,000 Euros as of 2009) so in this sense, the estimates clearly provide us with a lower bound for the (lifetime) migration elasticity. Thus, our parametric choices for the cost of migration appear to be plausible and in line with available empirical evidence.⁹ In contrast, under autarky (no migration, namely letting δ diverge to infinity), using the same parameters and invoking a *Rawlsian* objective, the optimal marginal tax rate levied on the individuals with the highest skill level would be set to 75%! Thus, even with an extremely egalitarian government seeking to maximize the well-being of the least well-off individuals, which lends itself to setting very high marginal tax rates on the top earners, the optimal progressive income tax schedule suggests levying fairly small marginal tax rates on individuals with the highest skill level due to the restraining effects of labor migration and tax competition.

6 Conclusion

One of the most controversial results in the optimal tax literature was the zero tax at the top property due to Sadka (1976) and Seade (1977), arguing that when the underlying skill distribution is bounded, the optimal marginal tax rate levied on the top-earning individual should be set to zero. A corollary of this result is the desirability of setting declining marginal tax rates at the top end of the income distribution. Diamond (1998), examining the quasi-linear utility specification, and Saez (2001), extending the analysis by incorporating income effects, have challenged the zero tax at the top property, showing that for empirically relevant unbounded skill distributions, the asymptotic optimal marginal tax rate should be bounded away from zero.

In this paper, we revisit the result, demonstrating that when migration consideration is taken into account, in a simple tax competition framework, there exists an income threshold above which the optimal marginal tax rate should be approximately set to zero. The optimal lump-sum tax levied on income levels at the top end should be set at its *Laffer* rate, maximizing tax revenues against the backdrop of migration threats. Moreover, we show that the range in which the optimal marginal tax is set to zero is widening, as migration costs decrease.

The literature on tax competition in the presence of labor migration usually demonstrates the restraining effect of migration on the redistributive power of an egalitarian government. Our focus in this paper is on the impact migration bears on the optimal marginal tax rates levied on top earners. Using plausible parametric assumptions, we show that the impact could be quite substantial: our simulations suggest that a non-negligible fraction of the population at the higher end of the income distribution should

⁹ It is important to note that there are only a few empirical studies that provide estimates of the elasticity of migration. The case of Denmark may not be necessarily representative of other countries, but still gives us some reference point against which to assess the plausibility of our parametric choices.

be subject to fairly small marginal tax rates. When compared with recent results in the literature in the case of autarky, which call for levying fairly high marginal tax rates on top earners, it seems that migration of high-skill workers should be seriously taken into account when designing the optimal tax-and-transfer system.

References

- Atkinson, A. (1990). Public economics and the economic public. *European Economic Review*, *34*, 225–248.
- Bierbrauer, F., Brett, C., & Weymark, J. (2012). Strategic nonlinear income tax competition with perfect labor mobility. Mimeo, Department of Economics, Vanderbilt University.
- Dahan, M., & Strawczynski, M. (2000). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates: Comment. *American Economic Review*, *90*, 681–686.
- Diamond, P. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review*, *88*, 83–95.
- Diamond, P., & Saez, E. (2011). The case for a progressive tax: From basic research to policy recommendations. *Journal of Economic Perspectives*, *25*, 165–190.
- Hamilton, J., & Pestieau, P. (2005). Optimal income taxation and the ability distribution: Implications for migration equilibria. *International Tax and Public Finance*, *12*, 29–45.
- Kleven, H., Landais, C., Saez, E., & Schultz, E. (2013) Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in Denmark. *Quarterly Journal of Economics* (forthcoming).
- Krause, A. (2009). Education and taxation policies in the presence of countervailing incentives. *Economica*, *76*, 387–399.
- Lehmann, E., Simula, L., & Trannoy, A. (2013) Tax me if you can: Optimal non-linear income tax between competing governments. Cesifo working paper #4351.
- Mankiw, G., Weinzierl, M., & Yagan, D. (2009). Optimal taxation in theory and practice. *Journal of Economic Perspectives*, *23*, 147–174.
- Morelli, M., & Yang, H. (2012). Competitive nonlinear taxation and constitutional choice. *American Economic Journal: Microeconomics*, *4*, 142–175.
- Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, *38*, 175–208.
- Piaser, G. (2007). Labor mobility and income tax competition. In G. Gregoriou & C. Read (Eds.), *International taxation handbook: Policy, practice, standards and regulation* (pp. 73–94). Oxford: CIMA Publishing.
- Piketty, T., & Saez, E. (2012). Optimal labor income taxation. *Handbook of Public Economics* (forthcoming).
- Piketty, T., & Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, *81*, 1851–1886.
- Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, *68*, 205–229.
- Saez, E. (2002). Optimal income transfer programs: Intensive versus extensive labor supply responses. *Quarterly Journal of Economics*, *117*, 1039–1073.
- Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies*, *43*, 261–268.
- Salanie, B. (2003). *The economics of taxation*. Cambridge: MIT Press.
- Seade, J. (1977). On the shape of optimal tax schedules. *Journal of Public Economics*, *7*, 203–236.
- Simula, L., & Trannoy, A. (2010). Optimal income tax under the threat of migration by top-earners. *Journal of Public Economics*, *94*, 163–173.
- Simula, L., & Trannoy, A. (2012). Shall we keep the highly skilled at home? The optimal income tax perspective. *Social Choice and Welfare*, *39*, 751–782.
- Tuomala, M. (1990). *Optimal income tax and redistribution*. New York: Oxford University Press.
- Wilson, J. (2006). Income taxation and skilled migration: The analytical issues. Mimeo, Department of Economics, Michigan State University.