

Blumkin, Tomer; Sadka, Efraim; Shem-Tov, Yotam

Working Paper

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

CESifo Working Paper, No. 3820

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Blumkin, Tomer; Sadka, Efraim; Shem-Tov, Yotam (2012) : International tax competition: Zero tax rate at the top re-established, CESifo Working Paper, No. 3820, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/58334>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

International Tax Competition:
Zero Tax Rate at the Top Re-established

Tomer Blumkin
Efraim Sadka
Yotam Shem-Tov

CESIFO WORKING PAPER NO. 3820

CATEGORY 1: PUBLIC FINANCE

MAY 2012

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

International Tax Competition: Zero Tax Rate at the Top Re-established

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 large range at the higher end of the income distribution, the optimal tax is approximately given by a 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.

JEL-Code: D600, H200, H500.

Keywords: tax competition, migration, zero marginal tax at the top.

Tomer Blumkin
Department of Economics
Ben-Gurion University
Beer-Sheba 84105 / Israel
tomerblu@bgumail.bgu.ac.il

Efraim Sadka
The Eitan Berglas School of Economics
Tel Aviv University
Tel-Aviv 69978
Israel
sadka@post.tau.ac.il

Yotam Shem-Tov
The Eitan Berglas School of Economics
Tel Aviv University
Tel-Aviv 69978
Israel
yotam121@yahoo.com

1. Introduction

The general setting examined in Mirrlees (1971) seminal paper provides fairly limited insights regarding the desirable properties of the optimal non-linear 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, assuming that the tax schedule is continuous in income, that the marginal tax rate cannot be monotonically rising with respect to income; that is the marginal tax rate must decline at high incomes [see also the elaborate discussion in Mankiw et al. (2009)]. Later simulations [see Tuomala (1990)] demonstrate the desirability of setting declining marginal tax rates for a broad range of high incomes.

Prima-facie, the zero tax at the top property and its implied marginal-tax regression within the range of high income levels, 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 piece-wise 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 are 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.

Nonetheless, the zero tax at the top result still remained highly controversial. In two influential papers, Diamond (1998) and Saez (2001) challenged this result. Diamond (1998), examining a quasi-linear utility and Saez (2001), extending the analysis to incorporate income effects, have demonstrated that with unbounded distributions, the zero tax at the top fails to hold under reasonable parametric assumptions regarding the skill distribution. Diamond and Saez (2011) conclude that the result is of little policy relevance and in any case, is only a local property which 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 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 backdrop of the disincentive associated with migration.

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 Section 3. Section 4 characterizes the optimal tax-transfer policy in the international tax competition equilibrium. Some useful numerical simulations are offered in Section 5. Section 6 concludes.

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), (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 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 $[0, \delta / 2]$ for each skill level θ .

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

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

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):

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

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

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

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:

$$(3) \quad V(\theta) - m \geq V_2(\theta),$$

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 $[0, \delta/2]$ 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 $\frac{2 \cdot f(\theta) \cdot m^*(\theta)}{\delta}$ 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.

Clearly, a more generous policy of the government in country $i=1$ towards 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 ($m^* = 0$).

The government in country $i=1$ is seeking to maximize the following welfare function:

$$(4) \quad \int_{\underline{\theta}}^{\infty} \varphi(V(t)) dF(t),$$

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

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

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

$$(6) \quad \eta(\theta) = f(\theta) \cdot (1 + 2 \cdot (V(\theta) - V_2(\theta)) / \delta),$$

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

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

Several remarks are in order. First, notice that the welfare function we invoke is assumed to depend on the utilities of the 'permanent residents' (individuals residing in the country in a *laissez-faire* allocation absent of government intervention in both countries) and not on the utilities of migrants. Clearly, in a symmetric equilibrium the latter would be of measure zero. 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

$\eta(\theta) \equiv f(\theta)$. 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:

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

4. Characterization of the Optimal Policy in Equilibrium

We next turn to solve the 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* yields:

$$(9) \quad H = \varphi(V) \cdot f + \lambda \cdot \left[(y - V - h(y/\theta)) \cdot \left((1 + 2 \cdot (V - V_2) / \delta) \right) \right] \cdot f + \mu \cdot \frac{y \cdot h'(y/\theta)}{\theta^2}$$

Formulating the necessary conditions for optimality yields:

$$(10) \quad \frac{\partial H}{\partial y} = \lambda \cdot \left(1 - \frac{h'}{\theta} \right) \cdot \eta + \mu \frac{h' + y \cdot h'' / \theta}{\theta^2} = 0,$$

$$(11) \quad -\frac{\partial H}{\partial V} = -\varphi'(V) \cdot f + \lambda \cdot f \cdot \left[\left(1 + 2 \cdot (V - V_2) / \delta\right) - \frac{2 \cdot (y - V - h(y / \theta))}{\delta} \right] = \mu'(\theta).$$

In addition the two *transversality* conditions hold:

$$(12) \quad \mu(\underline{\theta}) = \lim_{\theta \rightarrow \infty} \mu(\theta) = 0.$$

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:

$$(13) \quad V(\theta) \equiv V[\theta, \hat{c}(V_2(\theta)), \hat{y}(V_2(\theta))] = V_2(\theta).$$

In order to simplify our exposition, we will henceforth focus on the *Rawlsian* government case; namely, the case in which the government is seeking to maximize the well-being of the least well-off individual. Employing the first-order conditions in (10) and (11), the *transversality* conditions in (12) and the symmetry property, yields, following some rearrangements (see appendix A for details), the following expression that characterizes the optimal marginal tax rate:

$$(14) \quad \frac{T'(y)}{1 - T'(y)} = \left(\int_{\theta}^{\infty} \left(1 - \frac{2 \cdot (y - V - h(y / t))}{\delta} \right) \cdot f(t) dt \right) \cdot \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1}{f(\theta) \cdot \theta},$$

where ε_y denotes the labor supply elasticity.

It is straightforward to verify that in the limiting case of no migration (an autarky), where $\delta \rightarrow \infty$, the condition in (14) reduces to the standard condition found in the literature [see Diamond (1998) and Salanie (2003)]:

$$(14') \quad \frac{T'(y)}{1-T'(y)} = \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1-F(\theta)}{f(\theta) \cdot \theta}$$

The additional component, given by the first term of the expression on the right-hand side of condition (14), 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 (14) is quite complex, as the income-distribution is endogenously determined and is affected by the (optimally designed) non-linear labor income tax schedule. Following Saez (2001), instead of using the (actual) density of income associated with the tax schedule, $T(y)$, we simplify the expression in (14) by introducing $g(y)$, the (virtual) density of income (at the pre-tax income level, y) assuming a local linear approximation of the tax schedule; namely, assuming that the tax schedule, $T(y)$, is replaced by a (locally) linear schedule tangent to the schedule, $T(y)$, at the pre-tax income level, y . Denoting by G the cumulative distribution function associated with g [that is, $\partial G(y) / \partial y \equiv g(y)$], it follows by construction that:

$$(15) \quad G[y(\theta)] = F(\theta).$$

Differentiating the expression in (15) with respect to θ , employing the individual first-order condition and following some algebraic manipulations (see appendix B for details), one can obtain the following condition that relates the densities f and g [which replicates condition (13) in Saez (2001)]:

$$(16) \quad g(y) \cdot y \cdot (1 + \varepsilon_y) = f(\theta) \cdot \theta.$$

Substituting from the conditions in (15) and (16) into (14) and re-arranging, yields:

$$(17) \quad \frac{T'(y)}{1-T'(y)} = \left(1 - \frac{\int_y^\infty \frac{2 \cdot T(t)}{\delta} \cdot g(t) dt}{1-G(y)} \right) \cdot \frac{1}{\varepsilon_y} \cdot \frac{1-G(y)}{g(y) \cdot y}.$$

Whereas the differential equation given in (17) is fairly complex and does not admit for a general closed-form solution, one can still provide some partial characterization of its properties, summarized in the following proposition (the proof is relegated to appendix C):¹

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

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 exist 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) and Seade (1977)] only applies to the top-earner and is restricted to bounded income distributions. The subsequent literature [Diamond (1998) and 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

¹ In the proof of the proposition we invoke an additional technical assumption that $\lim_{y \rightarrow \infty} [1-G(y)]/g(y) \cdot y / \varepsilon_y < \infty$. This assumption is supported by standard assumptions in the literature [e.g., assuming an iso-elastic disutility from labor and that the income distribution is approximated by a *Pareto* distribution above a certain income level, as in Diamond (1998), amongst others].

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 $\delta/2$ (see the proof of the proposition in appendix C). The lump-sum tax is set at its *Laffer* rate; namely, set at the rate which 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 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 $\delta/2$, 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:

$$(18) \quad \eta(\theta) = f(\theta) \cdot (1 + 2 \cdot (V(\theta) - T - V_2(\theta)) / \delta).$$

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

$$(19) \quad TR(\theta, T) = T \cdot f(\theta) \cdot (1 + 2 \cdot (V(\theta) - T - V_2(\theta)) / \delta).$$

Differentiating the expression in (19) with respect to T and equating to zero, imposing the (symmetric) equilibrium condition $[V(\theta) - T = V_2(\theta)]$ yields:

$$(20) \quad \partial TR(\theta, T) / \partial T = f(\theta) \cdot [1 - 2T / \delta] = 0 \Leftrightarrow T = \delta / 2.$$

5. Numerical Simulations

As we were unable to obtain a closed-form solution for the differential equation in equation (17), 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 which the higher skilled individuals are faced with. 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 $\frac{1 - F(\theta)}{f(\theta) \cdot \theta} = 1 / \alpha$.

Employing (15) and (16), it then follows that:

$$(21) \quad \frac{1 - G(y)}{g(y) \cdot y} = \frac{1 - F(\theta)}{f(\theta) \cdot \theta / (1 + \varepsilon_y)} = \frac{1 + \varepsilon_y}{\alpha}.$$

We further assume, as is common in the literature [see Diamond (1998) and Salanie (2003), amongst others], that the disutility from labor takes an iso-elastic functional form,

namely, $h(l) = \frac{l^{1+1/\varepsilon_y}}{1+1/\varepsilon_y}$; hence, the pre-tax income elasticity, ε_y , is constant. Letting

$$B = \left[\frac{1 + \varepsilon_y}{\varepsilon_y} \right] \cdot \frac{1}{\alpha} \text{ and substituting into (17) yields:}$$

$$(22) \quad \frac{T'(y)}{1-T'(y)} = B - B \cdot \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)},$$

where B is a constant. Further simplifying the differential equation in (22) to eliminate the integral expression, we fully differentiate the expression in (22) with respect to y , yielding:

$$(23) \quad \frac{T''(y)}{(1-T'(y))^2} = -B \cdot \frac{2}{\delta} \cdot \frac{g(y)}{1-G(y)} \cdot \left[-T(y) + \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} \right].$$

Employing (21) and (22) and re-arranging yields,

$$(24) \quad T''(y) = \frac{\alpha \cdot (1-T'(y))^2}{y \cdot (1+\varepsilon_y)} \cdot \left[\frac{2 \cdot B}{\delta} \cdot T(y) - B + \frac{T'(y)}{1-T'(y)} \right].$$

We let the function $J(y)$ denote the total tax revenues from individuals whose income level is lower than or equal to y , formally given by:

$$(25) \quad J(y) \equiv \int_{\underline{y}}^y T(t) \cdot g(t) dt.$$

Hence,

$$(26) \quad J'(y) = T(y) \cdot g(y).$$

The revenue constraint faced by the government is characterized therefore by the differential equation given in condition (26) and the associated boundary condition,

$$\lim_{y \rightarrow \infty} J(y) = 0.$$

The optimal tax schedule is the solution to the system of differential equations given by (24) and (26) and three boundary conditions:

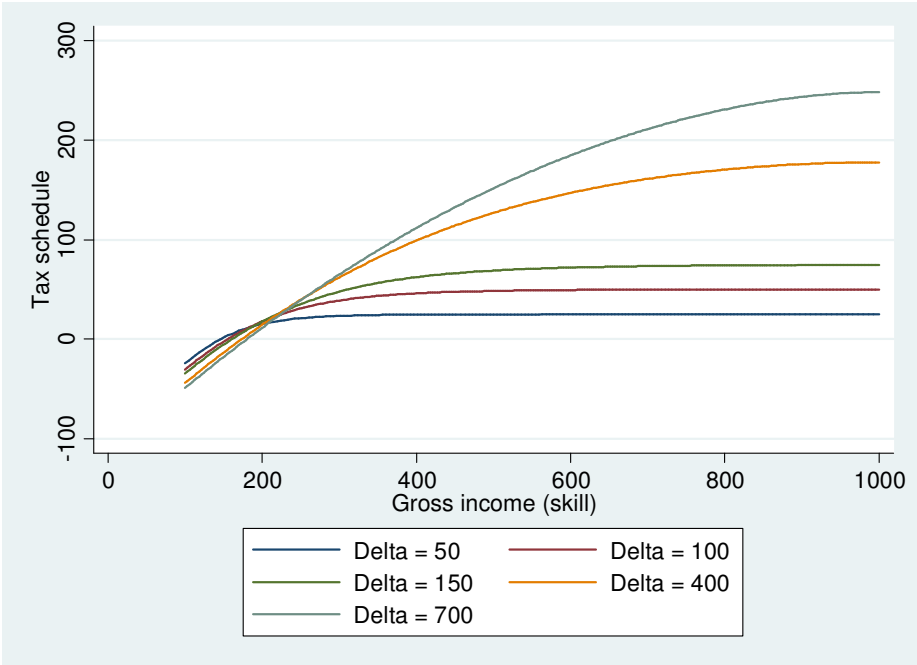
$$(27) \quad \lim_{y \rightarrow \infty} J(y) = 0, \lim_{y \rightarrow \infty} T'(y) = 0 \quad \text{and} \quad T'(\underline{y}) = \frac{B}{1+B}.$$

Notice that the second boundary condition follows from the fact that by virtue of the proposition the asymptotic marginal tax is zero. Further notice that the latter boundary condition is obtained by substitution into (22), employing our assumption that the total tax revenues raised by the government is zero.

We follow Gruber and Saez (2002) by assuming that the pre-tax income elasticity is $\varepsilon_y = 0.4$, and follow Saez (2001) by assuming that $\alpha = 2$. Finally, we assume that the support for of the income distribution is given by $y \in [100, 1000]$. We turn next to investigate the effect of migration on the optimal schedule.

Figure 1 below depicts the optimal tax schedule for different costs of migration. As can be readily observed from the figure, 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 $\delta/2$. 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, 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, this latter property can be proved under our parametric assumptions (see appendix D).

Figure 1: The effect of migration on the optimal tax schedule



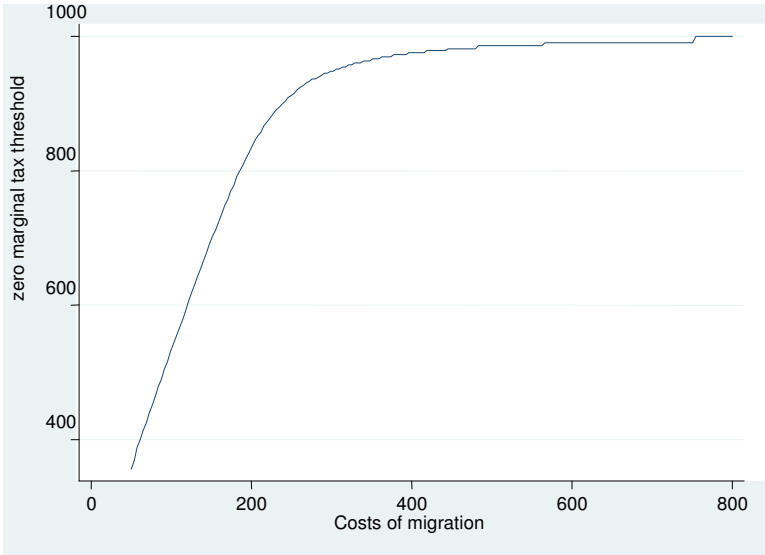
This result we obtain is an extension of Piaser (2007), who demonstrates that the patterns of the binding self-selection constraints crucially hinge on the level of migration costs. In a two type model Piaser (2007) demonstrates that when the costs of migration are sufficiently small, both self-selection 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 towards 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. Most importantly, we show that a case for levying a zero marginal tax at the top is established, even when the skill distribution is unbounded. Moreover the 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).

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 figure 2 below. The figure

depicts for different costs of migration the income threshold above which individuals are faced with a marginal tax rate lower than or equal to 1%. As can be observed, reduced migration costs imply that a higher fraction of the population is faced with an (almost) zero marginal tax rate.

Figure 2: The relationship between zero marginal tax and migration



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 are 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.

Appendix A: Derivation of Condition (14)

In the symmetric equilibrium, by construction, the tax schedules implemented by both countries are identical and, therefore, no migration takes place. By virtue of the symmetry property the condition in (11) simplifies to:

$$(A1) \quad -\frac{\partial H}{\partial V} = -\varphi'(V) \cdot f + \lambda \cdot f \cdot \left[1 - \frac{2 \cdot (y - V - h(y / \theta))}{\delta} \right] = \mu'(\theta)$$

Integrating the expression in (A1) and employing the *transversality* condition for the limiting skill level given in (12), yields:

$$(A2) \quad \mu(\theta) = \int_{\theta}^{\infty} \varphi'(V) \cdot f(t) dt - \lambda \cdot \int_{\theta}^{\infty} \left(1 - \frac{2 \cdot (y - V - h(y / t))}{\delta} \right) \cdot f(t) dt .$$

Employing the revenue constraint in (8) and the *transversality* condition for the lower-bound skill level, given in (12), yields:

$$(A3) \quad \lambda = \int_{\underline{\theta}}^{\infty} \varphi'(V(t)) \cdot f(t) dt .$$

Next we define the following function, which measures the average social marginal utility of income over the interval $[\theta, \infty)$:

$$(A4) \quad D(\theta) = \frac{1}{1 - F(\theta)} \cdot \int_{\theta}^{\infty} \varphi'(V(t)) \cdot f(t) dt .$$

Notice that $\lambda = D(\underline{\theta})$. Employing (A4), one can re-formulate the expression in (A2) to obtain:

$$(A5) \quad \mu(\theta) = D(\theta) \cdot [1 - F(\theta)] - D(\underline{\theta}) \cdot \int_{\theta}^{\infty} \left(1 - \frac{2 \cdot (y - V - h(y/t))}{\delta} \right) \cdot f(t) dt$$

The first order condition for the individual optimization program implies,

$$(A6) \quad 1 - T'(y) = h'(\theta) .$$

Denoting the net hourly wage-rate earned by an individual of skill-level θ by $\theta_n \equiv \theta(1 - T'(y))$, the first order condition in (A7) can be re-written as:

$$(A7') \quad h'(y/\theta) = \theta_n .$$

Differentiating the first-order condition in (A7') with respect to the net hourly wage-rate, θ_n , it is straightforward to derive the elasticity of the pre-tax income, which is then given by:

$$(A8) \quad \varepsilon_y = \frac{\frac{\partial y}{y}}{\frac{\partial \theta_n}{\theta_n}} = \frac{\theta_n \cdot \theta}{y \cdot h''} .$$

Employing (A7), (A7') and (A8) yields:

$$(A9) \quad h' + y \cdot h'' / \theta = \theta(1 - T'(y)) + \theta(1 - T'(y)) \cdot \frac{1}{\varepsilon_y} = \theta(1 - T'(y)) \left[1 + \frac{1}{\varepsilon_y} \right]$$

Substituting from (A5), (A6), (A7) and (A9) into (10) yields after re-arrangement:

$$(A10) \quad \frac{T'(y)}{1 - T'(y)} = \left(\int_{\theta}^{\infty} \left(1 - \frac{2 \cdot (y - V - h(y/t))}{\delta} \right) \cdot f(t) dt - \frac{D(\theta) \cdot (1 - F(\theta))}{D(\theta)} \right) \cdot \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1}{f(\theta) \cdot \theta},$$

We will hence focus our attention on the *Rawlsian* case, in which social welfare is given by $V(\underline{\theta})$, the utility level derived by the least well-off individual. Invoking a *Rawlsian* objective implies that $D(\theta) = 0$ for all $\theta > \underline{\theta}$. Substituting into the expression in (A11) then yields:

$$(A12) \quad \frac{T'(y)}{1 - T'(y)} = \left(\int_{\theta}^{\infty} \left(1 - \frac{2 \cdot (y - V - h(y/t))}{\delta} \right) \cdot f(t) dt \right) \cdot \left[1 + \frac{1}{\varepsilon_y} \right] \cdot \frac{1}{f(\theta) \cdot \theta}$$

Appendix B: Derivation of Expression (16)

By fully differentiating the individual first-order condition in (A6) with respect to θ , assuming a local linear approximation of the tax schedule, one obtains:

$$(B1) \quad 1 - T'(y) = h'' \cdot \left(\frac{y'(\theta) \cdot \theta - y}{\theta^2} \right),$$

which upon re-arrangement holds if-and-only-if:

$$(B2) \quad \frac{\theta^2 \cdot (1 - T'(y))}{h'' \cdot y} = \frac{y'(\theta) \cdot \theta}{y} - 1.$$

Substituting from (A8) into (B2) implies that:

$$(B3) \quad \varepsilon_y = \frac{y'(\theta) \cdot \theta}{y} - 1.$$

Differentiation of the expression in (15) with respect to θ yields:

$$(B4) \quad g[y(\theta)] \cdot y'(\theta) = f(\theta).$$

Substituting for the term $y'(\theta)$ from (B3) into (B4), and re-arranging, yields the expression given in (16).

Appendix C: Proof of the Proposition

We prove the proposition in several steps organized into a series of lemmas.

Lemma 1: When $T'(y) = 0$ then $T(y) = \delta / 2$.

Proof: Assume by negation that $T'(\hat{y}) = 0$ and $T(\hat{y}) < \delta / 2$. Re-arranging the expression in (17) yields,

$$(17') \quad \frac{T'(y)}{1-T'(y)} = B(y) - B(y) \cdot \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)},$$

where $B(y) = \frac{1}{\varepsilon_y} \cdot \frac{1-G(y)}{g(y) \cdot y} > 0$. Fully differentiating the expression in (17') with respect

to y yields,

$$(C1) \quad \frac{T''(y)}{(1-T'(y))^2} = -B(y) \cdot \frac{2}{\delta} \cdot \frac{g(y)}{1-G(y)} \cdot \left[-T(y) + \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} \right] - B'(y) \cdot \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} + B'(y)$$

Substituting from (17') into (C1) yields,

(C2)

$$\frac{T''(y)}{(1-T'(y))^2} = -B(y) \cdot \frac{g(y)}{1-G(y)} \cdot \left[-\frac{2}{\delta} \cdot T(y) - \frac{T'(y)}{1-T'(y)} \cdot \frac{1}{B(y)} + 1 \right] + \frac{B'(y)}{B(y)} \cdot \left(\frac{T'(y)}{1-T'(y)} - B(y) \right) + B'(y)$$

Substituting $T'(y) = 0$ into (C2) and re-arranging yields,

$$(C3) \quad T''(y) = -B(y) \cdot \frac{g(y)}{1-G(y)} \cdot \left[-\frac{2}{\delta} \cdot T(y) + 1 \right]$$

It follows from (C3) and by our presumption that $T(\hat{y}) < \delta/2$, that $T''(\hat{y}) < 0$; hence, by continuity (invoking a first-order approximation), $T'(\hat{y} + \varepsilon) < 0$ for sufficiently small $\varepsilon > 0$. That is, the marginal tax rate is negative within a small neighborhood to the right of \hat{y} . As the marginal tax rate is zero at \hat{y} , it follows from the condition in (17') that:

$$(C4) \quad \int_{\hat{y}}^{\infty} T(t) \cdot g(t) dt = \frac{\delta}{2} (1 - G(\hat{y})).$$

It follows by virtue of (C4) and our presumption that $T(\hat{y}) < \delta/2$ that there exist some income level $y' > \hat{y}$ for which $T(y') > \delta/2$. Hence, there exists some income level $\hat{y} < y'' \leq y'$ for which $T'(y'') > 0$. Then, by the intermediate value theorem there exists an income level for which the marginal tax rate is zero within the interval (\hat{y}, y'') . Let A denote the (non-empty and bounded) set of all income levels within the interval (\hat{y}, y'') for which the marginal tax rate is zero, and further denote by \tilde{y} the greatest lower-bound of the set A . By construction, it follows that $\tilde{y} > \hat{y}$. By virtue of (17') and the definition of \tilde{y} , it follows that:

$$(C5) \quad \int_{\tilde{y}}^{\infty} T(t) \cdot g(t) dt = \frac{\delta}{2} (1 - G(\tilde{y})).$$

It further follows that $T'(y) < 0$ for all $\hat{y} < y < \tilde{y}$. Hence, it follows that $T(y) < \delta/2$ for all $\hat{y} < y < \tilde{y}$, which, by virtue of (C5), implies that:

$$(C6) \int_{\hat{y}}^{\infty} T(t) \cdot g(t) dt < \frac{\delta}{2}(1 - G(\hat{y})).$$

Thus we obtain a contradiction to (C4).

In exactly the same manner (the formal steps are therefore omitted) one can prove by negation that it cannot be the case that $T'(y) = 0$ and $T(y) > \delta/2$. This concludes the proof.

Lemma 2: If $T'(y^*) = 0$ and $y^* < \infty$ then $\forall y, y^* \leq y, T'(y) = 0$.

Proof: Suppose by negation that for some $y, \hat{y} > y^*, T'(\hat{y}) \neq 0$. For concreteness, we assume that $T'(\hat{y}) < 0$ (the other case can be proved by symmetric arguments and is hence omitted). We first turn to show that $T(\hat{y}) > \delta/2$. Suppose by negation that $T(\hat{y}) \leq \delta/2$. As $T'(\hat{y}) < 0$, it follows by virtue of (17') that:

$$(C7) \int_{\hat{y}}^{\infty} T(t) \cdot g(t) dt > \frac{\delta}{2}(1 - G(\hat{y})).$$

Thus, by our presumption that $T(\hat{y}) \leq \delta/2$, it necessarily follow that $T(y) > \delta/2$ for some $y > \hat{y}$. Hence, there exists some $y, y > \hat{y}$, for which $T'(y) > 0$. By virtue of the intermediate value theorem, it follows that there exists an income level $y, y > \hat{y}$, for which the marginal tax rate is zero. Let A denote the (non-empty and bounded from below) set of all income levels within the interval (\hat{y}, ∞) for which the marginal tax rate is zero, and further denote by \tilde{y} the greatest lower-bound of the set A . By construction, it follows that $\tilde{y} > \hat{y}$. By virtue of (17') and the definition of \tilde{y} , it further follows that:

$$(C8) \int_{\tilde{y}}^{\infty} T(t) \cdot g(t) dt = \frac{\delta}{2}(1 - G(\tilde{y})).$$

It further follows that $T'(y) < 0$ for all $\hat{y} < y < \tilde{y}$. Hence, it follows that $T(y) < \delta/2$ for all $\hat{y} < y < \tilde{y}$, which, by virtue of (C8), implies that:

$$(C9) \int_{\hat{y}}^{\infty} T(t) \cdot g(t) dt < \frac{\delta}{2}(1 - G(\hat{y})).$$

Thus we obtain a contradiction to (C7). Thus we have established that:

$$(C10) T(\hat{y}) > \delta/2.$$

By virtue of lemma 1 and as $T'(y^*) = 0$, it follows that $T(y^*) = \delta/2$. It therefore follows that there exists some level of income y' , $y^* < y' < \hat{y}$, for which $T'(y') > 0$. Hence by our presumption that $T'(\hat{y}) < 0$, it follows that there exists some level of income y , $y^* < y' < y < \hat{y}$, for which $T'(y) = 0$.

Let A denote the (non-empty and bounded) set of all income levels within the interval (y', \hat{y}) for which the marginal tax rate is zero, and further denote by \tilde{y} the least upper-bound of the set A . By construction, it follows that $\tilde{y} < \hat{y}$. By virtue of the definition of \tilde{y} , it follows that:

$$(C11) T'(\tilde{y}) = 0.$$

It further follows by the definition of \tilde{y} that $T'(y) < 0$ for all $\tilde{y} < y \leq \hat{y}$. By virtue of lemma 1, $T(\tilde{y}) = \delta/2$, hence $T(y) < \delta/2$ for all $\tilde{y} < y \leq \hat{y}$, which implies that:

$$(C12) T(\hat{y}) < \delta/2. \text{ Thus we obtain a contradiction to (C10).}$$

This establishes the claim.

Lemma 3: $T'(y) \geq 0$ for all y .

Proof: By virtue of (17') the marginal tax rate faced by the individual with the lowest income level is given by:

$$(C13) \frac{T'(y)}{1-T'(y)} = \frac{1}{\varepsilon_y} \cdot \frac{1-G(y)}{g(y) \cdot y} > 0.$$

Suppose by negation that there exists an income level for which the marginal tax rate is negative. By the intermediate value theorem there exists an income level for which the marginal tax rate is zero. Let A denote the (non-empty and bounded from below) set of all income levels for which the marginal tax rate is zero, and further denote by \tilde{y} the greatest lower-bound of the set A . By construction, it follows that $\tilde{y} > \underline{y}$. By virtue of (17') and the definition of \tilde{y} , it follows that:

$$(C14) \int_{\tilde{y}}^{\infty} T(t) \cdot g(t) dt = \frac{\delta}{2} (1-G(\tilde{y})).$$

By virtue of lemma 2, it follows that $T'(y) = 0$ for all $y \geq \tilde{y}$. By construction, $T'(y) > 0$ for all $y < \tilde{y}$. Thus we obtain the desired contradiction.

Lemma 4: $\lim_{y \rightarrow \infty} T'(y) = 0$.

We first establish that $T(y) \leq \delta/2$ for all y . To see this, suppose by negation that there exists some income level, y' , for which $T(y') > \delta/2$. Then, as the marginal tax is non-negative for all y (by lemma 3) it follows that $T(y) > \delta/2$ for all $y \geq y'$. It follows that,

$$(C15) \int_{y'}^{\infty} T(t) \cdot g(t) dt > \frac{\delta}{2} (1-G(y')).$$

By virtue of (17') it then follows that $T'(y') < 0$, which contradicts lemma 3. We conclude that $T(y)$ is bounded from above by $\delta/2$. As $T(y)$ is non-decreasing, it follows that $T(y)$ converges to some finite limit. Let $\lim_{y \rightarrow \infty} T(y) = T < \infty$. We turn next to examine the marginal tax rate as $y \rightarrow \infty$. Taking the limit of the expression in (17') implies:

$$(C16) \quad \lim_{y \rightarrow \infty} \frac{T'(y)}{1 - T'(y)} = \lim_{y \rightarrow \infty} B(y) \times \lim_{y \rightarrow \infty} \left[1 - \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1 - G(y)} \right]$$

By our earlier assumption (see the discussion in footnote 1 in the main text),

$\lim_{y \rightarrow \infty} B(y) = B < \infty$. Applying *L'Hôpital's Rule* then implies:

$$(C17) \quad \lim_{y \rightarrow \infty} \frac{T'(y)}{1 - T'(y)} = B \cdot \left[1 - \frac{2}{\delta} \cdot T \right] < \infty.$$

As both $T(y)$ and $T'(y)$ converge to a finite limit when y goes to infinity, it follows that $\lim_{y \rightarrow \infty} T'(y) = 0$. This concludes the proof.

Appendix D: The marginal tax rate is declining with respect to income under a *Pareto* skill distribution and an iso-elastic disutility from labor

Re-arranging the expression in (17) yields,

$$(17') \quad \frac{T'(y)}{1 - T'(y)} = B(y) - B(y) \cdot \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1 - G(y)},$$

where $B(y) = \frac{1}{\varepsilon_y} \cdot \frac{1 - G(y)}{g(y) \cdot y} > 0$. Fully differentiating the expression in (17') with respect

to y yields,

$$(D1) \frac{T''(y)}{(1-T'(y))^2} = -B(y) \cdot \frac{2}{\delta} \cdot \frac{g(y)}{1-G(y)} \cdot \left[-T(y) + \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} \right] - B'(y) \cdot \frac{2}{\delta} \cdot \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} + B'(y)$$

With a *Pareto* skill distribution and an iso-elastic disutility from labor, $B'(y)=0$, hence, substitution into (D1) yields:

$$(D2) \frac{T''(y)}{(1-T'(y))^2} = -B(y) \cdot \frac{2}{\delta} \cdot \frac{g(y)}{1-G(y)} \cdot \left[-T(y) + \frac{\int_y^\infty T(t) \cdot g(t) dt}{1-G(y)} \right]$$

Now consider some level of income y' , for which $T'(y') > 0$. By virtue of lemma 3, $T'(y) \geq 0$ for all y , hence, as $T'(y') > 0$, it follows that:

$$(D3) \quad \frac{\int_{y'}^{\infty} T(t) \cdot g(t) dt}{(1 - G(y'))} > T(y').$$

Substituting into (D2) implies that $T''(y) < 0$. This concludes the proof.

References

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

Diamond, P. and E. Saez (2011) "The Case for a Progressive Tax: From Basic Research to Policy Recommendations", *Journal of Economic Perspectives*, 25, 165–90.

Gruber, J. and E. Saez (2000) "The Elasticity of Taxable Income: Evidence and Implications", *Journal of Public Economics*, 84, 1-32.

Mankiw, G. et al. (2009) "Optimal Taxation in Theory and Practice", *Journal of Economic Perspectives*, 23, 147-174.

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: Gregoriou, G., Read, C. (Eds.), *International Taxation Handbook: Policy, Practice, Standards and Regulation*, CIMA Publishing, Oxford, 73-94.

Saez (2001) "Using Elasticities to Derive Optimal Income Tax Rates", *Review of Economic Studies*, 68, 205-229.

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", MIT Press.

Seade, J. (1977) "On the Shape of Optimal Tax Schedules", *Journal of Public Economics*, 7, 203-236.

Tuomala, M. (1990) "Optimal Income Tax and Redistribution", New York, Oxford University Press.