Optimal income taxation with tax competition

Vilen Lipatov*   Alfons Weichenrieder†

April 2011

Abstract

We introduce tax competition for mobile labor into an optimal-taxation model with two skill levels and analyze a symmetric subgame-perfect Nash equilibrium of the game between two governments and two taxpayer populations. Tax competition reduces the distortion from the informational asymmetry and increases employment of the less productive individuals. When countries are heterogeneous, this effect is more pronounced in the smaller country.

JEL Classification Codes: F22, H21

Keywords: optimal income tax, migration, unemployment, tax competition, Leviathan government

*Goethe University Frankfurt, corresponding author: e-mail: Lipatov@em.uni-frankfurt.de, phone: (+49) 69 798-34777, address: Grüneburgplatz 1, 60323 Frankfurt am Main; Germany
†Goethe University Frankfurt, Vienna University of Economics and Business, and CESifo
1 Introduction

Recent years have seen a surge of research on tax competition. This is of little surprise, as in our globalized world the borders are becoming increasingly open; people, goods, and resources increasingly mobile; and government policies more interdependent. Nowadays, there is little doubt that a tax policy neglecting cross-border effects is no more than a (possibly convenient) abstraction.

A wide range of problems have been addressed within this blooming field, from tax-base erosion to redistribution and allocation of resources to coordination and harmonization proposals. Sinn (2003) provides an excellent overview of tax competition literature within a broader framework of systems competition. Capital tax competition has perhaps the longest tradition, as capital has early been recognized to be a mobile factor of production and, correspondingly, a most mobile tax base (for a seminal contribution, see Zodrow and Mieszkowski 1986). Income tax competition has also been analyzed, but mostly insofar as the mobile factors could affect it. Lately, mobility of individuals also has come into focus, especially in the context of European integration (e.g., Richter 2004).

Our paper contributes to this new strand of literature by merging tax competition for mobile labor with optimal-income-taxation approaches. In a novel article, Simula and Trannoy (2010) analyze how migration possibilities affect the optimal taxation formula in a single country. Although our paper is also based on connecting optimal taxation with labor mobility, unlike Simula and Trannoy we focus on the effect of tax competition on the employment of low-skilled workers.

We augment a standard two-skill-level optimal-income-taxation model with the possibility of migration for high-skilled workers. In this framework governments compete for these workers and their taxes in a simple Hotelling setting.

1Huber (1999) studies the effect of capital tax competition on the optimal income tax when labor is immobile. Osmundsen et al. (2000) analyze optimal income tax with mobile labor, but the asymmetric information in their model is about location preferences rather than productivity. Osmundsen et al. (1998) study a similar problem for firms.

2Other recent contributions to the analysis of optimal income tax with tax competition include Morelli et al. (2010) who focus on the political economy implications of tax competition, and Bierbrauer et al. (2011) who confirm a “race to the bottom” under the assumption of perfect labor mobility.
The main result of our analysis is that opening the borders increases employment of the low-skill workers. Intuitively, competitive pressure lowers the tax on the mobile high-skill workers. This allows the government to reduce the distortion from taxing the low-skilled without violating the incentive compatibility constraint. As a result, their employment increases. This is a clear, testable prediction that is robust to the choice of various objectives of the government and the relative size of the countries.

We also show that the smaller country lowers its tax on the high-skilled by more than the larger country does. This is consistent with the general intuition that the smaller entity is more aggressive in competition, as it has less revenue to lose from its own population, but a larger competitor’s tax base to gain from lowering the tax.

There is a clear contribution of our result to the policy discussion about the vices and virtues of tax competition: despite a negative effect on tax revenues, it also has a positive effect on the employment of low-skilled workers. This may be particularly important for the countries with low efficiency of the government sector, as tax competition tames Leviathan governments and improves the resource allocation.

The rest of the paper is structured as follows. Section 2 contains the basic Leviathan model; in section 3 alternative government objectives are discussed; in section 4 the model with asymmetric equilibrium is analyzed; limitations and extensions are discussed in the conclusion.

2 The Model

2.1 Closed economy

We use as a benchmark Stiglitz’s (1982) version of the Mirrlees (1971) model of income taxation, but introduce a different objective of the government. In a closed economy, individuals of measure 1 have identical preferences that can be represented by a utility function \( u(x, y) \), where \( x \geq 0 \) is consumption and \( 0 \leq y \leq 1 \) is the time worked. \( u \) is a strictly concave, continuously differentiable function, strictly increasing in \( x \) and strictly decreasing in \( y \).

There are two types of individuals in the economy: those with high productivity \( \theta_H \) constitute measure \( \gamma \), and those with low productivity \( \theta_L \) have correspondingly measure \( 1 - \gamma; \theta_H > \theta_L > 0 \). An individual of type \( i \) provides \( z_i = \theta_i y_i \) of labor while investing \( y_i \) of her time.
The government cannot observe \( \theta \), but it does observe income \( z \) and chooses income taxes \( \{t_L, t_H\} |_{t_i \leq z_i} \) to maximize the tax revenue
\[
R = \gamma t_H + (1 - \gamma) t_L
\]
subject to a satisfaction constraint \( u_L, u_H \geq u_0 \). This constraint makes it impossible for the living conditions of the poor to be set arbitrarily low and may be interpreted as a requirement of a modern welfare state. In a separating equilibrium, the individual \( i \) then chooses \((x_i, y_i)\) that maximizes \( u(x, y) \) subject to \( x_i \leq \theta_i y_i - t_i \), and corresponding incentive compatibility (IC) and participation constraints. For simplicity we assume that the utility thresholds that ensure participation are equal to \( u_0 \).

It is well known that the budget constraints, the participation constraint for the low type and the IC constraint either for the high type or for the low type, are binding in such problems (e.g., Stiglitz 1982). In the appendix we show that in our setting it is possible to rule out a binding IC constraint for the low productivity type.

The individual optimization will result in setting consumption and time for the low type at the levels satisfying
\[
x_i = \theta_i y_i - t_i,
\]
\[
\theta_i (1 - t'_i) u_x + u_y = 0.
\]
The Leviathan will then leave the less productive with their reservation utility, setting \( t_L \) to satisfy
\[
u(z_L - t_L, z_L/\theta_L) = u_0,
\]
and \( t_H \) to satisfy
\[
u(z_L - t_L, z_L/\theta_H) = u(z_H - t_H, y_H)
\]
and the revenue maximization condition. The government will not find itself better off in a pooling equilibrium in our setting, as shown by Stiglitz (1982). Nothing guarantees, however, that the corner with \( z_L = 0 \) is not hit.

Writing down the maximization explicitly (and in line with the literature), we can define the marginal tax rate as
\[
t'_i = 1 + \frac{u_y}{\theta_i u_x}.
\]
We set up the Lagrangian $L = \gamma t_H + (1 - \gamma) t_L + \mu (u(z_L - t_L; z_L/\theta_L) - u_0) + \lambda (u(z_H - t_H; z_H/\theta_H) - u(z_L - t_L; z_L/\theta_L))$ and denote for compactness $u^L := u(z_L - t_L; z_L/\theta_L); u^H := (z_H - t_H; z_H/\theta_H); u^{HL} := u(z_L - t_L; z_L/\theta_H)$. The corresponding FOCs are
\begin{align*}
t_L &: \quad 1 - \gamma - \mu u_x^L + \lambda u_x^{HL} = 0, \\
z_L &: \quad \mu (u_x^L + u_y^L/\theta_L) - \lambda (u_x^{HL} + u_y^{HL}/\theta_H) = 0, \\
t_H &: \quad \gamma - \lambda u_x^H = 0, \\
z_H &: \quad \lambda (u_x^H + u_y^H/\theta_H) = 0.
\end{align*}

The last equation immediately produces a “no distortion at the top” result: $u_x^H + u_y^H/\theta_H = 0 \implies t_H = 0$. From quasiconcavity, $dx/dy = -u_y/u_x$ is an increasing function of $y$. Thus, as long as $x_L < x_H$, we have $-u_y^{HL}/u_x^{HL} < -u_y^H/u_x^H$. Correspondingly, $u_x^{HL} + u_y^{HL}/\theta_H > u_x^H + u_y^H/\theta_H = 0$, and from (1b) $u_x^L + u_y^L/\theta_L > 0$, so that $t_L > 0$.

Denote the optimal tax rates in the autarky case by $\{t_L^a, t_H^a\}$.

The appendix shows that for a sufficiently high level of $\gamma$ the low-skilled will find it optimal not to participate in the labor force ($z_L^a = 0$). In what follows we assume that $\gamma$ is not too high for interior solution (Our basic result about the increased employment level of the “poor” from the tax competition for the high skilled will sustain in the corner solution).

### 2.2 Open economy

Suppose now we have two identical economies of the sort described above. Additionally, high-productivity individuals may migrate between countries in search of a better life. Low-productivity individuals are immobile. This is an extreme case of correlation between productivity and mobility decision, and we employ it for the sake of simplicity. Simula and Trannoy (2010) discuss why it seems reasonable to assume that higher-skilled workers are also more mobile. For example, skilled workers have better language skills and should have easier access to information on foreign countries.

Our high-productivity individuals differ in their propensity to migrate. Specifically, we assume that the initial population in each country is distributed on the interval $[0, 1]$ according to a continuously differentiable distribution function $F(a)$. Under this assumption we can use a Hotelling model for the analysis. Basically, our migration costs are similar in spirit to switching costs widely analyzed in the industrial organization literature (e.g., Farrell
and Klemperer 2007). The utility of the high-productivity individual located at \( a \) is \( u(x, y) - c(a) \), where \( c \) is a strictly increasing function with \( c(0) = 0 \). Thus, we assume that utility is additively separable with respect to migration costs.

One caveat related to this analysis is that upon migration the government can observe the type of individual and thus impose a perfect-information tax on her (or any other tax conditioned upon the fact of migration and hence potentially different than the tax on the rest of population). However, we can exclude such behavior by postulating that the government must treat migrants and nonmigrants equally (and this is indeed the case in many countries that have antidiscrimination laws) for the sake of horizontal equity.

The timing of the events is as follows. In the first stage, the governments simultaneously choose the tax schedules. In the second stage, the agents observe these tax schedules and decide which schedule to accept (equivalently, choose their labor-consumption pairs). The low type individuals are restricted to choose the tax menus from the country of their residence only; the high type individuals may also (at some cost) choose the tax menus offered by the other country.

Given a pair of taxes \((t^A_H, t^B_H)\) in two countries, if \( t^A_H < t^B_H \), all the individuals from country B with \( a < \hat{a} : u(z^A_H - t^A_H, z^A_H/\theta_H) - c(\hat{a}) = u(z^B_H - t^B_H, z^B_H/\theta_H) \) will migrate to country A; and analogously for country B. Correspondingly, now the Leviathan will want to maximize

\[
R^A = \gamma t^A_H \left( 1 + \int_0^{\hat{a}} dF(a) \right) + (1 - \gamma) t^A_L
\]

subject to the participation constraint

\[
u(\theta_L y_L - t_L, y_L) = u_0,
\]

the incentive compatibility constraint

\[
u(\theta_H y_L - t_L, y_L) \leq u(\theta_H y_H - t_H, y_H),
\]

which does not have to be binding any more, and individual rationality

\[
\theta_i (1 - t'_i) u_x + u_y = 0.
\]

\[3\text{To be concise, we do not explicitly consider the case with } t^A_H > t^B_H. \text{ However, it is easy to see that our formulation remains valid in this complementary case, if we additionally define functions } c \text{ and } F \text{ on the interval } [-1,0] \text{ by } c(-a) = -c(a) \text{ and } F(-a) = -F(a).\]
The solution to this program for given $t^B_H$ will give us a best-response function for country A. Writing this up a bit more explicitly, we have

$$\hat{a} = c^{-1} \left( u \left( z^A_H - t^A_H; z^A/\theta_H \right) - u \left( z^B_H - t^B_H; z^B/\theta_H \right) \right)$$

and the Lagrangian $L = \gamma t_H \left( 1 + \int_0^\hat{a} dF (a) \right) + (1 - \gamma) t_L + \mu (u (z_L - t_L, z_L/\theta_L) - u_0) + \lambda (u (z_H - t_H, z_H/\theta_H) - u (z_L - t_L, z_L/\theta_H))$, where $\lambda \geq 0$ and we omit superscript A for more parsimonious notation. The first order conditions are now

$$t_L : 1 - \gamma - \mu u^L_x + \lambda u^{HL} = 0,$$  

(3a)  

$$z_L : \mu \left( u^L_x + u^L_y/\theta_L \right) - \lambda \left( u^{HL} + u^{HL}/\theta_H \right) = 0,$$  

(3b)  

$$t_H : \gamma \left( 1 + \int_0^\hat{a} dF (a) - t_H f (\hat{a}) c^{-1} (\cdot) u^H_x \right) - \lambda u^H_x = 0,$$  

(3c)  

$$z_H : \gamma t_H f (\hat{a}) c^{-1} (\cdot) \left( u^H_x + u^H_y/\theta_H \right) + \lambda \left( u^H_x + u^H_y/\theta_H \right) = 0.$$

(3d)  

First, we can see that the conditions of the less productive are not affected by the migration possibility of the high-skilled. Second, the “no distortion at the top” result is still preserved, regardless of whether the IC constraint is still binding. Indeed, as in the last expression $\gamma t_H f (\hat{a}) c^{-1} (\cdot) + \lambda$ is strictly positive, it is necessary that at the optimum $u^H_x + u^H_y/\theta_H = 0$, that is, $t_H' = 0$. Third, the FOC with respect to $t_H$ has now an additional term $\int_0^\hat{a} dF (a) - t_H f (\hat{a}) c^{-1} (\cdot) u^H_x$. If the IC constraint were not binding, the choice of the tax on high-productivity individuals would be a simple trade-off between increasing the tax base and reducing the tax rate to maximize revenue. Otherwise, relaxing the IC constraint is an additional benefit of decreased tax:

$$1 + \int_0^\hat{a} dF (a) = t_H f (\hat{a}) c^{-1} (\cdot) u^H_x + \frac{\lambda}{\gamma} u^H_x.$$  

The shadow value of the constraint is changed from $\gamma/u^H_x$ in autarky to $\gamma \left( 1 + \int_0^\hat{a} dF (a) \right)/u^H_x - \gamma t_H f (\hat{a}) c^{-1} (\cdot)$ in the open economy.

If competition is very intense, the IC constraint may become nonbinding; in such a case the condition (3b) simplifies to $\mu \left( u^L_x + u^L_y/\theta_L \right) = 0$, and we have no distortion at the bottom: $u^L_x + u^L_y/\theta_L = 0$, as the participation constraint for the low-productivity individuals is binding\(^4\). This is a remarkable

---

\(^4\) As competition intensifies further, IC constraint of the low type may become binding. In this case “no distortion at the bottom” is preserved, but “no distortion at the top” disappears and there may be overprovision of hours worked for the high skilled. Our results in Propositions 1-3 remain unaltered.
result: in our model tax competition is a way to tame Leviathan, and it might even restore first-best solution in some cases.

**Example 1** In the extreme case of no switching costs, Bertrand competition decreases the tax on the high productivity type to zero. There is no distortion for the low type; its IC constraint may be binding.

**Proof.** In Bertrand equilibrium both governments get the first best revenue from their low type residents. Clearly, a deviation to any other tax on the “poor” is not profitable. A deviation to a higher tax on the “rich” does not change the revenue, because all the “rich” emigrate. A deviation to a negative tax on the “rich” decreases the revenue. □

The best response of country A is defined by the equations (3a)–(3d) and (2). By the inverse function theorem, \( c^{-1} (.) = 1 / c' (.) \). We now look at a symmetric (subgame-perfect) Nash equilibrium, defined by the pair of best responses \( t_A^H (t_B^H) \) and \( t_B^A (t_A^H) \) such that \( t_A^H = t_B^H = t_H^0 \). The condition (3c) can be rewritten as

\[
 t_H^0 = \frac{c' (0)}{f (0)} \left( \frac{1}{u_H^X} - \frac{\lambda}{\gamma} \right),
\]

and together with the conditions (3a)–(3d) it defines a symmetric Nash equilibrium in our model.

Notice that \( c' (0) \) reflects intensity of competition: for \( c' (0) = 0 \) there is no heterogeneity with respect to migration decision, so there is effectively Bertrand competition; for \( c' (0) \to \infty \) competition becomes ineffective, and we have the following lemma.

**Lemma 1** Consider autarky equilibrium tax rates \( \{ t_L^0, t_H^0 \} \). For \( c' (0) \to \infty \), the unique symmetric equilibrium in the tax competition game converges to \( t_L^0 = t_L^H, t_H^0 = t_H^H \).

**Proof.** Starting from autarky equilibrium values, from (1c) \( \gamma - \lambda u_H^X = 0 \). The condition (3c) as a best response to autarky equilibrium in another country can be rewritten as \(-\gamma t_H u_H^X f (0) / c' (0) < 0\), so there is an incentive to cut the tax, but this incentive vanishes in the limit of unbounded slope of the switching cost function. □

Thus, the autarky equilibrium is a limiting case of open-economy equilibrium with no effective tax competition.

**Proposition 1** Tax competition lowers the tax on the high-skilled, \( t_H^0 < t_H^H \).
Proof. From Lemma 1 we see that $t^o_H \neq t^a_H$. Now $t^0_H > t^a_H$ is not feasible: since the IC constraint is binding in autarky, a tax higher than in autarky on the “rich” is not possible without increasing the tax on the “poor”. But if that were possible without violating their participation constraint, such an increase would have been also optimal in autarky. Thus, the only possible case is $t^0_H < t^a_H$. ■

Proposition 2 Tax competition increases employment of the low-skilled: $z^o_L > z^a_L$.

Proof. From Proposition 1 we know that $t^0_H < t^a_H$. If the IC constraint is binding, $dz^L/dt^H < 0$ along the set of binding constraints (see appendix), hence $z^o_L > z^a_L$. If the IC constraint is not binding, from the condition of no distortion at the bottom, $z^o_L > z^a_L$. ■

The propositions assume existence of the equilibrium, and we establish it in the appendix.

An interesting policy-relevant observation obtains immediately: tax competition contributes to the employment of low-skilled labor, which is obviously a virtue. While such an increase does not improve the lot of the low-skilled, tax competition benefits the high-skilled at the expense of Leviathan. Conversely, tax coordination (autarky in our model) would increase tax revenue, but would be inferior to tax competition in terms of the employment and utility of the high-skilled\textsuperscript{5}.

3 Alternative objectives of the government

3.1 Rawlsian government

Suppose now government is not interested in its own rents, but has Rawlsian preferences, that is, it wants to maximize the utility of the low-productivity individuals subject to some budget constraint. The corresponding Lagrangian is then

\[ L = u(z_L - t_L, z_L/\theta_L) + \lambda(u(z_H - t_H, z_H/\theta_H) - u(z_L - t_L, z_L/\theta_H)) + \mu\left(\gamma t_H \left(1 + \int_0^\delta dF(a)\right) + (1 - \gamma) t_L\right). \]

We immediately see that the structure of the problem does not change, so

\textsuperscript{5}It can be noted that tax competition is not necessarily welfare-improving in models of Leviathan governments. See Edwards and Keen (1996) for details.
the structure of the solution to it stays the same. The difference is that whereas Leviathan takes all the rents away from the “poor”, the Rawlsian government, to the contrary, maximizes them. The FOCs are now

\begin{align}
    t_L & : \mu (1 - \gamma) - u^L_x + \lambda u^H_x = 0, \\
    z_L & : u^L_x + u^L_y / \theta_L - \lambda \left( u^H_x + u^H_y / \theta_H \right) = 0, \\
    t_H & : \mu \gamma \left( 1 + \int_0^{\tilde{a}} dF(a) - t_H f (\tilde{a}) c^{-1'}(.) \right) u^H_x - \lambda u^H_x = 0, \\
    z_H & : \mu \gamma t_H f (\tilde{a}) c^{-1'}(.) \left( u^H_x + u^H_y / \theta_H \right) + \lambda \left( u^H_x + u^H_y / \theta_H \right) = 0.
\end{align}

To see that this set of FOCs is equivalent to (3a)–(3d), divide them through by $\mu$ and re-denote $\mu_1 = 1 / \mu$, $\lambda_1 = \lambda / \mu$. Then Lemma 1 and the no-distortion results go through. Proposition 1 still holds, because the Rawlsian government does not want to increase the tax on the “poor”, and an increase in the tax on the “rich” is not possible without violating the IC constraint. Proposition 2 then remains intact, as it uses Proposition 1, the IC constraint (or no distortion at the bottom), and the “no distortion at the top” results.

Intuitively, it makes little difference whether the government wishes to tax the high-skilled to maximize its own rent or the utility of the poor. In both situations mobility of the high-skilled tends to ease the self-selection constraint that the government has to respect, allowing the poor to be less rationed on the labor market.

While in the Leviathan model tax competition has kept the utility of the poor constant, in the Rawlsian model their utility goes down and only the utility of the high-skilled goes up.

### 3.2 Utilitarian government

Now consider the probably most popular formulation, in which the government wants to maximize the sum of the utility of the individuals. A problem here is that it is not clear whether the utility of new immigrants should enter the government’s objective\(^6\). Given that in reality obtaining citizenship is often a long and painful process, we assume that the government cares only about the established residents. Then the Lagrangian is

\[ L = \gamma u (z_H - t_H, z_H / \theta_H) + (1 - \gamma) u (z_L - t_L, z_L / \theta_L) \]

\(^6\)For a discussion see Mirlees (1982) and Simula and Trannoy (2009).
\[ +\lambda (u(z_H - t_H, z_H/\theta_H) - u(z_L - t_L, z_L/\theta_H)) + \mu \left( \gamma t_H \left( 1 + \int_{0}^{\hat{a}} dF(a) \right) + (1 - \gamma) t_L \right). \]

The corresponding FOCs are

\[
\begin{align*}
t_L & : \mu (1 - \gamma) - (1 - \gamma) u_L^L + \lambda u_L^{HL} = 0, \quad (6a) \\
z_L & : (1 - \gamma) \left( u_x^L + u_y^L/\theta_L \right) - \lambda \left( u_x^{HL} + u_y^{HL}/\theta_H \right) = 0, \quad (6b) \\
t_H & : \mu \gamma \left( 1 + \int_{0}^{\hat{a}} dF(a) - t_H f(\hat{a}) c^{-1'}(.) u_x^H \right) - (\lambda + \gamma) u_x^H = 0, \quad (6c) \\
z_H & : \mu \gamma t_H f(\hat{a}) c^{-1'}(.) \left( u_x^H + u_y^H/\theta_H \right) + (\lambda + \gamma) \left( u_x^H + u_y^H/\theta_H \right) = 0. (6d)
\end{align*}
\]

This is not exactly equivalent to the previous problem, but we can immediately see that the “no distortion at the top” result survives, and so does the “no distortion at the bottom” in the case of a nonbinding IC constraint. The same is true for Lemma 1. In Proposition 1, the government has no incentive to increase taxes, as it cares about the utility of the “rich”. Proposition 2 holds by the same reasoning as with Rawlsian government.

To sum up, our result about the effect of tax competition on employment of the “poor” is robust to the changes in the specification of government’s objective function.

## 4 Asymmetric countries

Suppose now that the two countries we consider are of different size. Assume that whereas country B still has population of measure 1, country A has a population of measure \( m > 1 \). Otherwise the countries are identical; in particular, \( a \) is still distributed on a unit interval, only in the country A every point is \( m \) times more populated.

The following two FOCs are changed for the Leviathan in country A (we consider here the more relevant case of \( t_H^A > t_H^H \)):

\[
\begin{align*}
t_L & : (1 - \gamma) m - \mu u_x^L + \lambda u_x^{HL} = 0, \quad (7a) \\
t_H & : \gamma \left( m - m \int_{0}^{\hat{a}} dF(a) - m t_H f(\hat{a}) c^{-1'}(.) u_x^H \right) - \lambda u_x^H = 0. \quad (7b)
\end{align*}
\]

For country B, the only equation altered is

\[
t_H : \gamma \left( 1 + m \int_{0}^{\hat{a}} dF(a) - m t_H f(\hat{a}) c^{-1'}(.) u_x^H \right) - \lambda u_x^H = 0. \quad (8)
\]
We see that, compared to the symmetric situation, the relative importance of tax competition terms is increased for the small country (B) and reduced for the large country (A). Notice that Propositions 1 and 2 do not hinge on the symmetry assumption, so they are still valid in the asymmetric setup. The existence proof, however, uses symmetry and has to be reestablished (see appendix).

Intuitively, the small country is more aggressive in tax competition, since it has more to gain (through attracting a foreign tax base) and less to lose from it (through reduced taxes from the home tax base). This is confirmed by the following proposition:

**Proposition 3** In equilibrium of the asymmetric game, $t_A^H > t_B^H$. 

**Proof.** Suppose the contrary is true. The case of $t_A^H = t_B^H$ is clearly inconsistent with the sets of FOC above. Consider the case of $t_A^H < t_B^H$. In equilibrium the gain of each country from marginally changing the tax rate should be zero. Consider, for example, country B:

$$\frac{dR_B}{dt_B^H} = \gamma \left( 1 - \int_0^{\hat{a}} dF(a) - t_B^H \frac{f(\hat{a})}{c'(\hat{a})} \right) + \left(1 - \gamma\right) \frac{dt_L^B}{dt_H^B} = 0,$$

where $dt_L^B/dt_H^B$ is taken along the binding constraints. There may be 4 cases that differ according to which constraints are binding.

Case 1: the set of binding constraint is identical across countries. Hence, $dt_L^A/dt_H^A = dt_L^B/dt_H^B$. Substituting this into the change in revenue for country A, we get

$$\gamma \left( m + \int_0^{\hat{a}} dF(a) - t_A^H \frac{f(\hat{a})}{c'(\hat{a})} \right) + \left(1 - \gamma\right) m \frac{dt_L^A}{dt_A^H} = \gamma \left( (m + 1) \int_0^{\hat{a}} dF(a) + (mt_H^B - t_A^H) \frac{f(\hat{a})}{c'(\hat{a})} \right) > 0,$$

so country A is better off raising its tax. Thus, the case $t_A^H < t_B^H$ cannot be an equilibrium.

Case 2: the IC constraint of the high type is not binding for the small country only. The high skilled must be better-off in country B, hence it must be that $t_A^H > t_B^H$. 

12
Case 3: the IC constraint of the high type is not binding only in the big country. We have $dt^A_L/dt^A_H = 0$ and hence from revenue maximization

$$t^A_H = \left( m + \int_0^{\hat{a}} dF(a) \right) \frac{c'(\hat{a})}{f(\hat{a})},$$

$$t^B_H = \left( 1 - \int_0^{\hat{a}} dF(a) + \frac{1 - \gamma}{\gamma} dt^B_L/dt^B_H \right) \frac{c'(\hat{a})}{f(\hat{a})}.$$

As $dt^B_L/dt^B_H < 0$, we must have $t^A_H > t^B_H$.

Case 4: the IC constraint of the high type is not binding in both countries; the “no distortion at the top” is binding in one country, whereas the ICC for the low type is binding in the other country. Since the ICC of the high type is not binding, we have $dt^A_L/dt^A_H = dt^B_L/dt^B_H = 0$ and hence from revenue maximization

$$t^A_H = \left( m + \int_0^{\hat{a}} dF(a) \right) \frac{c'(\hat{a})}{f(\hat{a})},$$

$$t^B_H = \left( 1 - \int_0^{\hat{a}} dF(a) \right) \frac{c'(\hat{a})}{f(\hat{a})}.$$

Clearly, $t^A_H > t^B_H$.

Thus, for any configuration of the sets of binding constraints, we arrive at contradiction with our assumption that $t^A_H < t^B_H$. □

While more aggressive behavior of the smaller country is a robust result in tax competition models (e.g., Haufler 2001, ch. 5), Proposition 3 allows us to formulate a new testable hypothesis: The positive effect of opening borders on employment of low-skilled workers is more pronounced in a small country.

5 Conclusion

We have analyzed tax competition in a simple optimal-income-taxation model. We show that the tax on the high-skilled decreases and employment of the low-skilled increases with respect to autarky. Our results are robust to a number of modifications concerning the government’s objective function and symmetry of the two competing countries.

There are important limitations that we share with many optimal-taxation models. First, there is no account of capital, although it should be even more
mobile than high-skilled labor. We focus on income taxation because we want to clearly identify the effect of combining competition with the principal–agent framework that underlies optimal taxation models. Second, due to the simple linear production technology in one good economy, there are no general-equilibrium or trade effects of the wage changes that could lead to repercussions on the effects discussed.

We see several new directions for future research in the framework we have considered. Extensions of our model could assume countries that differ with respect to the national objective function or could allow for some mobility of low-skilled workers. We also hope that this paper will encourage empirical work on the labor market effects of migration opportunities. Based on our model, we would expect that tax competition for mobile high-skilled workers has more pronounced implications for low-skilled workers in small countries.

6 Appendix

6.1 To Proposition 2

For the proof of proposition 2 we need to establish that lower \( t_H \) implies higher \( z_L \).

Consider a change in employment that the government offers for the low type, \( dz_L \). Along the participation constraint, we have

\[
    u_x \left( 1 - \frac{dt_L}{dz_L} \right) + u_y \frac{1}{\theta_L} = 0.
\]

Since \( \theta_H > \theta_L \),

\[
    u_x \left( 1 - \frac{dt_L}{dz_L} \right) + u_y \frac{1}{\theta_H} > 0.
\]

IC constraint of the high type is binding. Therefore, as a result of an increase in \( t_H \), the change in the utility of the high type is the same when honest or mimicking:

\[
    du (z_H - t_H, z_H/\theta_H) = du (z_L - t_L, z_L/\theta_H) < 0.
\]

In terms of total derivative we have

\[
    \frac{du^H}{dtH} = \left( u_x \left( 1 - \frac{dt_L}{dz_L} \right) + u_y \frac{1}{\theta_H} \right) \frac{dz_L}{dtH} < 0.
\]
Since the expression in the brackets is positive, we must have $dz^L/dt^H < 0$, Q.E.D.

6.2 To the corner solution

The consequences of competition for mobile labor that we have analyzed suggest that the Leviathan may want to force the “poor” not to work, $z_L = 0$, and the “rich” to work as much as possible and to tax them as much as possible as well. This depends on the value of $\gamma$.

Remark 1 For sufficiently high $\gamma$, a tax-revenue-maximizing allocation is characterized by $z^a_L = 0$.

Proof. Suppose $z^a_L > 0$. Then a small reduction in $z_L$ will lead to an increase in $z_H$ that will keep the IC constraint satisfied. From the participation and “no distortion at the top” constraints, that will also reduce $t_L$ and increase $t_H$ by amounts $dt^L/dz^L < 1$ and $dt^H/dz^H > 1$ correspondingly. Obviously, as long as $\gamma > \frac{dt^L/dz^L - dz^H/dz^H - dH/dt^L}{dt^L/dz^L}$, such a change will increase tax revenue without violating any constraint. Thus, at the optimum $z^a_L = 0$. 

Note that in situations with $z^a_L = 0$ we have $t^a_L < 0$ from the participation constraint, that is the “poor” receive a subsidy.

6.3 On the Existence of Equilibrium

6.3.1 Symmetric model

The existence of the equilibrium in our game hinges on two assumptions: (i) the conditions (3a)–(3d) and (2) define best responses; (ii) the intersection of these best responses is nonempty. For (i) it is necessary and sufficient that the conditions define an interior solution and second-order conditions are satisfied (they actually are, given our assumptions on the utility function and appropriate assumptions on the functions $c$ and $F$, plus any parameter restrictions that ensure an interior solution). For (ii) we have to study the best response on the interval $[0, t^a_H]$. By proposition 1 it is necessary and sufficient that the best response intersect the $45^o$ line on this interval. There are no discontinuities in our problem, so the best-response function must be continuous. As we have shown, $BR(t^a_H) < t^a_H$. On the other hand,

\footnote{The latter under a technical condition $u^H_{xy} < \min \{-\theta_H u^H_{xx}, -u^H_{yy}/\theta_H\}$.}
\[ BR(0) \geq 0, \] as a negative tax on the rich can not be revenue-maximizing. By continuity then there exists an intersection (or intersections) with the 45° line on the interval \([0, t_H^a]\), and hence an equilibrium exists. Moreover, this equilibrium (or equilibria) is symmetric, because the best responses are identical.

6.3.2 Asymmetric model

We follow the same logic as for the symmetric situation. The appropriate parameter restrictions ensure that conditions from (i), modified correspondingly as in (7a)–(8), define best responses. For (ii), we need the two best responses to intersect. It is still true that \( BR(t_H^a) < t_H^a \) and \( BR(0) \geq 0 \) for each country, so by continuity there exists at least one intersection on the interval \([0, t_H^a]\).

6.4 On the Uniqueness

The symmetric equilibrium analyzed is unique whenever the system of equations (3a)–(3d), (2), and \( t_A^H = t_B^H \) has a unique solution.

6.5 Inexistence of closed economy equilibrium with binding IC constraint for the low productivity type

In the closed economy setup with Leviathan government, suppose the IC constraint for the low type is binding. Formally, \( L = \gamma t_H + (1 - \gamma) t_L + \mu \left( u(x_L - t_L, \theta_L) - u_0 \right) + \lambda \left( u(x_L - t_L, \theta_L) - u(x_H - t_H, \theta_L) \right), u^{LH} := u(x_H - t_H, \theta_L) \). The FOCs are

\[
\begin{align*}
    t_L & : 1 - \gamma - \mu u_x^L - \lambda u_x^L = 0, \\
    z_L & : (\mu + \lambda) \left( u_x^L + u_y^L / \theta_L \right) = 0, \\
    t_H & : \gamma + \lambda u_x^{LH} = 0, \\
    z_H & : -\lambda \left( u_x^{LH} + u_y^{LH} / \theta_L \right) = 0.
\end{align*}
\]

Since \( \gamma > 0, \lambda > 0 \) and \( u_x^{LH} > 0 \), the third line can never be satisfied. Since \( L_{t_H} = \gamma + \lambda u_x^{LH} > 0 \), it would be optimal for the government to set the highest possible tax on the high productivity type, thus hitting its participation constraint. This, however, cannot be the case, as “the rich”
would prefer to mimic the poor. Thus, our initial assertion that the IC constraint for “the poor” is binding must be wrong.

Another way to see this is through the “no distortion at the bottom” result that characterizes an equilibrium with a binding IC constraint for the low productivity type (line two on the display). At the optimal $t_L$, $dx/dz\big|_{u^L = u_0} = 1$. At optimal $t_H$, $dx/dz\big|_{u^L = u_0} > 1$, as otherwise the high type would want to mimic the lower type. Clearly, the government could offer the high type the same tax scheme as the lower type and get higher revenue than before (tax revenue is maximized at $dx/dz = 1$ given a condition $u = u_0$). Such a pooling situation would certainly also not be an optimum, as the government could get higher revenue offering the “rich” a menu for which their IC constraint is binding. Thus, an equilibrium with “no distortion at the bottom” and a binding IC constraint of the “poor” cannot exist, if the government objective is revenue maximization.

Intuitively, it may seem surprising that regardless of the value of $\gamma$, i.e. also when there are very few “rich”, the government will prefer not to distort their labor supply decision at the expense of distorting the choice of the “poor”. The thing is that as $\gamma$ approaches zero, the distortion imposed on the lower type also vanishes. Formally, from (1a)-(1c) we have

$$t_{L}^{'} = \frac{\gamma (u^H_{x} + u^H_{y} / \theta_H)}{(1 - \gamma) u^H_{x} + \gamma u^H_{x}},$$

that is the marginal tax rate on the “poor” approaches zero as their share approaches unity. Thus, we have the “no distortion on the bottom” result in the limit, but never for $\gamma > 0$.

The inexistence result is robust to the changes in the government objective considered in this paper. Indeed, a Rawlsian or Utilitarian government will always prefer to keep the IC constraint of the high (rather than the low) type binding (in the former case because they only care about the poor; in the latter case on pure efficiency grounds).

**References**


