

A Positive Model of Overlapping Income Taxation in a Federation of States*

Esteban F. Klor[†]

W. Allen Wallis Institute of Political Economy
107 Harkness Hall, Box 27-0158
University of Rochester
Rochester, NY 14627

Abstract

This paper develops a positive theory of overlapping income taxation in a federation of states. Its main motivation comes from the observation that in the U.S. states income tax rates are significantly lower than the federal income tax rate. The analysis shows that in a federal system total productivity dispersion between the states determines the federal tax rate. In fact, there exists a positive relation between the level of productivity dispersion and the federal tax rate, even if the income of the decisive voter is above the mean income. When the individuals' income is endogenous, the higher the implemented federal tax rate is, the lower the resulting state tax rate will be, even if the decisive voter at the state level has zero pre-tax income. Empirical evidence obtained from a panel data set on tax schedules at the state level supports the main hypothesis of the paper. Most notably, the data points to the existence of a significant trade-off between the states tax rates and the federal tax rate, explained through productivity dispersion between the states.

KEYWORDS: Fiscal Federalism, Political Economy, Income Taxation.
JEL Classification: D72, H23, H77

*I am grateful to Sophie Bade, Ronny Razin, and especially to Alessandro Lizzeri, Antonio Merlo and Efe Ok for helpful discussions. The paper has benefited from the comments of seminar participants at CUNY, Georgetown University, Hebrew University, New York University, Stanford University GSB, Tel Aviv University and University of Pennsylvania, as well as participants at the conferences of the Social Choice and Welfare Society (Pasadena, 2002), the Society of Economic Design (New York, 2002) and the Society of Public Choice (San Diego, 2002). All views and mistakes in the paper remain my own.

[†] E-mail: efk201@nyu.edu. <http://home.nyu.edu/~efk201>.

1 Introduction

The main motivation for this paper comes from the observation that states income tax rates in the U. S. are significantly lower than federal income tax rates. According to the TAXSIM model, in the last twenty five years the average effective marginal income tax rate at the federal level oscillates around 30% while for a majority of states the average effective state income tax rate is below 5%, with nine states during the studied period exhibiting a zero income tax rate.¹

From a political economy standpoint the extremely low income tax rates at the state level are a puzzling observation. It is well established in the positive literature of income taxation that if the income of the median voter is below the mean (this is the case for every state within the U.S.), then a majority of voters (namely those whose income is less than the mean) should prefer large scale expropriation and redistribution. Several explanations exist for the fact that, in practice, the rich are not expropriated through the tax system, the most prominent related to the deadweight loss from taxation (Meltzer and Richard, 1981). But even if voters take into account the deadweight loss from taxation, the low levels of state income tax rates exhibited in the U.S. are difficult to explain. (This is especially the case for the nine states that have never implemented a positive income tax rate.) Of course, it would be a mistake to treat states, which belong to a federation, as separate political entities. While it is the case that in the U.S. states are free to impose their own income tax rates, that freedom is restricted by the very existence of a federal income tax. For in a federal system where tax bases are joint property, state and federal tax settings decisions are interdependent.

This paper builds a simple model of taxation and redistribution in a two-tier federal system consisting of a single central government and two state governments. The federation's political process works as follows. In the first stage individuals vote for a federal tax schedule that applies to all the residents of the federation, regardless of their state of residence. At a second stage residents of each state vote over tax schedules in that particular state. Governments at all levels use only linear tax-transfer schedules to redistribute income. The political mechanism considered for all the elections is majority rule.

In the model, individuals are endowed with a productivity level and choose the amount of labor they supply as a function of the selected tax schedules. This introduces a trade-off between the level of output and its distribution, as was first modeled in a political economy context by Romer (1975), Roberts (1977), and Meltzer and Richard (1981). The point of departure of this paper is that individuals, who are assumed immobile, reside in two different states and face overlapping taxes on their income. This framework creates a new source of heterogeneity in which individuals differ not only with respect to their productivity level, but also with respect to their state of residence.² Consequently, new

¹The TAXSIM data referred to is available at <http://www.nber.org/~taxsim>.

²This paper abstracts from mobility of individuals between the states. This assumption allows me to highlight a different mechanism influencing federal and states' income tax schedules. See the conclusions for a detailed discussion of the influence of mobility on the results

considerations (besides their own income) influence the individuals' preferences over tax schedules.

When forming their preferences over the state and federal income tax schedules, individuals know that significant differences exist between these taxes. First, individuals are aware that state and federal tax settings decisions are interdependent. In particular, individuals rationally anticipate that, because of incentive considerations, the higher the implemented federal tax rate is, the lower the implemented state's tax rate will be, even if the decisive voter at the state level has zero pre-tax income. The second important difference is that while the state tax schedule redistributes income *within* each state, the federal tax schedule implicitly redistributes income *between* the states.

The results show that the existence of productivity dispersion between the states plays a crucial role in the analysis. Residents of the relatively rich state always oppose a positive federal tax rate. Individuals with low productivity residing in the rich state prefer a zero federal income tax rate in order to maximize the redistribution at the state level. Individuals with high productivity in this state also oppose federal taxation but simply because taxation, in any form, reduces their utility. In contrast, residents of the poor state favor a positive federal tax rate, its level depending on each individual's productivity level.

Since individuals' preferences over the federal tax schedule are not monotonic in their productivity level, a coalition of poor individuals (which constitute a majority of the federal population) never emerges. In fact, the income of the decisive voter at the federal level is always above the median federal income and may even be above the mean federal income. This voter's preferred federal tax rate is an increasing function of productivity dispersion between the states. So if this relatively rich individual is from a relatively poor state she will support a positive federal income tax rate. The objectives of this individual are twofold. First, a positive federal tax rate redistributes income toward the individual's state. And second, the demand for state redistribution decreases as the federal tax rate increases. This trade-off between the federal tax rate and the state tax rate thereby provides another reason for high productivity individuals in the poor state to support a high federal tax rate, ultimately bringing a low state tax rate.

From an efficiency standpoint, a federal social contract allowing a two-tier income taxation system is, in general, not optimal.³ Although the federal tax rate has a significant impact on the equilibrium income tax schedule in both states, this externality is partially ignored under a decentralized system of decision-making. In particular, individuals in one state do not take into account the impact of the federal tax schedule on redistribution in the other state. Obviously, a policy that takes this externality into account might achieve a welfare improvement for all the residents of the federation, with greater redistribution and lower taxation.

of the paper.

³A similar conclusion was obtained in the related normative literature (see Gordon, 1983; Johnson, 1988; Wildasin, 1991; Boadway et al., 1998).

This provides a possible role for the federal government: to implement policies that undo the nonoptimal outcome that arises from decentralization. Gordon (1983) and Wildasin (1991) argue that one such policy is the implementation of a system of federal matching grants whereby the federal government shares a proportion of the cost of states' redistribution. The current paper finds that such a system would tend to decrease rather than increase total welfare of the federation's residents. The reason being that a federal matching grant system reduces the cost of state redistribution paid by the state's population. Consequently, individuals support higher income tax rates to finance greater redistribution at the state level. This in turn implies a high cost in federal matching grants, which have to be covered with high federal tax rates. Thereby, the implementation of this system causes higher levels of taxation and lower levels of redistribution.⁴

In the last section of the paper I estimate a number of hypotheses derived from the theoretical model at hand. Most notably, this empirical exercise corroborates that productivity dispersion between the states has a statistically significant positive impact on the level of federal income taxes. Furthermore, it is found that an increase in the federal income tax leads to a decrease in the states income taxes. These findings are in concert with the implications of the present model.

2 The Model

Consider a federation of two states, A and B (the analysis is easily generalized to any number of states). There is a unit mass of individuals living in the federation, a share p_A of them resides in state A . Individuals cannot move between the states. Each individual is endowed with a productivity level w and has no non-labor wealth. Thus, individuals are heterogeneous with respect to their productivity level and their state of residence. The population of each state is divided into two classes; in each state $i = A, B$ there is a mass $n_i^l > 1/2$ of low-productivity individuals (with productivity equal to zero), and a mass $n_i^h = 1 - n_i^l$ of high-productivity individuals with $w_i > 0$.

Individuals choose the amount of labor they provide on a competitive market. Income is measured in units of consumption and is produced using a constant returns to scale technology. Hence, an individual with productivity $w > 0$ who supplies y/w units of labor earns pre-tax income y .

The federation has a two-tier taxation system: there is a federal and a state income tax schedule. Both tiers impose linear taxes that are used to collect revenues. These revenues are redistributed lump-sum to the population of individuals that are subject to that particular tax. The political system of the federation is such that the federal tax is imposed first on the individuals' pre-tax

⁴This result contrasts with the one obtained in Wildasin (1991). In his paper, a higher-level government using corrective matching grants achieves a welfare improvement. In the present model, however, I assume that all tax schedules are chosen by majority rule. This political mechanism increases the externalities arising from decentralized decision-making, and thus the inefficiency of a system of overlapping income taxation.

income; later on every state imposes its own tax schedule on the remaining of the individuals' pre-tax income.⁵

Formally, the *federal tax schedule* is represented by a tax rate $f \in [0, 1]$ and a redistribution level $r_f \in \mathbf{R}_+$ such that the federal budget constraint,

$$r_f = f [p_A n_A^h y_A + p_B n_B^h y_B],$$

is satisfied. Similarly, the *tax schedule of state i* is given by the tax rate $s_i \in [0, 1]$, where the state's redistribution level $r_i \in \mathbf{R}_+$ is obtained from

$$r_i = s_i(1 - f)p_i n_i^h y_i.$$

The individuals' net income is

$$c_i = (1 - s_i)(1 - f)y_i + r_f + \frac{r_i}{p_i}.^6 \quad (1)$$

Given both tax schedules, an individual with productivity w_i chooses pre-tax income $y_i(w_i, s_i, f)$ to maximize $u(c, \frac{y}{w})$ subject to (1). Throughout the paper I assume the following quasi-linear preferences over consumption and labor supply:

$$u(c, \frac{y}{w}) = c - \frac{\alpha}{\beta + 1} \left(\frac{y}{w}\right)^{\beta + 1}, \quad c, y \geq 0, \quad (2)$$

where α is a positive constant and $(1/\beta)$ is the (constant) elasticity of labor supply. Under this class of preferences, redistribution at either governmental level does not affect labor supply decisions and every individual with positive productivity level chooses to work.

While this is a highly restrictive specification of preferences, it captures the incentive effects of taxes (consumption-leisure trade-off). Moreover, this specification removes a source of considerable complication in the analysis that follows. In particular, under a general specification, redistribution may induce productive individuals to refrain from working. If every individual in a state has zero pre-tax income, then a continuum of equilibrium tax rates for this state exists. This indeterminacy complicates the analysis when solving for the equilibrium federal tax rate. Finally, the specification above (which is widely used in studies of income taxation; cf. Diamond, 1998; Bohn and Stuart, 2001; De Donder and Hindricks, 2001) is considerably more tractable than a general one, allowing us to obtain clear, intuitive results.

⁵As an objection to the previous assumption one may argue that the current practice in the U.S. is that both taxes are paid simultaneously. Assuming that taxes are paid simultaneously would not change the nature of any of the following results. The adoption of the sequential timing of the events, which is common in the related literature (cf. Boadway et al., 1998), only tries to reflect the strategic considerations of the residents of a given state when choosing their own state tax schedule. When doing so, it is reasonable to suppose that these residents take the federal tax schedule as given to them.

⁶This particular form of individuals' net income is a consequence of the sequential timing in which the taxes are imposed. Had I assumed instead that both taxes are imposed simultaneously, we would have obtained that $c_i = (1 - s_i - f)y_i + r_f + r_i/p_i$ (see Gouveia and Masia, 1998). As already pointed out, adopting this different specification will not change the nature of any result of this paper.

For these preferences, given the federal and state tax schedules, the optimal pre-tax income of an individual with productivity $w_i > 0$ is

$$y_i = w_i \left[\frac{(1 - s_i)(1 - f)w_i}{\alpha} \right]^{\frac{1}{\beta}}. \quad (3)$$

Note that under the assumed preferences pre-tax income is not a function of redistribution, either at the federal or state level.

The next section solves for the equilibrium state and federal tax schedules.

3 Federal and States' Tax Schedules

Voting takes place in two stages: first a federal income tax schedule is chosen, and afterwards each state chooses its own state tax schedule, taking the elected federal tax schedule as given. Taxes at both the federal and state levels are determined according to majority rule equilibrium. Whenever the above criterion is satisfied by several tax schedules each is implemented with equal probability.

This section first solves for the states' tax schedules and then for the equilibrium federal tax schedule because individuals anticipate the effects of the federal tax on the state tax.

3.1 Preferences Over States' Tax Schedules

Given the individuals' pre-tax income the federal level of redistribution is

$$r_f = f \left(\frac{1 - f}{\alpha} \right)^{\frac{1}{\beta}} \sum_{i=A,B} p_i n_i^h w_i [(1 - s_i)w_i]^{\frac{1}{\beta}}, \quad (4)$$

and state i 's level of redistribution is given by

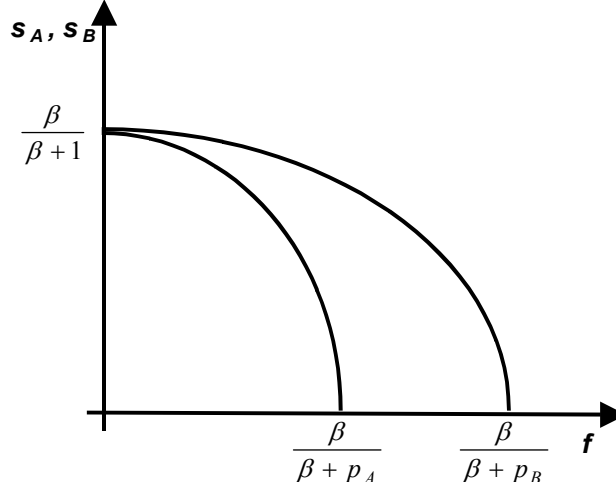
$$r_i = s_i(1 - f)p_i n_i^h w_i \left[\frac{(1 - s_i)(1 - f)w_i}{\alpha} \right]^{\frac{1}{\beta}}. \quad (5)$$

Since individuals with low productivity constitute a majority of the population in each state, they choose which tax schedule will be implemented in their state. To find her preferred state tax schedule, a low-productivity individual from state i maximizes her indirect utility function over the set of feasible state's tax schedules. This individual's indirect utility function is obtained by substituting equations (4) and (5) back into (2):

$$V_i^l = s_i(1 - f)p_i n_i^h w_i \left[\frac{(1 - s_i)(1 - f)w_i}{\alpha} \right]^{\frac{1}{\beta}} + f \left(\frac{1 - f}{\alpha} \right)^{\frac{1}{\beta}} \sum_{i=A,B} p_i n_i^h w_i [(1 - s_i)w_i]^{\frac{1}{\beta}}.$$

The solution to the associated maximization problem follows. All proofs are in Appendix A.

Figure 1: States' Tax Reaction Function (for $p_A > p_B$)



Lemma 1. *State i 's equilibrium tax rate s_i is characterized by*

$$s_i = \begin{cases} \frac{\beta}{\beta+1} - \frac{fp_i}{(1-f)(\beta+1)} & \text{if } f < \frac{\beta}{\beta+p_i}, \\ 0 & \text{otherwise.} \end{cases}$$

This lemma illustrates the relation between federal and states' tax rates: *states tax rates are decreasing in the federal rate* (see Figure (1)). An existing trade-off between state redistribution and federal redistribution is the main reason behind this result. Notice that the state's federal tax bill falls when taxable incomes in the state are reduced by increased state's tax rates. Consequently, federal redistribution is decreasing in s_i . On the other hand, the state level of redistribution increases in s_i for any $s_i < \beta/(\beta+1)$. Therefore, only when f (and consequently r_f) is equal to zero low-productivity individuals choose to maximize state redistribution. To choose such a high state tax rate when the federal tax rate is positive would be too costly in terms of federal redistribution. So as f increases the equilibrium state tax rate decreases, and the federation shifts redistribution from the state level to the federal level. When $f = \beta/(\beta+p_i)$, a positive state tax rate hurts federal redistribution more than what it contributes to state redistribution. Thus, when the federal tax rates is above $\beta/(\beta+p_i)$, low-productivity individuals choose a zero state tax rate.

In addition, the state's income tax rate decreases with the state's share of the overall population p_i . As p_i increases (for a fixed f) more of the total income in state i is transferred from the state to the federal level and will be

used for redistribution *between* the states. Therefore, less income is available for redistribution *within* the state. A reduction in the state income tax rate is required to partially offset this disincentive.

Finally, note that s_i is a concave function of f . This is because of the sequential structure of the taxation process. From equation (3) follows that the individuals' optimal pre-tax income is not a linear function of the sum of the tax rates. Due to this multiplicative structure, low-productivity individuals have to compensate high-productivity individuals with more significant decreases in s_i (for a given increase in f) as the level of f increases.⁷ Had low-productivity individuals not compensated high-productivity individuals this way, the state's total income would decrease considerably with f , and consequently the state's level of redistribution would decrease as well.

The following subsection solves for the equilibrium federal tax rate.

3.2 Preferences Over the Federal Tax Schedule

In the first stage a federal tax schedule is chosen by majority rule. The individuals' preferences over the federal tax rate are obtained by maximizing their indirect utility function over the set of feasible federal tax schedules, subject to the states' reaction functions found in Lemma 1. The resulting preferences over f are a function of the total productivity dispersion between the states. Let us define

$$x \equiv \frac{n_B^h}{n_A^h} \left(\frac{w_B}{w_A} \right)^{\frac{\beta+1}{\beta}} \quad (6)$$

as the measure of dispersion in total productivity between the states. When x is close to one, productivity dispersion between the states is relatively low. The farther away x is from one, the more unequal the states' total productivity levels.

The dispersion index x combines the original inequality in productivity between the states together with the elasticity of labor supply. For high values of β the relative importance of the individuals' pre-tax income is low, and dispersion between the states is mainly determined by the ratio of the share of high-productivity individuals. As the elasticity of labor supply increases, the difference between the individuals' productivity plays a more significant role in the resulting dispersion between the states. Whenever both states impose the same tax rate x is equal to the ratio of the states' total income.⁸

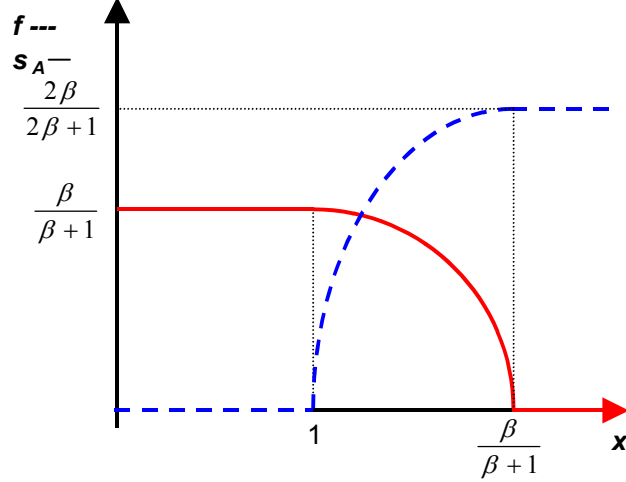
⁷Mathematically, totally differentiating y and setting $dy = dw = 0$ we obtain

$$0 = (1 - s)ds + (1 - f)df,$$

which implies that $ds/df < 0$. From a second differentiation of the last expression we obtain that also $\frac{d^2s}{df^2} < 0$.

⁸When the analysis is extended to m states, state j 's relevant measure of productivity dispersion is $x_j \equiv \left(\sum_{i=1}^m n_i^h w_i^{\frac{\beta+1}{\beta}} - n_j^h w_j^{\frac{\beta+1}{\beta}} \right) / (m - 1)n_j^h w_j^{\frac{\beta+1}{\beta}}$.

Figure 2: Preferred Federal Tax Rate for Low-Productivity Individuals from State A.



The proposition below presents the preferred federal tax rate for low-productivity individuals from state A when the population is evenly distributed between the two states. I adopt this simplifying assumption from now on because it allows me to obtain closed-form solutions that highlight the basic forces at work in the model.

Proposition 1. *The preferred federal income tax rate for low-productivity individuals from state A , f_t^A , is*

$$f_t^A = \begin{cases} 0 & \text{if } x \leq 1, \\ \frac{2\beta(x-1)}{x+\beta(x-1)} & \text{if } 1 < x \leq \frac{\beta+1}{\beta}, \\ \frac{2\beta}{2\beta+1} & \text{if } \frac{\beta+1}{\beta} < x. \end{cases}$$

Figure (2) depicts these preferences. To understand the intuition behind Proposition 1 we need to analyze the impact of the federal income tax rate on the utility of low-productivity individuals. On the one hand, a positive federal tax rate implies a positive level of federal redistribution. On the other, more federal redistribution implies less state redistribution. Therefore, low-productivity individuals evaluate which tax rate they should increase to maximize their own utility, knowing that in equilibrium the other tax rate will decrease.

Suppose, for example, that the total productivity dispersion between the states is less than one. If the federal tax rate is positive, low-productivity

individuals residing in A will receive federal redistribution. A positive federal tax rate, however, has two negative effects on these individuals' utility. First, it implicitly transfers income from state A to state B ; and second, it lowers their state's redistribution level. If instead the federal tax rate equals zero, low-productivity individuals in A appropriate to themselves part of the transfer between the states through their state income tax schedule. Low-productivity individuals in state A prefer this last alternative as it maximizes their total income. If it is the case that $x > 1$ low-productivity individuals in A are now the recipients of redistribution between the states. Thus, they prefer a positive federal tax rate.

The level of their preferred federal tax rate is determined by the productivity dispersion between the states. As x increases, the gains from redistribution between the states for low-productivity individuals residing in A increase as well.⁹ Therefore, they prefer a higher federal tax rate. This explains why f_t^A is increasing in x . The cost that this group pays for these gains is a lower redistribution at the state level. Eventually s_A reaches zero and there is no more room to trade-off an increase in the federal rate for a decrease in their state's rate of income taxation, even for greater levels of productivity dispersion between the states. A further increase of the federal tax rate above this level (without an accompanying decrease in s_A) has a large disincentive effect, lowering federal redistribution. This defines the second threshold value of x , above which both the federal and state income tax rates are constant, namely $f_t^A = 2\beta/(2\beta + 1)$ and $s_A = 0$.

Due to the symmetry of the analysis, the preferences of low-productivity individuals in state B are the exact opposite to the ones presented in Proposition 1. Their preferred federal income tax rate, f_t^B , is decreasing in x and the relevant threshold values of productivity dispersion are the inverse of the ones found for f_t^A . In particular, f_t^B reaches a maximum of $2\beta/(2\beta + 1)$ when $x < \beta/(\beta + 1)$, and a value of zero when productivity dispersion is greater than or equal to one. In the intermediate range $f_t^B = 2\beta(1 - x)/[1 + \beta(1 - x)]$.

From the previous argument follows that preferences over f are not monotonic in the individuals' productivity level; i.e., individuals with the same productivity level residing in different states have opposing preferences over the federal tax rate.¹⁰ Consequently, low-productivity individuals (which constitute a majority of the population) never form a coalition to extract as much income as possible from high-productivity individuals. In other words, the individual with the median productivity level is not the decisive voter in this framework. Therefore it is necessary to study high-productivity individuals' preferences over f to find out whether some consensus may emerge between the different groups in the federation. Only under such a consensus a federal tax rate able to reach the required support of at least half of the population against any other feasible tax rate will exist. The preferences of high-productivity individuals from state A

⁹Whenever the implications of changes in x are analyzed, total productivity at the federal level ($n_h^A w_A + n_h^B w_B$) is assumed constant.

¹⁰Note that, as the individuals' indirect utility function is strictly quasiconcave in f , their preferences over the federal tax rate are single-peaked.

appear next.

Proposition 2. *The preferred federal income tax rate for high-productivity individuals from state A, f_h^A , is*

$$f_h^A = \begin{cases} 0 & \text{if } x \leq \underline{x}, \\ \frac{2\beta(x - \underline{x})}{x + \beta(x - \underline{x})} & \text{if } \underline{x} < x \leq \bar{x}, \\ \frac{2\beta}{2\beta + 1} & \text{if } \bar{x} < x. \end{cases}$$

where $\underline{x} \equiv 1 + \frac{1}{(\beta+1)n_A^h}$ and $\bar{x} \equiv \frac{\beta+1}{\beta} + \frac{1}{\beta n_A^h}$.

It is readily seen from comparing the previous proposition to Proposition 1 that f_l^A and f_h^A are very similar. Indeed, both functions are increasing in x , meaning that also high-productivity individuals prefer a positive federal tax rate for high levels of productivity dispersion. In fact, the same intuition applies here as in Proposition 1, with one caveat. High-productivity individuals also derive a benefit and suffer a cost from the federal tax schedule. Unlike low-productivity individuals, the federal tax bill of high-productivity individuals is positive when the tax rate is positive, this is the cost. The benefits are both experienced at the state level (from a decrease of their state's tax rate) and at the federal level (from federal redistribution). For a high enough dispersion level the benefits exceed the costs, and thus the preference for positive federal tax rates.

The main difference between f_l^A and f_h^A is their threshold values. While f_l^A is positive for any x above one, f_h^A remains equal to zero until x reaches a higher value. To understand this remember from Lemma 1 that s is a concave function of f ; that is, when f is low an increase in f causes a relatively small decrease in s . Therefore, for low levels of x the gains that high productivity individuals in state A obtain from a positive federal tax rate are small compared to the losses they suffer in terms of higher overall taxation. As productivity dispersion between the states increases the gains that these individuals accrue from federal redistribution increase as well. Eventually, benefits outweigh costs, defining the threshold value \underline{x} above one.

Combining Propositions 1 and 2 follows that $f_l^B \geq f_h^B \geq f_h^A \geq f_l^A$ for $x < 1$ and $f_l^B \leq f_h^B \leq f_h^A \leq f_l^A$ for $x > 1$.¹¹ Thus, the federal tax rate proposed by a high productivity individual always obtains a majority over a tax rate proposed by a low productivity individual. That is,

Corollary 1. *The decisive voter over the federal income tax schedule is a high productivity individual. Consequently, the equilibrium federal income tax rate is given by $f_h^i(x)$.*

¹¹For $x < 1$ we obtained in Propositions 1 and 2 that $f_l^A = f_h^A = 0$. When $x > 1$ it is simple to show that

$$\frac{2\beta(x-1)}{x+\beta(x-1)} > \frac{2\beta(x-\underline{x})}{x+\beta(x-\underline{x})} \leftrightarrow x > 0,$$

which is always the case.

According to this corollary the decisive voter's productivity level is above the median productivity.¹² This result stands in sharp contrast to the one obtained in similar models with only one tier of income taxation (Romer, 1975; Roberts, 1977, Meltzer and Richard, 1981). In those models, monotonicity between the individuals' productivity level and the rate of their preferred tax schedule is obtained. Thus, under universal suffrage, the decisive voter is the individual with the median productivity level. This is not the case in a federal system of income taxation. Rather, in a federal system the decisive voter is a relatively productive individual who, for a wide range of parameter values, chooses to implement a positive federal income tax rate. More strikingly, an interesting situation may arise where the decisive voter's productivity is above the federation's mean productivity, yet this voter selects a positive federal income tax rate. The following proposition formalizes this observation for individuals in state A .

Proposition 3. *If $\underline{x} < x < (2 - n_A^h)/n_A^h$ then the decisive voter's productivity is above the federation's mean productivity level, yet the equilibrium federal tax rate is positive.*

According to Proposition 3, for a certain range of x , as income inequality within a state increases, redistribution in that state decreases. As an illustration, consider the predictions of the model when w_A decreases while w_B increases. Those changes increase the productivity dispersion between the states. Hence, if the implemented federal income tax schedule is f_h^A the resulting federal tax rate will increase as well. Consequently, both states' income tax rates will decrease, even though income inequality in state A increases while in state B decreases; that is, either a positive or negative relation between the productivity of the decisive voter and the state's level of redistribution may arise in a federal system of income taxation.¹³ This is perhaps the reason why several studies using data from the states concluded that there is no empirical support to the claim that a positive relation between income inequality and government redistribution exists - the main hypothesis of Meltzer and Richard (1981).¹⁴

The next section develops an efficiency analysis of a federal social contract with a two-tier income taxation system.

4 Efficiency Analysis

This section shows that the current federal social contract consisting of a two-tier income taxation system leads to an equilibrium on the downward sloping side of the federation's Laffer curve. As already mentioned in the introduction, when choosing the federal income tax rate an individual in state i ignores the

¹²Persson and Tabellini (1996) obtain similar results in a different framework. In a model where the federal policy achieves two main goals (risk sharing and redistribution) they find that transfers between the states exacerbates interstate conflict, in the sense that residents of the rich state tend to prefer lower federal tax rates than residents of the poor state.

¹³While r_A will certainly decrease, whether r_B decreases or increases depends on the relative change in w_B .

¹⁴See for example Gouveia and Masia (1998), and Rodriguez (1999).

impact of f on the redistribution level of the other state. Yet, the federal tax rate has a significant impact on the income tax schedule implemented in the other state. Therefore, the federal tax rate creates an externality that is ignored under a decentralized system of decision-making. A policy that takes this externality into account might achieve a welfare improvement for all the citizens of the federation, with greater redistribution and lower taxation.

This nonoptimal redistributive policy is reminiscent of Gordon (1983), Johnson (1988), Wildasin (1991) and Boadway et al. (1998). These papers develop a normative analysis of taxation in a federation of states. In their analysis, a benevolent social planner (maximizing a Benthamite welfare function over the utilities of current residents of a state) fails to take into account either vertical or horizontal externalities. (A vertical externality relates to the effects of the states' policies on federal revenues. A horizontal externality is caused by the mobility of individuals between states and the impact of the states' taxes on nonresidents of a particular state.) As a consequence of these two externalities inefficiency arises.

Several differences between the current paper's approach and the one adopted in the previous literature are worth mentioning. First, the previous papers abstract from political economy considerations, the main focus of the present paper. Second, in this paper the horizontal externality is assumed away since individuals are immobile. Finally, given that individuals (and not a social planner) choose the federal tax rate through majority vote, the vertical inefficiency has an opposed effect to the one obtained in previous papers. Both in Johnson (1988) and in Boadway et al. (1998) states ignoring the effects of their taxes on federal revenues tend to implement a higher than optimal income tax rate. In contrast, in the current paper states' tax rates are low while the federal tax rate tends to be higher than optimal.

To formalize matters, let us consider how the federation's total redistribution level reacts to changes in the productivity dispersion between the states. The federation's total redistribution is given by

$$R(x) \equiv r_f + r_A + r_B = \frac{\mathcal{P}}{2} \left[\frac{(1-s)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} [f + s(1-f)] \quad (7)$$

where

$$\mathcal{P} \equiv n_A^h w_A^{\frac{\beta+1}{\beta}} + n_B^h w_B^{\frac{\beta+1}{\beta}}$$

is the overall productivity level of the federation, which is constant. If the equilibrium federal tax rate is f_h^A , then R is constant for $x < \underline{x}$ and $x > \bar{x}$. For intermediate values of x , however, total redistribution is strictly decreasing in x .

Lemma 2. *$R(x)$ is strictly decreasing in x for $\underline{x} < x < \bar{x}$.*

In this same range of productivity dispersion, total income taxation on the individuals' income increases with x . That is, for any $\underline{x} < x$ the federation as a whole ends up at the decreasing part of its Laffer curve. This is a nonoptimal outcome as a reduction in income taxes would increase total redistribution.

Proposition 4. *If $x > \underline{x}$ the equilibrium tax rates of the federation are on the downward sloping side of its Laffer curve.*

The intuition is straightforward. When the federal income tax rate is positive the loss in redistribution in the more productive state (which is not taken into account by the decisive voter) more than offsets the gain in redistribution in the less productive state. Proposition 4 does not state, however, that a decrease in the rate of a given tax will result in a Pareto improvement. Rather, the proposition establishes that in equilibrium the combination of taxes selected by the individuals delivers an outcome inside the federation's Pareto frontier.

An important reason behind this inefficiency is that the federal tax schedule is the only available policy instrument that redistributes income between the states. Clearly, a social planner implementing lump-sum transfers between the states and eliminating one layer of taxation would achieve a Pareto improvement. But even abstracting from lump-sum transfers, a Pareto improvement can be achieved for sufficiently high levels of productivity dispersion by eliminating states' income tax schedules.

Proposition 5. *There exists a critical productivity dispersion level $x^* < \bar{x}$ such that for every $x > x^*$ eliminating states' income tax schedules results in a greater utility level for all the individuals in the federation.*

The individuals' impossibility to credibly commit to a non-equilibrium strategy drives the federation to this inefficient outcome. Suppose low-productivity individuals promise to choose a state income tax rate of zero for any federal income tax rate. In this situation the resulting federal tax rate equals $\beta/(\beta + 1)$, the one preferred by low-productivity individuals in both states. Notwithstanding their promise, in the second stage of the political process low-productivity individuals in both states will select a positive state income tax rate as it allows them to enforce more redistribution. High-productivity individuals anticipate such a deviation in the first period and behave accordingly. Hence, the resulting inefficient equilibrium is inescapable without a commitment mechanism.

But even if a device that outlaws income tax schedules exists only in state A , an efficient outcome will not be reached.¹⁵ If a device that outlaws income tax schedules exists only in state A , low-productivity individuals in state B are enjoying the best of both worlds. First, they receive more federal redistribution given that high-productivity individuals in state A have higher incomes due to incentive effects. And second, they also receive redistribution at the state level which there is no reason to give up by setting the state's tax rate at zero. Thus, for low-productivity individuals the implementation of a state income tax schedule is a dominant strategy, that leads the federation to an inefficient equilibrium.

Another policy instrument that may achieve an efficiency improvement is the implementation of a system of federal matching grants. The next subsection analyzes this policy instrument.

¹⁵In some states, like Tennessee, the constitution does not allow the implementation of an income tax schedule. In others states a supermajority is required for tax rates increments (this is the case in more than 15 states).

4.1 Federal Matching Grants

In models of fiscal competition the federal government might design corrective schemes that undo the nonoptimal outcomes that arise from decentralized state decision-making. Wildasin (1991), for example, shows that a system of matching grants from a federal government to state governments can neutralize the horizontal externalities created by states' policies, leading the federation to an efficient outcome.¹⁶ It is then natural to explore the implications of such a policy using the current framework whereby all the taxes are determined according to majority rule equilibrium, rather than by social planners.

Under a federal matching grants program state i 's budget constraint is given by

$$\delta r_i = s_i(1-f)p_i n_i^h y_i, \quad (8)$$

where $\delta \in (0, 1)$ measures the state's share of the cost of a dollar's worth of redistribution. The balanced budget constraint condition at the federal level implies that

$$r_f + (1-\delta)(r_A + r_B) = f [p_A n_A^h y_A + p_B n_B^h y_B]. \quad (9)$$

Given that under a matching grants program states are responsible for only a share of their redistribution expenses, individuals choose more state redistribution. A higher federal tax rate is required to pay for part of this greater redistribution level. As a result we obtain higher income tax rates and lower total redistribution. Thus, a system of federal matching grants in general enhance the resulting inefficiencies.¹⁷

Proposition 6. *Under a federal matching grants program, if*

$$x \leq \underline{x} \quad \text{or} \quad x \geq 1 + \frac{2(\delta + n_A^h)}{n_A^h [(1+\delta)(\beta+1) - 2]},$$

then the overall implemented income tax rates are greater and the federations' total redistribution level is smaller than without the matching grants program.

That is, a federal matching grants program under a democratic system might have the opposite effect than the one obtained when the federal government is represented by a benevolent social planner. This provides a critical assessment to the use of federal matching grants as a corrective device that undo externalities arising from decentralized decision making.

The next section provides an empirical analysis of the main implications of the model.

¹⁶A similar suggestion without a formal analysis appears in Gordon (1983).

¹⁷The result in the following proposition cannot be derived for any productivity dispersion level without imposing some (sufficient) conditions on the parameters. For example, notice that the range where the result is not guaranteed is empty for $\beta \geq (1+2n_A^h+\delta)/(1-2n_A^h-\delta)$, with $1-2n_A^h-\delta > 0$.

Another possibility would be to impose restrictions only on δ . In fact, there exists one critical value δ^* such that the stated inefficiency is obtained for all $\delta < \delta^*$, for any x .

5 Empirical Analysis

The main goal of this empirical exercise is to estimate the predictions of the model at hand. The equations I wish to estimate, linearizations of the model-based reduced forms relating states income tax rates, federal income tax rate, productivity dispersion between the states and states' populations are

$$\ln f_t = \theta_0 + \theta_1 \ln x_t + \theta_2 \ln p_t + \varepsilon_t, \quad (10)$$

and

$$s_{i,t} = \gamma_0 + \gamma_1 \ln p_{t,i} + \gamma_2 f_{t-1} + d_i + \epsilon_{i,t}, \quad (11)$$

where i indexes countries and t indexes time, d_i is a set of state fixed effects, the θ 's and γ 's are parameters to be estimated, and $\epsilon_{i,t}$ and ε_t are random error terms.

There are two empirical predictions that follow from the theoretical analysis. The first refers to the equilibrium level of the federal tax rate. In particular, we expect θ_1 to be positive. The second is related to the existent interdependency between federal and state tax rates. This prediction focuses on the slope of the reaction function and not on equilibrium tax levels. Accordingly, we expect γ_2 to be negative.¹⁸ Combining the two results we can draw conclusions about the impact of productivity dispersion on states income taxes.

Several econometric issues arise in the estimation of equations (10) and (11). First, states within the U.S. are likely to vary considerably in their preferences over income taxes.¹⁹ Thus, the unobserved d_i , reflecting persistent differences across states in taxation preferences, is likely to be correlated with the regressors. To remedy this, state fixed effects are included in (11).²⁰ Second, since state income tax rates are zero for some states for some years, least squares estimates of equation (11) are biased. For this reason the equation is not log-linearized and Tobit estimation is performed. Finally, the Newey and West's robust, consistent estimator for autocorrelated disturbances with an unspecified structure is used to estimate equation (10). This estimator solves the problem of serial correlation of disturbances across periods.

5.1 Data

The main focus of the empirical exercise is on the variables determining states and federal income tax rates as well as their interaction. This paper uses annual

¹⁸The use of the lagged value of the federal tax rate is due to the assumption that the federal government acts as a Stackelberg leader.

¹⁹Tennessee, for example, one of only nine states that do not tax personal income, is known to be a state with a long history of aversion to income taxes. This aversion dates back to 1931, when Tennessee Supreme Court ruled that income taxes were unconstitutional. The aversion to income taxes is still maintained in the year 2002 when, in dealing with a growing budget deficit, the legislature decided to increase sales taxes rather than to enact an income tax.

²⁰Note that the federal tax rate captures the effects of national changes in policy that may jointly affect trends in the variables. Therefore, its inclusion in the estimated equation precludes me from adding time dummies.

data on the states for the years 1977 to 2000 inclusive. I begin by discussing the sources and constructions of the different variables used in the empirical estimation. All the variables, with their definitions, means, and standard deviations, are reported in Appendix B.

States and federal income taxes are taken from the TAXSIM model. In particular, I use the average effective marginal state income tax rate on wages over the years 1977 to 2000.²¹ To find these rates, TAXSIM uses the same nationally representative sample from 1995, properly deflated, for each state and year. This approach allows for comparisons of law without confusing changes in income and deductions with changes in law.²²

The data show that income taxes vary considerably across states. For every year in the studied sample there are at least 8 states with zero income tax rates.²³ On the contrary, several states exhibit over the years income taxes above 5%. Hawaii and Oregon have the highest average rate (8.35% and 8.25% respectively).

The average effective marginal federal income tax is obtained from the same source. This tax presents some variation over the years as well. It ranges from 22.92% in 1991 and 1992 to 32.53% in 1981.

I use data on average hourly earnings of production workers on manufacturing payrolls and the civil labor force of each state to build a proxy for the productivity dispersion between the states. In order to obtain a number for the measure of productivity dispersion, a fix value is assigned to the individuals' constant elasticity of labor supply. Since Stern's (1976) contribution, it is well known in the literature of income taxation that the results are influenced by the particular value assigned to this parameter. For that reason I carried on several estimations varying the value of the elasticity of labor supply. The next section reports the results for β equal to 1/2, 1 and 2. Once the value of productivity dispersion for each state is obtained, in accordance with the theoretical model, the estimation uses the median state's value of productivity dispersion.

5.2 Results

Table 1 reports the basic findings. From the estimation of equation (10) the general pattern that emerges is that an increase in productivity dispersion between the states leads to an increase in the federal income tax rate. The estimation indicates that a 1% point rise in productivity dispersion leads to a statistically significant increase of around 2.5% point in the federal income tax rate. Moreover, from a comparison of columns (1), (2) and (3) follows that the impact of

²¹Sauter (1985) points to several advantages of using this particular definition instead of the ratio of income tax revenue to personal income which is used in related empirical studies.

²²For further details about the TAXSIM model and the data used in this paper see Feenberg and Coutts (1993).

²³The states of Florida, Nevada, New Hampshire, South Dakota, Tennessee, Texas, Washington and Wyoming exhibit a zero income tax rate for every year in the studied sample. Alaska imposes a zero income tax rate starting from 1979 and Connecticut has a zero income tax rate until 1990.

productivity dispersion on the federal income tax decreases as the elasticity of the individuals' labor supply increases.

The estimation of equation (11) also confirms the hypothesis advanced. The results show that there is a highly statistically significant negative interdependence between states income tax rates and the federal income tax rate; the coefficient of -0.067 implies an elasticity of -0.4 computed at the means.

Table 1: Parameter Estimates

Variable	$\ln f$			s
	$1/\beta = 0.5$	$1/\beta = 1$	$1/\beta = 2$	
	(1)	(2)	(3)	(4)
Constant	23.521 (3.863)	21.672 (3.251)	16.212 (2.102)	5.579 (3.342)
$\ln x$	2.962 (0.553)	2.542 (0.381)	2.260 (0.373)	
$\ln p_t$	-1.626 (0.312)	-1.476 (0.262)	-1.039 (0.169)	
f_{t-1}				-0.067 (0.009)
$\ln p_{t,i}$				-0.035 (0.385)
σ				0.735 (0.017)
Log Likelihood Function				-1058.63
N	24	24	24	1150

Number in parenthesis are robust Standard Errors. Variables are defined in Appendix B. Equation (4) is estimated with state effects.

This result contributes to the ongoing debate in the related empirical literature of strategic interaction among governments. Inspired by the theoretical analysis of Boadway and Keen (1996), several studies attempt to estimate tax reactions functions. Besley and Rosen (1998) along with Esteller-Moré and Solé-Ollé (2001) study interaction in the U.S., focusing on commodity taxes and income taxes respectively; Goodspeed (2000, 2002) conducts similar analyses focusing on income taxes in Europe and Hayashi and Boadway (2001) focus on Canadian corporate income taxes. While Besley and Rosen (1998) as well as Esteller-Moré and Solé-Ollé (2001) find upward-sloping state reactions functions, the estimations of Goodspeed (2000, 2002) and Hayashi and Boadway (2001) point to a downward-sloping reaction function for lower-level governments.²⁴

Combining the two estimations we conclude that productivity dispersion between the states significantly decreases states tax income rates. From the results above follow that the elasticity of states income tax rates with respect to productivity dispersion evaluated at the means oscillates between -1.179, -1.012 and -0.9 for $1/\beta$ equal to $1/2$, 1 and 2 respectively.

²⁴ Among the studies using data on personal income tax, the different results may be in part attributable to the fact that Esteller-Moré and Solé-Ollé (2001) is the only one that includes horizontal effects.

Finally, the results point to a negative relation between the population's size and the income tax rates. While highly significant at the federal level, this relation is not statistically significant at the state level.

6 Conclusion

This paper developed a positive theory of overlapping taxation in a federation of states. The main result shows that the existence of productivity dispersion between the states plays a crucial role in the analysis. Individuals residing in different states have opposing preferences over the federal tax rate. As a consequence, the decisive voter at the federal level has a productivity level above the median. Yet, this high-productivity individual's preferred federal tax rate is an increasing function of the productivity dispersion between the states; so if this relatively productive individual is from a poor state, she will support a positive rate of federal income tax. A high rate of federal income tax causes, because of incentive considerations, a low rate of taxation at the state level, ultimately bringing low state income taxes.

Another result worth mentioning is that a federal social contract allowing a two-tier income taxation system is, in general, not optimal. The reason for this is that under a decentralized system of decision-making the federal tax schedule creates externalities that are partially ignored. Moreover, it is also found that a system of federal matching grants brings higher tax rates. Higher tax rates further decrease total redistribution, and thus the total welfare of the federation's residents. This contrasts with previous results obtained in the related normative literature where a system of federal matching grants is welfare improving.

Although it delivers new and interesting results, the model is highly stylized. Several caveats are worth nothing.

First, the model abstracts from mobility of individuals between the states. Including this feature to the model adds another layer to the individual's problem. Relatively higher tax rates in one state may lead to the emigration of productive individuals to the other state. Yet, the concentration of rich individuals in one state may lead to the immigration of poor individuals. Given that poor individuals comprise a majority of the population, their immigration will result in higher tax rates. An equilibrium in such a framework is a fixed-point in which no individual wishes to move or alter its labor supply, and no state wishes to change its tax rate given the tax rate chosen by the other state. To guarantee the existence of such an equilibrium is a challenging task. Using this framework it is very difficult to come up with a set of simple sufficient conditions for existence. A change in policy implies migratory movements that imply a change in the composition of the population and, subsequently, another change in policy. This cycle may continue endlessly.

In any event, I presume that the inclusion of mobility considerations may help us understand the coexistence of high federal income tax rates with low state income tax rates. Simply put, the federal government has a monopoly on

the power and ability (however imperfect) to coerce citizens into paying taxes. In the stylized democracy model of this paper, the federal government is nothing more than the aggregation of the preferences of a majority of individuals. Given that at the federal level the poor population will always constitute a majority, federal tax rates will tend to be high. While federal income taxes are inescapable to the rich population, such is not the case with states income tax schedules. Tax competition among the states will emerge and drive states income tax rates to low levels.

Another important extension is to introduce a general distribution of productivity levels for each state. I believe that the obtained monotonicity of the preferences of the individuals of a given state over the federal tax rate would be preserved. Yet, it is not simple to prove it since the characterization of the states income tax rate as a function of the federal tax rate is cumbersome. Moreover, this characterization is carried over when solving for the individuals' preferences over the federal tax rate.

Similar difficulties arise when trying to generalize the individuals' preferences to any utility function exhibiting "nice behavior." For general utility functions the individuals' labor supply is a function of the state and federal redistribution level. Boadway and Keen (1996) show, in a different framework, that the slope of states' reactions functions to the federal tax is ambiguous. Whether this is the case as well in a positive analysis without transfers between governments is still an open question.

Finally, the analysis is restricted to linear taxes for technical reasons. In reality, however, taxes at both the state and federal level exhibit increasing marginal rates. Unfortunately, allowing for a richer set of tax schedules might lead to voting cycles. Hopefully, future research will help us understand the dynamics of overlapping income taxation in a federal system imposing fewer restrictions on the set of feasible taxes.

A Appendix A: Proofs

Proof of Lemma 1. Low-productivity individuals' decision problem takes the form:

$$\max_{s_i \in [0,1]} s_i(1-f)n_i^h w_i \left[\frac{(1-s_i)(1-f)w_i}{\alpha} \right]^{\frac{1}{\beta}} + f(1-f)^{\frac{1}{\beta}} \sum_{i=A,B} p_i n_i^h w_i \left[\frac{(1-s_i)w_i}{\alpha} \right]^{\frac{1}{\beta}}.$$

Strict concavity in s_i is easily verified on the relevant domain, and the first-order conditions directly yield the stated results. ■

Proofs of Propositions 1 and 2. The proof below solves the maximization problem for an individual living in state A . A similar procedure yields the results for a resident of state B .

The individuals' maximization problem takes the form:

$$\begin{aligned} & \max_{s_i, f \in [0,1]} s_A(1-f)n_h^A w_A \left[\frac{(1-s_A)(1-f)w_A}{\alpha} \right]^{\frac{1}{\beta}} + \frac{f}{2}(1-f)^{\frac{1}{\beta}} \sum_{i=A,B} n_i^h w_i \left[\frac{(1-s_i)w_i}{\alpha} \right]^{\frac{1}{\beta}} \\ & + \gamma \frac{\beta}{\beta+1} \left(\frac{1}{\alpha} \right)^{\frac{1}{\beta}} [(1-s_A)(1-f)w_A]^{\frac{\beta+1}{\beta}} \\ & \text{subject to } s_i = \begin{cases} \frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)}, & \text{if } f < \frac{2\beta}{2\beta+1}, \\ 0, & \text{otherwise.} \end{cases}, \quad f \in [0,1], \end{aligned}$$

where $\gamma = 0$ for the preferences of low-productivity individuals (Proposition 1) and $\gamma = 1$ for the preferences of high-productivity individuals (Proposition 2).

To simplify the exposition I denote the total weighed productivity in each state by

$$a \equiv n_A^h w_A^{\frac{\beta+1}{\beta}} \quad \text{and} \quad b \equiv n_B^h w_B^{\frac{\beta+1}{\beta}}.$$

The Lagrangian for the maximization problem above is:

$$\begin{aligned} \mathcal{L}(f, s_A, s_B, \lambda_A, \lambda_B, \theta_A, \theta_B, \theta_f; x) &= s_A(1-f)a \left[\frac{(1-s_A)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} + \\ & \frac{f}{2}(1-f)^{\frac{1}{\beta}} \sum_{i=A,B} n_i^h w_i \left[\frac{(1-s_i)w_i}{\alpha} \right]^{\frac{1}{\beta}} + \gamma \frac{a\beta}{n_A^h(\beta+1)} \left(\frac{1}{\alpha} \right)^{\frac{1}{\beta}} [(1-s_A)(1-f)]^{\frac{\beta+1}{\beta}} \\ & + \sum_{i=A,B} \left\{ \theta_i s_i - \lambda_i \left[\frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)} - s_i \right] \right\} + \theta_f f. \end{aligned}$$

The corresponding first-order conditions are

$$\begin{aligned} 1. \quad \mathcal{L}_f &\Rightarrow -s_A \frac{\beta+1}{\beta} a \left[\frac{(1-s_A)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} + \left(\frac{1-f}{\alpha} \right)^{\frac{1}{\beta}} \frac{\beta(1-f)-f}{2\beta(1-f)} \left[a(1-s_A)^{\frac{1}{\beta}} + b(1-s_B)^{\frac{1}{\beta}} \right] \\ & - \gamma \frac{a(1-s_A)}{n_A^h} \left[\frac{(1-s_A)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} + \sum_{i=A,B} \frac{\lambda_i}{2(\beta+1)(1-f)^2} + \theta_f = 0. \\ 2. \quad \mathcal{L}_{s_A} &\Rightarrow \frac{(1-f)a}{\beta(1-s_A)} [\beta(1-s_A) - s_A] \left[\frac{(1-f)(1-s_A)}{\alpha} \right]^{\frac{1}{\beta}} - \\ & \frac{af}{2\beta(1-s_A)} \left[\frac{(1-f)(1-s_A)}{\alpha} \right]^{\frac{1}{\beta}} - \gamma \frac{a(1-f)}{n_A^h} \left[\frac{(1-s_A)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} \end{aligned}$$

$$+\lambda_A + \theta_A = 0$$

$$3. \mathcal{L}_{s_B} \Rightarrow -\frac{bf}{2\beta(1-s_B)} \left[\frac{(1-f)(1-s_B)}{\alpha} \right]^{\frac{1}{\beta}} + \lambda_B + \theta_B = 0$$

$$4. \mathcal{L}_{\lambda_i} \Rightarrow s_i \geq \frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)}, \lambda_i \geq 0, \text{ with complementary slackness.}$$

$$5. \mathcal{L}_{\theta_i} \Rightarrow s_i \geq 0, \theta_i \geq 0, \text{ with complementary slackness.}$$

$$6. \mathcal{L}_{\theta_f} \Rightarrow f \geq 0, \theta_f \geq 0, \text{ with complementary slackness.}$$

With two inequality constraints and three non-negative variables, there are 32 possible patterns of equations and inequalities. Let us see which ones offer candidates for a maximum.

The first candidate is given by $\theta_f \geq 0$, $\theta_i = 0$, and $\lambda_i \geq 0$. These constraints imply that $f = 0$ and $s_i = \beta/(\beta + 1)$. Substituting these values into the FOCs yields that $\theta_f \geq 0$ if and only if $x \leq 1 + \frac{\gamma}{n_A^h(\beta+1)}$. The inequalities $\lambda_i \geq 0$ are satisfied for any value of x .

The second candidate is given by $\lambda_i \geq 0$, and $\theta_i = \theta_f = 0$. This case correspond to the interior solution where all the tax rates are positive. From the first three FOCs we obtain the following two equations:

$$\lambda_A = \left[\frac{(1-s_A)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} \left[\frac{af}{2\beta(1-s_A)} + \gamma \frac{a(1-f)}{n_A^h} - \frac{(1-f)a}{\beta(1-s_A)} [\beta(1-s_A) - s_A] \right] > 0, \quad (12)$$

$$\lambda_B = \left[\frac{(1-s_B)(1-f)}{\alpha} \right]^{\frac{1}{\beta}} \frac{fb}{2\beta(1-s_B)} > 0 \quad (13)$$

and replacing λ_i and s_i in \mathcal{L}_f according to (12), (13), and Lemma 1 yields

$$\begin{aligned} & \frac{1+x}{2\beta(1-f)} [\beta(1-f) - f] - \left[1 - \frac{f}{2\beta(1-f)} \right] - \frac{\gamma(2-f)}{2(\beta+1)(1-f)n_A^h} \\ & + \frac{1}{2(\beta+1)(1-f)^2} \left[\frac{f(1+x)(1-f)(\beta+1)}{\beta(2-f)} - \frac{f(1-f)(\beta+1)}{\beta(2-f)} + \frac{\gamma(1-f)}{n_A^h} \right]. \end{aligned}$$

Tedious manipulations of the equation above deliver

$$f_i^A = \frac{2\beta(x-1)}{x+\beta(x-1)} \text{ and } f_h^A = \frac{2\beta(x-\underline{x})}{x+\beta(x-\underline{x})},$$

the results stated in Proposition 1 and Proposition 2.

Finally, the solution for high levels of productivity dispersion is reached when $\theta_f = 0$, $\lambda_i \geq 0$ and $\theta_i \geq 0$, $i = A, B$. In this case $s_A = s_B = 0$ and $f = 2\beta/(2\beta+1)$. Substituting the values of the tax rates into the FOCs we obtain that $\lambda_i > 0$ for any parameters values but $\theta_i \geq 0$ if and only if $x \geq \frac{\beta}{\beta+1} + \frac{\gamma}{\beta n_A^h}$.

Second order sufficient conditions are satisfied for all the different solutions. I choose to omit these conditions here because they are lengthy and do not

provide additional insights. The details can be obtained from the author upon request. ■

Proof of Proposition 3. The condition $x < \frac{2-n_A^h}{n_A^h}$ implies that

$$n_A^h w_A^{\frac{1+\beta}{\beta}} + n_B^h w_B^{\frac{1+\beta}{\beta}} < 2w_A^{\frac{1+\beta}{\beta}}$$

which directly yields $y_A > \frac{1}{2}(n_A^h y_A + n_B^h y_B)$. Finally, by Proposition 2 we know that f_h^A is positive for $1 + \frac{1}{(\beta+1)n_A^h} < x$. ■

Proof of Lemma 2. If $x < \bar{x}$ by Proposition 2 we know that $f < 2\beta/(2\beta+1)$, which implies (by Lemma 1) that

$$s = \frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)}. \quad (14)$$

Substituting (14) into (7) yields

$$R(x) = \psi(2-f)^{\frac{1}{\beta}}(f+2\beta)$$

where

$$\psi \equiv \frac{a+b}{4(\beta+1)} \left[\frac{1}{2\alpha(\beta+1)} \right]^{\frac{1}{\beta}}$$

is a constant. Differentiating R with respect to x we obtain

$$\frac{\partial R}{\partial x} = \psi(2-f)^{\frac{1}{\beta}} \left[1 - \frac{f+2\beta}{\beta(2-f)} \right] \frac{\partial f}{\partial x},$$

which is always negative for $\beta > 0$. ■

Proof of Proposition 4. From Lemma 2 we know that total redistribution is strictly decreasing in this range. It remains to show that total taxation on the individuals' income is increasing in x .

From Lemma 1 and Proposition 2 follows that

$$s + f = \frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)} + f.$$

Differentiating the previous expression with respect to x we obtain that

$$\frac{d(s+f)}{dx} > 0 \text{ if and only if } f < 1 - \frac{1}{4(\beta+1)^2}.$$

The above inequality is always satisfied since the maximum possible value of f in equilibrium, $\frac{2\beta}{2\beta+1}$, is strictly less than $1 - \frac{1}{4(\beta+1)^2}$. ■

Proof of Proposition 5. When states taxes are outlawed, low-productivity individuals in both states have the same preferences over the federal tax schedule.

Hence, they would form a majority at the federal level. Their preferred federal tax rate in this case is

$$\widehat{f} = \arg \max_f r_f \quad \text{s.t. } s_i = 0; \quad i = A, B.$$

When $x \geq \bar{x}$ the equilibrium income tax schedules are $s_i = 0$ and $f_h^A = 2\beta/(2\beta + 1)$. Thus, in this range state redistribution is zero. Since \widehat{f} is uniquely defined and $\widehat{f} < f_h^A$ we obtain that

$$r_f(\widehat{f}) > r_f(f_h^A).$$

That is, for $x \geq \bar{x}$, according to the outcome under \widehat{f} taxes are lower and redistribution is greater than according to the equilibrium tax rate f_h^A . Thus, for $x \geq \bar{x}$, the utility of all the individuals in the federation is greater under \widehat{f} than under f_h^A . In fact, this is the case for every $x > x^*$, where x^* is defined by

$$r_f[f(x^*), s(x^*)] + r_A[f(x^*), s(x^*)] = r_f[f^*, 0]^{25}$$

Proof of Proposition 6. Under a federal matching grants program, the indirect utility level of low-productivity individuals in state A is

$$\begin{aligned} V_l^A &= s_A(1-f)n_A^h w_A \left[\frac{(1-s_A)(1-f)w_A}{\alpha} \right]^{\frac{1}{\beta}} \left[\frac{1-(1-\delta)p_A}{\delta} \right] + \\ &f \left(\frac{1-f}{\alpha} \right)^{\frac{1}{\beta}} \sum_{i=A,B} p_i n_i^h w_i [(1-s_i)w_i]^{\frac{1}{\beta}} - \frac{(1-\delta)}{\delta} r_B \end{aligned} \quad (15)$$

obtained by substituting equations (3), (8), and (9) back into (2). The implemented state's tax schedule in this state is obtained by maximizing (15) over the set of feasible state taxes. The solution to that maximization problem yields²⁶

$$\widehat{s}_A(\delta) = \begin{cases} \frac{\beta}{\beta+1} - \frac{\delta f p_A}{(1-f)(\beta+1)[1-(1-\delta)p_A]}, & \text{if } f < \frac{\beta[1-(1-\delta)p_A]}{\beta[1-(1-\delta)p_A] + \delta p_A}, \\ 0, & \text{otherwise.} \end{cases}$$

Substituting $p_i = 1/2$ and solving for the preferred federal tax rate as in Propositions 1 and 2, we obtain

$$f_l^A(\delta) = \begin{cases} 0 & \text{if } x \leq 1, \\ \frac{\beta(x-1)(1+\delta)^2}{2x\delta + \beta(x-1)(1+\delta)} & \text{if } 1 < x \leq \frac{(1+\delta)(\beta+1)}{\beta(1+\delta) - (1-\delta)}, \\ \frac{\beta(1+\delta)}{\beta(1+\delta) + \delta} & \text{if } \frac{(1+\delta)(\beta+1)}{\beta(1+\delta) - (1-\delta)} < x, \end{cases}$$

²⁵Note that the level of redistribution at either the state or federal level is continuous in x . Furthermore, total redistribution is strictly decreasing in x in the relevant range. Hence, such an x^* exists and is uniquely defined.

²⁶The indirect utility function V_l^A is strictly concave in s_A on the relevant domain. Hence, this is the unique solution, obtained directly from the first order conditions.

for low-productivity individuals, and

$$f_h^A(\delta) = \begin{cases} 0 & \text{if } x \leq 1 + \frac{2\delta}{n_A^h(\beta+1)(1+\delta)}, \\ \frac{\beta(1+\delta)[(\beta+1)n_A^h(x-1)(1+\delta)-2\delta]}{[2x\delta+\beta(x-1)(1+\delta)](\beta+1)n_A^h-2\beta\delta} & \text{if } 1 + \frac{2\delta}{n_A^h(\beta+1)(1+\delta)} < x \leq 1 + \frac{2(\delta+n_A^h)}{n_A^h[(1+\delta)(\beta+1)-2]}, \\ \frac{\beta(1+\delta)}{\beta(1+\delta)+\delta} & \text{if } 1 + \frac{2(\delta+n_A^h)}{n_A^h[(1+\delta)(\beta+1)-2]} < x. \end{cases}$$

for high-productivity individuals.

As it is the case without federal matching funds, preferences are not monotonic in w . Therefore, the equilibrium federal tax rate is $f_h^A(\delta)$. Note that $f_h^A(\delta) \geq f_h^A(1)$ for

$$x \leq \bar{x} \text{ or } x \geq 1 + \frac{2(\delta + n_A^h)}{n_A^h[(1+\delta)(\beta+1)-2]},$$

establishing the desired result. ■

B Appendix B: Summary Statistics

	Mean	St. Dev.	Median	Max	Min
s	4.54	2.722	5.035	9.87	0
f	26.97	3.197	26.155	32.53	22.92
$x_{0.5}$	0.984	0.023	0.984	1.025	0.937
x_1	0.982	0.025	0.981	1.019	0.92
x_2	0.99	0.03	0.995	1.026	0.92

^a Sources: States and federal average effective marginal tax rates are taken from the TAXSIM model. The measure of productivity dispersion between the states for the different values of β is built using data on average hourly earnings of production workers on manufacturing payrolls and the civil labor force, both found in Employment and Earnings, published by the U.S. Department of Labor, Bureau of Statistics.

References

- [1] **Besley, Timothy and Rosen, Harvey S.** “Vertical Externalities in Tax Setting: Evidence from Gasoline and Cigarettes.” *Journal of Public Economics*, December 1998, 70(3), pp. 383-398.
- [2] **Boadway, Robin, Marchand, Maurice and Vigneault, Marianne.** “The Consequences of Overlapping Tax Bases for Redistribution and Public Spending in a Federation.” *Journal of Public Economics*, June 1998, 68(3), pp. 453-478.
- [3] **Boadway, Robin and Keen, Michael.** “Efficiency and the Optimal Direction of Federal-States Transfers.” *International Tax and Public Finance*, May 1996, 3(2), pp. 137-156.
- [4] **Bohn, Henning and Stuart, Charles.** “Voting over Non-Linear Taxes in a Stylized Representative Democracy.” Mimeo, University of California at Santa Barbara, 2001.
- [5] **De Donder, Philippe and Hindricks, Jean.** “The Politics of Progressive Income Taxation with Incentive Effects.” *Journal of Public Economics*, forthcoming.
- [6] **Diamond, Peter A.** “Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates.” *American Economic Review*, March 1998, 88(1), pp. 83-95.
- [7] **Feenberg, Daniel R. and Coutts, Elisabeth.** “An Introduction to the TAXSIM model.” *Journal of Policy Analysis and Management*, Winter 1993, 12(1), pp. 189-194.
- [8] **Goodspeed, Timothy J.** “Tax Structure in a Federation.” *Journal of Public Economics*, March 2000, 75(3), pp. 493-506.
- [9] **Goodspeed, Timothy J.** “Tax Competition and Tax Structure in Open Federal Economies: Evidence from OECD Countries with implications to the European Union.” *European Economic Review*, forthcoming.
- [10] **Gordon, Roger H.** “An Optimal Taxation Approach to Fiscal Federalism.” *Quarterly Journal of Economics*, November 1983, 98(4), pp. 567-586.
- [11] **Gouveia, Miguel and Masia, Neil A.** “Does the Median Voter Model Explain the Size of the Government?: Evidence from the States.” *Public Choice*, October 1998, 97(1-2), pp. 159-177.
- [12] **Hayashi, Masayoshi and Boadway, Robin.** “An Empirical Analysis of Intergovernmental Tax Interaction: the Case of Business Income Taxes in Canada.” *Canadian Journal of Economics*, May 2001, 34(2), pp. 481-503.
- [13] **Johnson, William R.** “Income Redistribution in a Federal System.” *American Economic Review*, June 1988, 78(3), pp. 570-573.

- [14] **Meltzer, Allan and Richard, Scott F.** “A Rational Theory of the Size of the Government.” *Journal of Political Economy*, October 1981, 89(5), pp. 914-927.
- [15] **Persson, Torsten and Tabellini, Guido.** “Federal Fiscal Constitutions: Risk Sharing and Redistribution.” *Journal of Political Economy*, October 1996, 104(5), pp. 979-1009.
- [16] **Roberts, Kevin W. S.** “Voting over Income Tax Schedules.” *Journal of Public Economics*, December 1977, 8(3), pp. 329-340.
- [17] **Rodriguez, Francisco C.** “Does Distributional Skewness lead to Redistribution? Evidence from the United States.” *Economics and Politics*, July 1999, 11(2), pp. 171-199.
- [18] **Romer, Thomas.** “Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax.” *Journal of Public Economics*, February 1975, 4(2), pp. 163-186.
- [19] **Sauter, John J.** “On the Construction of Marginal Federal Personal and Social Security Tax Rates in the US.” *Journal of Monetary Economics*, January 1985, 15(1), pp. 121-135.
- [20] **Stern, Nicholas H.** “On the Specification of Models of Optimum Income Taxation.” *Journal of Public Economics*, July 1976, 6(1-2), pp. 123-162.
- [21] **Wildasin, David E.** “Income Redistribution in a Common Labor Market.” *American Economic Review*, September 1991, 81(4), pp. 757-774.