Federal Progressivity and State Regressivity*

Esteban F. Klor†

Department of Economics, New York University

269 Mercer St. 7th Floor, New York, NY 10003

December 2001

Abstract

This paper develops a positive theory of taxation in a federation of states. Its main motivation comes from two observations about tax schedules in the US. First, state income tax rates are low compared to federal income tax rates, in most cases below 5%. And second, the overall effective tax schedule at the state level is regressive in every state. In a federal system, with overlapping income taxation, total productivity dispersion between the states determines the federal tax rate. In fact, there exists a positive relation between the productivity dispersion and the federal tax rate, even if the income of the decisive voter is above the mean income. When individuals' income is endogenous, because of incentives considerations, the higher the implemented federal tax rate is the lower the resulting state tax rate will be, even if a relatively poor individual chooses this tax. Empirical evidence obtained from a new data set on nonlinear tax schedules at the state level supports the hypothesis of the paper. Most notably, the data points to the existence of a significant trade-off between the progressivity of the states and federal tax schedules, explained through productivity dispersion between the states.

KEYWORDS: Fiscal Federalism, Political Economy, Income Taxation.

JEL classification: D72, H23, H77

---

*Useful comments were provided by Sophie Bade, Jean-Pierre Benoît, Ronny Razin, Edward Wolff, as well as seminar participants at the University of Pennsylvania. I am especially indebted to Alessandro Lizzeri, Antonio Merlo and Efe Ok. All views and mistakes in the paper remain my own.

†Http://home.nyu.edu/~efk201. Email address: efk201@nyu.edu

1. Introduction

The main motivation for this paper comes from two observations about the tax schedules implemented in the states within the US. The first one is that state income tax rates are low, in most cases below 5%, with nine states having a zero income tax rate. The second perhaps more surprising regularity is that when we take into account the overall tax system in each state (income taxes, property taxes and sales taxes), the effective tax schedule of almost every state is regressive; that is, the average tax rate decreases with pre-tax income.[1] These facts (which are in stark contrast with the federal income taxation practice of any OECD country) are perhaps a consequence of the present federal political system that allows for overlapping income taxation.

Table 1, reproduced from the Institute on Taxation and Economic Policy (1996), illustrates these observations with data from 1995. This table shows the US's average personal income tax (P.I.T.) and total effective average tax (T.E.T.) for the 50 states and the District of Columbia.

Table 1: Personal Income Tax and Total Effective Average Tax, US Averages, 1995.

	Lowest	Second	Middle	Fourth	Top 20%		
	20%	20%	20%	20%	Next 15%	Next 4%	Top 1%
P.I.T.	1.2%	2.2%	2.8%	3.1%	3.5%	3.9%	4.6%
T.E.T.	12.5%	10.5%	9.8%	9.5%	9.0%	8.4%	7.9%

Source: Institute on Taxation and Economic Policy (1996).

As it is seen from the table, the personal income tax rate increases mildly with income yet it remains at low levels even for incomes at the top of the distribution. The total effective tax schedule at the state level includes sales taxes and property taxes on top of income taxes. Since the marginal propensity to consume (either durable and no durable goods) decreases with income, these two proportional statutory taxes have a regressive *effective* incidence. Given that income tax rates are low, the regressive effect of sales and property taxes more than offsets the progressive effect of income taxes. Consequently, the effective average tax rate of

[1] This is the case in 47 of the 50 states. If we add to each state's tax system the federal deduction offset, regressive taxes obtain in every state.

the overall tax system at the state level is decreasing with income; i.e., the overall tax system is regressive for all the states within the US.[2]

This paper constructs a simple model of taxation and redistribution in a two-tier federal system. The federal system has a single central government and two state governments. The federation's political process consists of two stages. In the first stage individuals vote for a federal tax schedule that applies to all the residents of the federation. In the second stage the residents of each state decide which tax schedule to implement in their home state. All the elections are decided by 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 immobile, reside in two different states and face overlapping taxes on their income. This framework brings a new source of heterogeneity to the analysis. Accordingly, individuals differ not only on their productivity level, but also in their state of residence.[3] Consequently, the individuals' preferred tax schedules is not entirely determined by their productivity level. The intuition is simple. While the state tax redistributes income *within* each state, the federal tax redistributes income *between* the states. To see this note that the richer the individuals in one state, the higher their federal tax liability, and thus the higher this state's proportional contribution to federal tax revenues. Yet, each state's proportion of federal redistribution (which is distributed lump-sum) is given by the size of its population, independently of the state residents' income. Hence, the effect of the federal tax schedule is to shift income from the rich to the poor state.

The existence of productivity dispersion between the states plays a crucial role in the analysis. Residents of the rich state always oppose a positive federal tax rate.[4] In contrast, residents of the poor state favor a positive federal tax rate, its

[2] Washington State exhibits the greatest difference of total effective tax between the lowest 20% and top 1% of the population (13.2%). Delaware exhibits the lowest difference (-1.3%). When we take into account the federal deduction offset this difference is positive even in Delaware (1.4%).

[3] This paper abstract from mobility of individuals between states. This assumption allows me to highlight a different mechanism influencing federal and state income tax schedules. See the conclusions for a detailed discussion of the influence of mobility on the results of this paper.

[4] Note that this is the case even for a low-productivity individual in the rich state. The

size depending on the individual's productivity level. That is, individuals residing in different states do not have monotonic preferences over the federal tax schedule with respect to their income. Consequently, even though low-productivity individuals constitute a majority of the federal population, they never form a coalition. Actually, the productivity of the decisive voter at the federal level is always above the median federal productivity, and may even be above the mean productivity. This voter's preferred federal tax rate is an increasing function of the productivity dispersion between the states; so if this high-productivity individual is from a poor state, she will support a positive rate of federal income tax. The reason being that federal taxation accomplishes two goals in the eyes of a high-productivity individual from a poor state: first, it redistributes income between the states benefiting even the more productive residents of the poor state; and second, because of incentives considerations, the higher the implemented federal tax rate is, the lower the implemented state tax rate will be, even if the decisive voter at the state level has zero pre-tax income. 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 federal progressivity, ultimately bringing state regressivity.

From an efficiency standpoint, a federal social contract allowing a two-tier income taxation system is, in general, not optimal.[5] 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, when choosing the federal income tax rate, individuals in one state do not take into account the impact of the federal tax schedule on redistribution in the other state.

A policy that takes the externality created by the decentralized political process into account will bring a welfare improvement for all the residents of the federation, with greater total redistribution and lower taxation. This provides a possible role for the federal government: to implement policies that undo the non-optimality arising from decentralization. Gordon (1983) and Wildasin (1991) argue that one such policy is the implementation of federal matching grants. According to a system of federal matching grants the federal government shares a proportion of the cost of each state's redistribution.

This paper finds that such a system would tend to decrease rather than in-

motivation of this individual is to keep as much as possible of her state's income for redistribution within the state.

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

crease total welfare of the federation's residents. A federal matching grants system reduces the cost of state redistribution paid by the state's residents. 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 lower levels of redistribution at even higher levels of taxation.

The second part of the paper expands the set of feasible tax schedules (both at the state and federal level) to allow for nonlinear taxation. Under this specification marginal tax rates can vary with income. This enables us to consider average rate regressive as well as average rate progressive tax schedules, in order to address the second observation depicted in Table 1: the regressivity of the overall tax system of every state of the US.

Using this framework I derive similar theoretical implications to the ones obtained in the first section. I should caution the reader, though, about several shortcomings of the model. Under nonlinear taxation, I am able to solve for the equilibrium state tax schedule but not for the federal tax schedule. Moreover, in this static model, without residential choices, only one kind of tax is imposed on the individuals. Regressivity of the overall tax system of every state obtains, however, when we combine income, property, and sales tax schedules. Embedding the theoretical model in a dynamic framework where individuals also have savings and residential choices would be a closer approximation to reality. Unfortunately, such a generalization remains elusive at this preliminary stage of the analysis. These generalizations are left for future research.

The last section of the paper estimates a number of implications of the theoretical model at hand using a new data set constructed by the Institute on Taxation and Economic Policy. Most notably, this empirical exercise corroborates that a significant negative relationship exists between the progressivity of the states' tax schedules and the progressivity of the federal tax schedule. Furthermore, it is found that productivity dispersion between states is a major force explaining this significant relationship between the progressivity of the two levels of taxation. Finally, the dispersion of the overall income distribution within each state (as measured by its variance) rather than the income of the median voter affects the progressivity of each state's tax schedule. These findings are in concert with the implications of the present model.

2. Linear Income Tax Schedules

2.1. The Benchmark Model

Consider a federation of two states, A and B.[6] 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. Therefore, there are two sources of heterogeneity between individuals: their productivity level and their state of residence. The population of each state is divided into two classes; in each state there is a continuum of low-productivity individuals (with productivity equal to zero) with mass $n_i^l > 1/2$, and a continuum of high-productivity individuals with $w_i > 0$, with mass $n_i^h = 1 - n_i^l$, $i = A, B$.

Individuals choose the amount of labor they sell on a competitive market and receive a wage rate equal to their productivity. The production sector exhibits constant-returns-to-scale so that the wage rate is constant. 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 tier 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 process of the federation is such that the federal tax is imposed first on the individuals' pre-tax income; later on every state imposes its own tax schedule on the remaining of the individuals' pre-tax income. As an objection to the previous assumption one may argue that the current practice in the US 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 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.[7]

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

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

[6] The analysis is easily generalized to any number of states.

[7] The assumption that the federal government acts as a Stackelberg leader is common in the literature (see for example Boadway et al., 1998).

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 $r_i \in \mathsf{R}_+$ is obtained from

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

Thus, individuals' net income is [8]

$$c_i = (1-f)(1-s_i)y_i + r_f + \frac{r_i}{p_i}. \tag{2.1}$$

Given both tax schedules, an individual with productivity w_i chooses pre-tax income $y_i(w_i, s_i, f)$ that maximizes $u(c, \frac{y}{w})$ subject to (2.1). Throughout this section I assume the following quasi-linear preferences over consumption and labor supply:

$$u(c, \frac{y}{w}) = c - \frac{\alpha}{\beta+1}\left(\frac{y}{w}\right)^{\beta+1}, \quad c, y \geq 0, \tag{2.2}$$

where α is a positive constant and $(1/\beta) > 0$ 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, in the presence of distorting taxes, redistribution may induce productive individuals to refrain from working, causing an indeterminacy of the state's tax rate.[9] Finally, this specification (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.

[8] This 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 $c_i = (1-f-s_i)y_i + r_f + r_i/p_i$ (see Gouveia and Masia, 1998). As already pointed out, adopting this different specification will not change the nature of any result of this paper.

[9] If no individual in the state has a positive income, there exists a continuum of equilibrium tax rates for this state. This indeterminacy creates complications when solving for the equilibrium federal tax rate.

For these preferences, given the federal and state tax schedules, the optimal pre-tax income of individual with productivity $w_i > 0$ is[10]

$$y_i = w_i \left[\frac{(1-s_i)(1-f)w_i}{\alpha} \right]^{\frac{1}{\beta}}. \tag{2.3}$$

Note that for these 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.

2.2. Federal and States' Tax Schedules

The political process is such that in the first stage the entire federation votes over the federal tax schedule, and in the second stage states' tax schedules are elected. Only Condorcet winners are implemented in equilibrium; that is, taxes that obtain the support of at least half of the population against any other feasible tax. 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 as a function of the federal tax schedule, and then for the implemented federal tax schedule.

2.2.1. Preferences Over the States' Tax Schedules

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

$$r_f = f \left(\frac{1-f}{\alpha} \right)^{\frac{1}{\beta}} \sum_{i=A,B} p_i n_i^h w_i \left[(1-s_i) w_i \right]^{\frac{1}{\beta}}, \tag{2.4}$$

while state i's level of redistribution is given by

$$r_i = p_i s_i (1-f) n_i^h w_i \left[\frac{(1-s_i)(1-f)w_i}{\alpha} \right]^{\frac{1}{\beta}}. \tag{2.5}$$

Since individuals with zero productivity constitute a majority in each state, they choose by majority rule which tax schedule will be implemented in their

[10] For this and all subsequent maximization problems, first order necessary conditions and second order sufficient conditions are satisfied. The details can be obtained from the author upon request.

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. The individual's indirect utility function is given by substituting equations (2.4) and (2.5) back into (2.2):

$$V_i^l = s_i(1-f)n_i^h w_i \left[\frac{(1-s_i)(1-f)w_i}{\alpha}\right]^{\frac{1}{\beta}} + f\left(\frac{1-f}{\alpha}\right)^{\frac{1}{\beta}} \sum_{j=A,B} p_j n_j^h w_j \left[(1-s_j)w_j\right]^{\frac{1}{\beta}}. \tag{2.6}$$

The solution to the associated maximization problem is given next.

Lemma 1. *Given a federal tax rate f, state i's equilibrium tax rate is:*

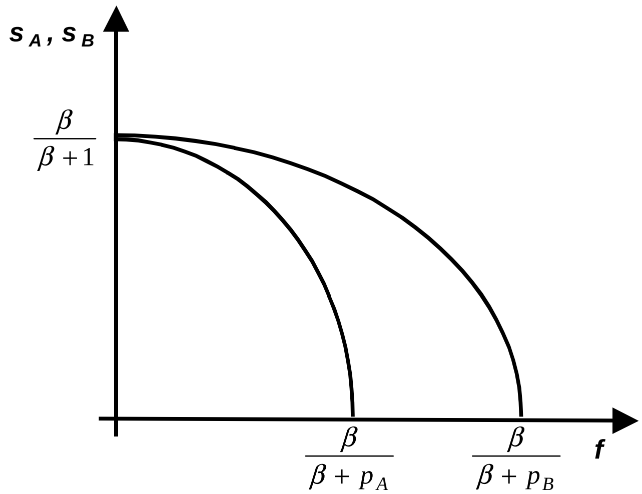
$$s_i = \begin{cases} \dfrac{\beta}{\beta+1} - \dfrac{fp_i}{(1-f)(\beta+1)} & \text{if } f < \dfrac{\beta}{\beta+p_i}, \\ 0 & \text{otherwise.} \end{cases}$$

This lemma illustrates the trade-off between federal and states' tax rates: *states' tax rates are decreasing in the federal rate* (see Figure 2.1). Incentive considerations are the primary reason behind this result. If the federal tax rate is already high, a high state tax rate *lowers* the state's redistribution level. That is, the maximum level of state's redistribution $r_i^{\max}$ (the maximum point of the state's Laffer curve) is a decreasing function of the federal tax rate. When f is equal to zero $r_i^{\max}$ attains its greatest value, which decreases monotonically in f, until $r_i^{\max}$ reaches zero for $f \geq \beta/(\beta + p_i)$.

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; that is, the required decrease in s_i (for a given increase in f) is increasing in the level of f. The sequential structure of the taxation process is the reason behind this result. From equation (2.3) we see that the individuals' optimal pre-tax income is not a linear function of the sum of the tax rates $s_i + f$. Because of this multiplicative structure, low-productivity individuals have to compensate high-productivity individuals with more signifi-

Figure 2.1: States' Tax Reaction Functions (for $p_A > p_B$)



cant decreases in s (for a given increase in f), as the level of f increases.[11] If low-productivity individuals would not compensate high-productivity individuals this way, the state's total income would considerably decrease with f, and consequently the state's level of redistribution would decrease as well.

The following subsection solves for the equilibrium federal tax rate.

2.2.2. Preferences Over the Federal Tax Schedule

In the first stage of the political process a federal tax schedule is chosen by majority rule. When deciding on their preferred federal tax rate, individuals take into account that states' tax rates react to the federal rate according to Lemma 1. A difficulty arises upon choosing the federal tax rate because preferences over f are not monotonic in w. Since individuals are heterogeneous in their state of

[11] Mathematically, 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 $\frac{d^2 s}{df^2} < 0$.

residence, low-productivity individuals in the federation never form a coalition to extract as much income as possible from high-productivity individuals. In fact, if neither of the four groups constitutes a majority of the population (this is the case if $p_i n_i^l < 1/2,\ i = A, B$) the decisive voter for the federal tax rate is an individual with high productivity, even though this productivity level is above the median productivity level. More surprisingly, if total productivity dispersion between the states is above a certain threshold, high-productivity individuals from the poor state choose to implement a high federal tax rate. By behaving this way, these individuals' aim is to increase redistribution between the states and decrease redistribution within each state.

The individuals' preferences over tax rates are given by maximizing their indirect utility function (equation (2.6)) 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

$$x \equiv \frac{n_B^h}{n_A^h} \left(\frac{w_B}{w_A} \right)^{\frac{1+\beta}{\beta}} \tag{2.7}$$

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 total 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 given by the ratio of the share of high-productivity individuals. As the elasticity of labor supply increases the difference between the individuals' productivities 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.[12]

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 ($p_A = p_B = 1/2$). This simplifying assumption allows us to obtain closed-form solutions that highlight the basic forces at work in the model. The general case exhibits basically the same properties; its analysis is relegated to the

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

appendix.[13]

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

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

Figure (2.2) depicts these preferences. To understand the intuition behind Proposition 1, we need to realize that the federal income tax rate provides a benefit to low productivity individuals, but also imposes a cost on them. The benefit is straightforward: a positive federal tax rate implies a positive federal redistribution level. The cost imposed by the federal tax rate is a consequence of the trade-off between federal and state redistribution levels: 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.

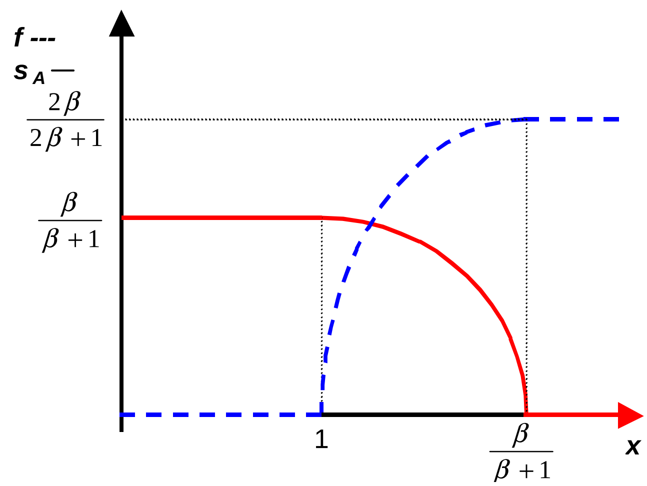
For example, suppose that the total productivity of state A is higher than that of state B. In this case, low-productivity individuals living in A receive federal redistribution if the federal tax rate is positive. A positive federal tax rate, however, also has two negative effects on these individuals' welfare: it transfers income from state A to state B, and it lowers their state's redistribution level. If instead low-productivity individuals in A manage to set the federal tax rate equal to zero, they appropriate (through an increase in s_A) part of the transfer between the states to themselves. Low-productivity individuals in state A prefer this last alternative as it maximizes their total income. If it is the case that $x > 1$, the opposite argument applies, and low-productivity individuals in A prefer a positive federal tax rate.

How high is their preferred federal tax rate? It depends entirely on x. As inequality between the states increases, the gains from redistribution between the states for low-productivity individuals residing in A increase as well.[14] Therefore,

[13] The assumption that $p_{\mathsf{A}} = p_{\mathsf{B}} = 1/2$ is maintained in all the forthcoming results that appear in the main body of the paper.

[14] Whenever the implications of changes in x are analyzed, total productivity at the federal level ($n_{\mathsf{h}}^{\mathsf{A}} w_{\mathsf{A}} + n_{\mathsf{h}}^{\mathsf{B}} w_{\mathsf{B}}$) is assumed constant.

Figure 2.2: Preferred Federal Tax Rate for Low Productivity Individuals from State A



they prefer a higher federal tax rate. This explains why f_l^A is increasing in x. The cost 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 a decrease in s_A) has a large disincentive effect and actually lowers federal redistribution. This defines the second threshold value of x, above which both the federal and state income tax rates are constant: $f_l^A = 2\beta/(2\beta+1)$ and $s_A = 0$ respectively.

Due to the symmetry of the analysis, the preferences of low-productivity individuals in state B are exactly opposed to the ones presented in Proposition 1. For these individuals the relevant cutoff values of productivity dispersion are equal to x^{-1}. More specifically, the preferred federal income tax rate of low-productivity individuals in B, f_l^B, is decreasing in x. It reaches a maximum of $2\beta/(2\beta+1)$ when $x < \beta/(\beta+1)$, and a value of zero when the dispersion level is greater or equal to one. In the intermediate range $f_l^B = [2\beta(1-x)] \,/\, [1+\beta(1-x)]$.

From this argument follows that preferences over f are not monotonic in the productivity level. As a consequence of that the individual with the median income is not the decisive voter in this framework. Given that the overall population is equally divided between the two states, none of the four different groups of individuals comprises a majority. Therefore we need to study high-productivity individuals' preferences over f to find out if some consensus may emerge between the different groups. 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 possible tax rate will exist. The preferences of high-productivity individuals in A appear in

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

$$f_h^A = \begin{cases} 0 & \text{if } x \leq \underline{x}, \\ \dfrac{2\beta\left[(\beta+1)(x-1)n_A^h - 1\right]}{\beta\left[(\beta+1)(x-1)n_A^h - 1\right] + (\beta+1)xn_A^h} & \text{if } \underline{x} < x \leq \overline{x}, \\ \dfrac{2\beta}{2\beta+1} & \text{if } \overline{x} < x. \end{cases}$$

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

Comparing the previous proposition to Proposition 1 one may conclude 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. Actually, the same intuition applies here as in Proposition 1, with one caveat. As it was the case with low-productivity individuals, also high-productivity individuals derive a benefit and suffer a cost from the federal tax schedule. For positive federal income tax rates, individuals with high productivity have to pay a proportion of their income in federal taxes, 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 the cutoff 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. In order 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 face (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 cutoff value $\underline{x}$ above one.

Combining Propositions 1 and 2 we observe that $f_l^B \geq f_h^B \geq f_h^A \geq f_l^A = 0$ for $x < 1$, and $f_l^B \leq f_h^B \leq f_h^A \leq f_l^A$ for $x > 1$.[15] Hence, the federal tax rate proposed by a high-productivity individual always obtain a majority over a tax rate proposed by a low-productivity individual, even though this last group comprises more than half of the overall population. 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 $f_h^i(x)$.*

According to this corollary, the decisive voter's productivity level is above the median productivity. This was somehow expected from the moment we established that preferences over federal tax rates were not monotonic on productivity. Due to this non-monotonicity, the one-to-one relation between the individuals' productivity and their preferred tax rate does not hold; hence the result above.

Persson 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 no coalition of voters is formed across borders: all voters in the rich state want smaller transfers than all voters in the poor state. In fact, the citizens of one state support total exploitation of the other state. When transfers are not allowed and the policy space of the federal government consists of a federal income tax, Persson and Tabellini (1996) still find that the state of residence has an important effect on the individuals' preferences, in the sense that residents of the rich state tend to prefer lower federal tax rates than residents of the poor state.

The next section focuses on the question whether a federal social contract with a two-tier income taxation system is efficient.

[15] For $x < 1$ we obtained in Propositions 1 and 2 that $f_{\mathsf{h}}^{\mathsf{A}} = f_{\mathsf{l}}^{\mathsf{A}} = 0$. When $x > 1$ it is simple to show that

$$\frac{2\beta(x-1)}{x+\beta(x-1)} > \frac{2\beta\left[(\beta+1)(x-1)n_{\mathsf{h}}^{\mathsf{A}} - 1\right]}{\beta\left[(\beta+1)(x-1)n_{\mathsf{A}}^{\mathsf{h}} - 1\right] + (\beta+1)xn_{\mathsf{A}}^{\mathsf{h}}}$$

if and only if $x > 0$, which is always the case.

2.3. Efficiency Analysis

This section shows that the current federal social contract consisting of a two-tier income taxation system is not efficient for a wide range of productivity dispersion levels. As already mentioned in the introduction, when choosing the federal income tax rate an individual in state i ignores the effect 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 will bring 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 them, 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 approach adopted in this paper and the one carried on in the previous literature are worth mentioning. First, the previous papers abstracted from political economy considerations, the main focus of the present paper. Second, in this paper, horizontal externality is assumed away since individuals are immobile. In addition, given that individuals (and not a social planner) choose the federal tax rate through majority vote, 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 state 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. Total redistribution of the federation is given by

$$R(x) \equiv r_f + r_A + r_B = \frac{\mathcal{P}}{2}\left[\frac{(1-s)(1-f)}{\alpha}\right]^{\frac{1}{\beta}}[f + s(1-f)] \tag{2.8}$$

where

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

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 $x > \overline{x}$.[16] For intermediate values of x, however, total redistribution is strictly decreasing in x.

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

In this same range of dispersion, total income taxation 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, since a reduction in income taxes would increase total redistribution.

Proposition 3. *The federal social contract consisting of a two-tier income taxation system is not optimal whenever $x > \underline{x}$.*

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.

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 transfer between the states and eliminating one layer of taxation would achieve a Pareto improvement. But even abstracting from lump-sum transfers, for sufficiently high levels of productivity dispersion, a Pareto improvement can be achieved eliminating states' income tax schedules.

Proposition 4. *There exists a critical dispersion level $x^* < \overline{x}$ such that for every $x > x^*$ eliminating states income tax schedules results in a greater utility level for all the individuals in the federation.*

What is the reason that drives the states to this inefficient outcome, even when a different feasible strategy would imply greater utility for all the individuals? The answer to this question is related to a usual source of inefficiency: the impossibility to credible commit to a non-equilibrium strategy. Since low-productivity

[16] Although the results are stated for the case where $p_{\mathsf{A}} = p_{\mathsf{B}}$, as I said earlier, they could be generalized to the case where $p_{\mathsf{A}} > p_{\mathsf{B}}$. Thus, from now on I take $f_{\mathsf{h}}^{\mathsf{A}}$ as the equilibrium federal income tax rate.

individuals constitute a majority in each state, they choose the equilibrium income tax schedule at the state level. Suppose these individuals promise to choose a state's tax rate of zero, no matter what is the level of the 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, low-productivity individuals in both states will select in the second stage of the political process a positive state's tax rate that 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 one of the states, an efficient outcome will not be reached.[17] In this situation, low-productivity individuals in the other state are enjoying the best of both worlds. They receive more federal redistribution given that high-productivity individuals in the other state have higher incomes due to incentive effects, and 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. Hence, for low-productivity individuals the implementation of a state's income tax schedule is a dominant strategy. Consequently, we obtain 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.

2.3.1. Federal Matching Grants

In models of fiscal competition one role for the federal government is to implement policies that undo the nonoptimality arising from decentralized state decision-making. For example, Wildasin (1991) shows how a system of matching grants from a federal government to state governments can neutralize the horizontal externalities created by states' policies, helping the federation reach an efficient outcome.[18] It is then natural to explore the implications of such a policy using the current framework.

The mentioned papers assume that the federal government will act as a benevolent social planner correcting inefficiencies created by lower level governments.

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

[18] A similar suggestion, without a formal analysis, appears in Gordon (1983).

Notwithstanding the insights of that normative approach, in this paper I am committed to a democratic principle, which is a different normative guideline. In particular, I assume that the decisions at the federal level are still adopted by majority rule. When this is the case, a system of federal matching grants in general increments, rather than diminishes, the resulting inefficiencies; that is, for a wide range of inequality levels, total income taxation further increases and total redistribution decreases. While this conclusion may look surprising at first, upon reflection it is even expected. Because matching grants reduce the cost of redistribution at the state level (which is financed in part now by the federal government) individuals in the poor state choose a higher equilibrium state tax rate to finance greater redistribution at the state level. This in turns implies greater matching grants, and consequently, a higher federal income tax rate.

Under a system of federal matching grants, *state i's tax schedule* is given by

$$\delta r_i = p_i s_i (1-f) n_i^h y_i, \tag{2.9}$$

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 the *federal tax schedule* is equal to

$$r_f + (1-\delta)(r_A + r_B) = f\left[p_A n_A^h y_A + p_B n_B^h y_B\right]. \tag{2.10}$$

As it is readily seen from the previous two equations, under a matching grants program the state is responsible for only a share of its redistribution expenses. As a consequence, there is a popular support for more redistribution. Since the federal government pays part of this greater level of redistribution, a higher federal tax rate is required by the balanced budget constraint condition. This leads to higher income taxation and lower total redistribution.[19]

Proposition 5. *Under a federal matching grants program, if*

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

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

[19] This result 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 $\beta \ge (1 + 2n_{\mathsf{A}}^{\mathsf{h}} + \delta)/(1 - 2n_{\mathsf{A}}^{\mathsf{h}} - \delta)$, with $1 - 2n_{\mathsf{A}}^{\mathsf{h}} - \delta > 0$.

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.

That is, the effect of a federal matching grants program under a democratic system is the opposite of the effect obtained when the federal government is represented by a benevolent social planner.

So the original question that motivated the previous analysis remains: Is it possible to obtain a Pareto improvement in a democratic federation of states? This is a very interesting issue, but an exhaustive analysis of it remains beyond the scope of the present paper. The answer is then left open for future research.

The next subsection provides an empirical illustration of the main implications of the model.

2.4. Empirical Implications

Corollary 1 above already pointed out the first implication of the model: in a federal system of taxation there is not a one-to-one relation between the individuals' income and their preferred federal tax schedule. 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. Hence, under universal suffrage the decisive voter is the individual with the median productivity level.

This is not the case in a federal system of taxation. Rather, in a federal system of taxation, the decisive voter is a relatively productive individual who, for a wide range of parameters values, chooses to implement a positive federal income tax rate. More strikingly, an interesting situation may arise when the decisive voter's productivity is above the federation's mean productivity, yet she selects a positive federal income tax rate. The following proposition formalizes this observation for individuals in state A.

Proposition 6. *If $\underline{x} < x \leq (2-n_A^h)/n_A^h$ 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 6, for a certain range of x, as income inequality within a state increases, redistribution in that state decreases. To see this, consider what the model predicts when the productivity of high-productivity individuals in state A decreases while the productivity of high-productivity individuals in B increases. Such changes increase the productivity dispersion between the states. Hence, if the implemented federal income tax schedule is f_h^A (an increasing function of x), the resulting federal tax rate will increase as well. Consequently, both states' income tax rates will decrease, even if the median voter's productivity relative to the mean productivity 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 model of income taxation.[20] This is perhaps the reason why several studies using data from the US 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).[21]

So what should we expect in a federal system of income taxation? According to this model, what really matters is not productivity dispersion *within* each state, but productivity dispersion *between* the states. In fact, total productivity dispersion between the states is the key variable determining both f and s. As stated in Proposition 2, there exists a positive relation between the productivity dispersion level and the federal tax rate, while there exists a negative relation between x and the states' tax rates. Are these hypotheses supported by the data? While it is difficult to provide a definitive answer, Figures (2.3) and (2.4) present a first exploration.

Figure (2.3) depicts the total income of the nine regions of the US between 1940 and 2000. As it is well documented in the growth literature, the figure shows a high level of income convergence across these regions (Barro and Sala-i-Martin, 1995, chapter 11). Obviously, convergence of states' incomes means that the dispersion between them decreases.

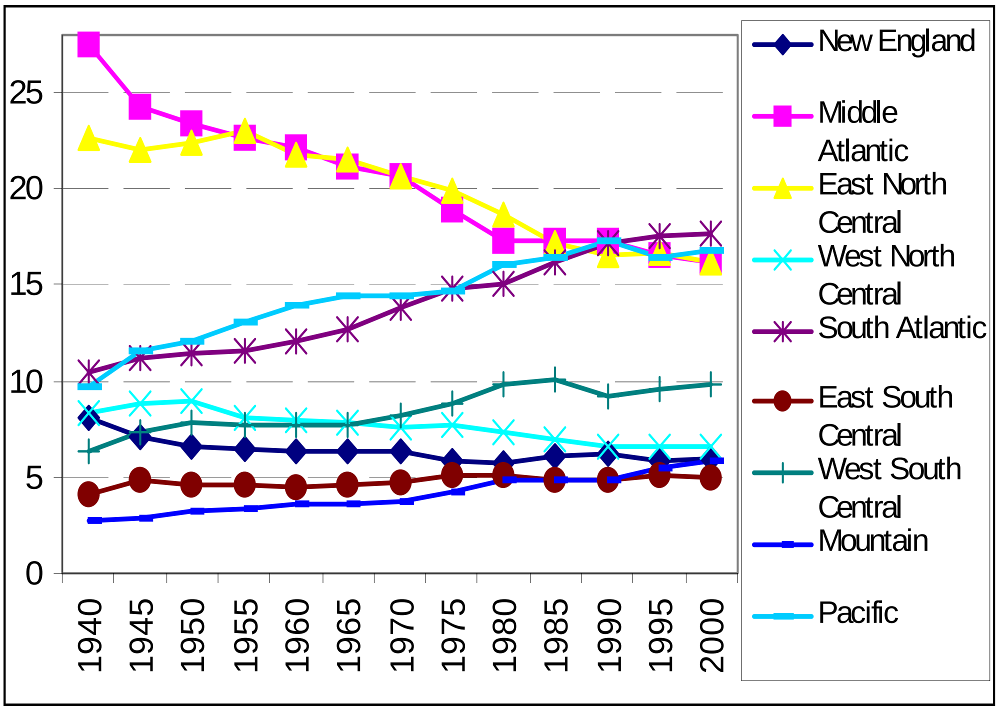
The resulting decrease in productivity dispersion implies, according to the model, lower federal marginal tax rates and higher states' marginal tax rates. These two sequences are depicted in Figure (2.4). We can see in this figure the significant decrease in the federal income tax's top marginal rate. This top rate oscillated between 91% and 88% right after World War II until the Kennedy-Johnson's tax cut brought it to 70% in 1965. The marginal federal tax rate at the top bracket remained at that level until the Tax Reform Act of 1981 lowered it by 20%. The Tax Reform Act of 1986 implemented a further decrease, cutting the top bracket to almost 30% during the late 1980s. Finally, in the 1990s the top federal marginal rate increased by almost 10%, to 39.5%.

The second sequence in the same figure depicts the average top marginal income tax for the 50 states and the District of Columbia. While the absolute

[20] While r_{A} will certainly decrease, whether r_{B} decreases or increases depends on the relative change in w_{B}.

[21] See for example Goveia and Masia (1998), and Rodriguez (1999).

Figure 2.3: Total Income for the Nine Regions of the United States, 1940-2000 (Percentage of U.S. Averages)

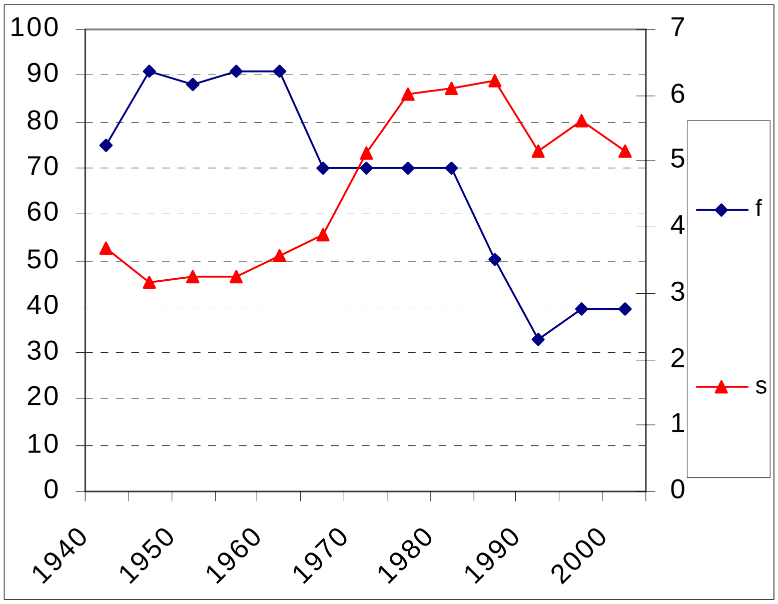


changes in this sequence are not as significant as the ones presented in the federal tax rate, there is a clear tendency upward. During the analyzed period the average top marginal income tax rate increased more than 50%, from 3.5% in 1950 to 5.5% in 2000.[22] More importantly, the number of states implementing an income tax increased steadily during this period. By 1940, only 30 states collected income taxes, albeit with very low rates.[23] This number increased to 40 by the mid 1970s,

[22] Feenberg and Rosen (1986), in a rigorous empirical study, calculated the states' effective marginal tax rate for high, middle and low-income individuals for the years 1977-1983. They found that the states' marginal tax rates increased from 3.27% to 4.19% for low-income individuals; from 4.4% to 5.62% for middle-income individuals; and from 5.5% to 5.93% for high-income individuals.

[23] Already by 1930 an economic model of taxation developed by the National Industrial Conference Board recommended a top marginal rate for the state income tax of about 6% (National

Figure 2.4: Top Marginal Rates: Federal and Average of the States, 1940-2000



with seven states implementing an income tax schedule between 1967 and 1971, right after Johnson lowered the top federal marginal income tax by almost 20%. Connecticut, in 1991, was the last state so far to implement a positive income tax rate.

Clearly, these figures are insufficient to prove that the reduction in productivity dispersion between states is the reason (certainly not the sole reason) behind the changes in marginal income taxes. Although the facts that the figures present support the main implication of the model, a richer data set accompanied by a rigorous empirical analysis are required to determine whether the federal and the states' income tax schedules are connected through the productivity dispersion level between the states. I make a first pass at this issue below.

In the next section, the analysis shift to non-linear tax schedules in order to account for the regressivity of the overall tax system of every state in the US.

Industrial Conference Board, 1930). At the time the average top marginal rate was below 3%. From the figures just presented it turns out that the 6% level was only reached 55 years later.

3. Non-Linear Tax Schedules

3.1. Theoretical Framework

I now turn to the second observation mentioned in the introduction: the regressivity of the overall tax system for every state in the US. Even if this fact is puzzling at a first look, upon reflection is not so surprising after all. To see this, notice that states relay mainly on three different taxes to collect revenue: income taxes, sales taxes, and property taxes. Given that the marginal propensity to consume decreases with income, the effective sales and property taxes are average rate regressive; that is, for these taxes the ratio of tax liabilities over the individual's income decreases with income. Since the income tax (which is the only progressive tax schedule used) is low, as a result the overall effective tax system at the state level is regressive.

To explain this strong regularity we need to enlarge the set of feasible tax schedules. In particular, we need to allow also for nonlinear tax schedules so the marginal tax rate can vary with income. This enables us to consider regressive as well as progressive tax schedules, in order to examine the link between federal progressivity and state regressivity.[24] Allowing for nonlinear tax schedules introduces new difficulties when solving the general model. Therefore, in this section I only solve for each state's tax schedule as a function of the federal tax schedule. From this result I derive a number of empirical implications that are tested in the next section using the available data. The theoretical model presented in this section builds on a static version of Bénabou (2000), adding a second tier of taxation.[25]

Formally, this section focuses only in one state of a federation of states. This state is populated by a continuum of individuals denoted by $i \in [0,1] \subset P$, where P is the federation's population. Each individual is endowed with a productivity level w. The productivity level is assumed to be lognormal distributed: $\ln w^i \sim$

[24] One way to obtain regressive tax schedules in the space of linear taxes is to allow for a negative tax rate. In fact, the previous model will account for regressive tax schedules (a negative tax rate) for high productivity dispersion levels if we adopt this extension of the tax space. I choose to change the set of feasible tax schedules, however, for two reasons. First, it seems unrealistic that every state of the U.S. implements negative tax rates. And second, the new specification will include in the analysis a higher moment of the states' income distribution -the variance- which has an important explanatory power of the progressivity of each state's tax schedule.

[25] This simple static version of Bénabou's model is rich enough to account for the main question raised in this paper, but fails to capture the richness of his model.

$\mathcal{N}(m, \Delta^2)$.[26] Individual i's preferences are represented by the following utility function:

$$u(c, \frac{y}{w}) = \ln c - \frac{\alpha}{\beta+1}\left(\frac{y}{w}\right)^{\beta+1}, \tag{3.1}$$

where all the parameters were defined in page (7).

As it was assumed in the linear taxation model, the federation's fiscal policy comprises a two-tier taxation system: there is a federal tax schedule τ_f and a state tax schedule τ_s, imposed sequentially on the individuals' income. In order to focus on the issue of tax progressivity, attention is restricted to a class of tax schedules that can be ordered according to their global degree of progressivity. I thus posit that the federal tax schedule satisfies

$$y - \tau_f(y) = (y_i)^{1-f}(r_f)^f = \widehat{y_i},$$

where $\widehat{y_i}$ is the income of individual i after paying the federal tax, and f is a scalar. Parameter f (called the residual progression of the federal tax schedule) represents the elasticity of post-tax income to pre-tax income. A tax schedule is progressive (i.e., the average tax rate increases with pre-tax income) if its residual progression is positive. If $f < 0$ the tax schedule is regressive. Furthermore, the progressivity of the tax schedule (according to the Lorenz domination criterion) increases with its degree of residual progression. In other words, for any given distribution of pre-tax income and tax schedules τ_f^1 and τ_f^2, if $f_1 > f_2$, then the distribution of post-tax income under τ_f^1 Lorenz dominates the one under τ_f^2 (Jakobsson, 1976). In the analysis that follows confiscatory rates ($f > 1$) are excluded as not incentive compatible.

The balanced government budget constraint determines the level of redistribution r_f. From the definition of τ_f it follows that

$$\int_P (y_i)^{1-f}(r_f)^f di = \int_P y_i di.$$

The state tax schedule τ_s is imposed on the individuals' income after federal taxes have been paid. This tax is given by:

$$\widehat{y_i} - \tau_s(\widehat{y_i}) = (\widehat{y_i})^{1-s}(r_s)^s = c_i, \tag{3.2}$$

[26] The lognormal distribution is a good approximation of empirical income distributions (Lydall, 1968). It also leads to tractable results, and allows for an unambiguous definition of inequality, as increases in the variance (Δ^2) shift the Lorenz curve outward.

where the break-even level r_s is determined by

$$\int_0^1 (\widehat{y}_i)^{1-s}(r_s)^s di = \int_0^1 \widehat{y}_i di.$$

The interpretation of the parameters of the state tax schedule is analogous to the one presented for the parameters of the federal tax schedule.

Given τ_f and τ_s, the optimal pre-tax income of individual i is

$$y_i = w_i \left[\frac{(1-s)(1-f)}{\alpha}\right]^{\frac{1}{\beta+1}}. \tag{3.3}$$

Substituting (3.3) into (3.2) we obtain the state's level of redistribution r_s,

$$s \ln r_s = f s \ln r_f + s(1-f)m + (1-f)^2(2s-s^2)\frac{\Delta^2}{2} + \frac{s(1-f)}{\beta+1} \ln\left[\frac{(1-s)(1-f)}{\alpha}\right], \tag{3.4}$$

which implies that individual i's indirect utility function is

$$\begin{aligned} V(w_i, \tau_f, \tau_s) &= (1-s)(1-f)\ln w_i + f \ln r_f + s(1-f)m + (1-f)^2(2s-s^2)\frac{\Delta^2}{2} \\ &+ \frac{(1-f)}{\beta+1} \ln\left[\frac{(1-s)(1-f)}{\alpha}\right] - \frac{(1-s)(1-f)}{\beta+1}. \end{aligned} \tag{3.5}$$

Individual i's preferred policy is given by the first-order condition $\partial V_i/\partial s = 0$:[27]

$$(1-s)(1-f)\Delta^2 + (m - \ln w_i) - \frac{s}{(\beta+1)(1-s)} = 0.$$

The above equation, which is quadratic in s, always has a unique solution where s is less than one. The fully characterization of it is

$$s = 1 + \frac{(m-\ln w_i)(\beta+1)+1}{2(1-f)\Delta^2(\beta+1)} - \frac{[(m-\ln w_i)^2(\beta+1)+4(1-f)\Delta^2]^{\frac{1}{2}}}{2(1-f)\Delta^2(\beta+1)^{\frac{1}{2}}}. \tag{3.6}$$

The previous equation delivers several intuitive results. First, the degree of progression of the individual's preferred tax schedule decreases with the individual's productivity. Also, a negative relation exists between the absolute value of

[27] When solving the maximization problem I assume that the state is small (both with respect to its income and its population). Therefore, the federal level of redistribution is not a function of the state's tax schedule. That is, $\partial r_f/\partial s = 0$.

s and the elasticity of labor supply. This is because a more elastic labor supply increases the deadweight loss from taxes and transfers, whether progressive or regressive, which cause individuals to distort their labor supply away from the first-best level. And finally, more relevant for the purpose of the current analysis, this solution provides several empirical implications, which are gathered in the following proposition.

Proposition 7. *For any $s < 1$,*

(1) *The residual progression of the state's tax schedule decreases with the residual progression of the federal tax schedule.*

(2) *If the decisive voter is the individual with the median productivity level, then the residual progression of the state's tax schedule increases with the state's income inequality.*

The first claim is nothing but the non-linear version trade-off between federal and states' tax progression: states' tax progression is decreasing in the degree of federal tax progression. As before, this is because of incentive considerations. If the federal tax schedule is already extracting most of the income of rich individuals, a very progressive state tax schedule reduces significantly the labor supply of productive individuals and thus the state's aggregate income. As a matter of fact, this trade-off may account for regressive states' tax schedules: there exists a level of f, $f^* < 1$, such that for every $f \geq f^*$, τ_s is regressive $(s < 0)$. Whether such a progressive federal tax schedule emerges in equilibrium is an important question. As pointed out before, I am still unable to answer it; solving for the federal tax schedule in this framework is complicated given that individuals' preferences over f are not monotonic in the individuals' productivity level. My conjecture is that an equivalent analysis to the one presented with the linear tax schedules will arise in this framework. Note that preferences over f are monotonic in the individuals' productivity for a given state of residence. Consequently, there shall be a decisive-voter in the pivotal state that will split the federal electorate in half.[28] The poorer the state of the decisive voter, or the productivity level of this voter, the more progressive the federal tax schedule should be. Eventually, a critical productivity dispersion level between the states will make the decisive voter choose a federal tax schedule as progressive as f^*.

The second claim is a particular case of Proposition 7 in Bénabou (2000,

[28] A state is pivotal if its productivity dispersion level is such that less than 50% of the federal population lives in states with higher x, and less than 50% of the population lives in states with lower x.

p. 109). This result states that in a more unequal society, without efficiency gains from redistribution, there is greater political support for redistribution. As inequality within the state increases, so does the progression of the tax schedule, even if this will reduce aggregate income.

A sufficient condition for this result is that the individual with the median productivity level is the decisive voter. As argued among others by Bénabou, this is not a realistic assumption. It is well known that poor and less educated individuals have a relatively low propensity to register, turn out to vote, and give political contributions (Rosenstone and Hansen, 1993). Consequently, the income of the decisive voter turns out to be located above the median income.[29] If the decisive voter's income is above the median income, we obtain a U-shaped relationship between the residual progression of the tax schedule and income inequality. That is, for a certain range of inequality levels, progression declines with inequality.[30]

The next subsection tests the empirical implications stated in Proposition 7.

3.2. Empirical Analysis

3.2.1. Data Description

This subsection presents empirical tests of the main implications of the model using cross-sectional data. Ideally, one would use a panel data set consisting of the states' residual progression index, the federal residual progression index, and the variance of income at the state level, together with each state's measure of productivity dispersion. With such a data set one would be able to test for the impact over time of productivity dispersion between the states on the federal and states degree of progression. Moreover, using the income inequality within each state, it could also be possible to account for the existent differences in residual progression among states at any given point in time. The binding constraint, however, is data: no such panel exists. The closest substitute is a new data set constructed by the Institute on Taxation and Economic Policy, which allows me

[29] How strong is this departure from the one-person one-vote rule? Bénabou (2000, p. 107) provides such a calculation. According to his results, the income of the decisive voter is located at the 56th percentile of the income distribution. While the income bias is relatively moderate, it is much stronger for other forms of political participation, like money contributions (74th percentile), and campaign related work (68th percentile).

[30] Due to the inexistence of the relevant data, I am unable to test empirically this particular formulation of the model. Therefore, I decided not to include a formalization of this scenario. The interested reader can find a formal model and a thorough discussion in Bénabou (2000).

to calculate the degree of residual progression in 1995 for each state of the US and the District of Columbia.[31]

Using the aforementioned data set I am able to calculate s, the degree of residual progression of the overall tax schedule for each state. Formally, the degree of residual progression is defined as the elasticity of post-tax income to pre-tax income:

$$RP(y) = \frac{1 - M(y)}{1 - T(y)},$$

where M denotes the marginal tax rate and T denotes the average tax rate.[32] $RP(y)$ measures the actual percentage increase in post-tax income. A reduction in $RP(y)$ must be interpreted as an increase in progression. Whenever $RP(y) < 1$ (this is the case if and only if $M < T$) the tax schedule is progressive at income level y; the tax schedule is regressive at income y whenever $RP(y) > 1$.

It is inconvenient that residual progression, defined this way, decreases when the tax becomes more progressive. Therefore, I use the following normalization

$$RP^{*}(y) = \frac{1}{RP(y)}$$

which was proposed by Lambert (1993, p. 161) to solve that particular problem.

Clearly, $RP^{*}(y)$ is a function of income. Yet, in the theoretical model, the residual progression of the tax schedule is assumed constant for every income level. Therefore, the average degree of residual progression is used as a proxy to *the* residual progression of a tax schedule.

To further link this measure to the one used in the theoretical model, I subtract one from the average RP^{*}; that is,

$$s = RP^{*} - 1.$$

Hence, as defined in the model, a tax schedule is progressive if and only if its index of residual progression is positive; and the progressivity of the tax schedule increases with s.

[31] To the best of my knowledge, this is the only data set available that provides the effective average tax liability for nine different income ranges of the population, for every state. Unfortunately this data set is only available for 1995. A detailed description of the data sources appears in Appendix A.

[32] This measure was first proposed by Musgrave and Thin (1948) and is still widely used in the literature of income taxation.

The federal tax schedule's index of residual progression is obtained using a similar procedure. To calculate this index I use a data set constructed by the Congressional Budget Office (1998). This data set provides the effective average tax rates by income distinguishing between five different types of families.[33] Using this data, and each state's demographic characteristics, a residual progression index of the federal tax schedule is obtained for each state. Given that the federal tax schedule is shared by all the states, there is not a lot of variability on f across the states. While this may weaken the results to certain extent, the level of residual progression of the federal tax schedule at the state level captures the different way the same statutory tax schedule affects the effective tax liability of different populations.[34]

Finally, the different parameters from the income distribution of each state are obtained from the Current Population Survey of 1992.[35] These parameters are calculated at the family level to match the similar definition used in the progression indexes.[36] In particular, using the family's adjusted gross income I calculate the variance of the income distribution for each state, and also construct an index of income inequality between the states, which is used as a proxy to x. The formal generalization of x to an expression for 51 observations is

$$x_j = \frac{\sum_{i=1}^{51} \mu_i - \mu_j}{50\mu_j}$$

where μ_i is the mean family adjusted gross income of state i. As defined in the theoretical model this index is always positive; and there is a negative relationship between x_j and the mean family income of state j: the greater x_j is, the poorer the state.

Table 2 presents descriptive statistics for the variables used in the empirical exercise.

[33] In particular, it provides the effective average tax for families with children under 18 with: (a) one adult; or (b) two or more adults. And for individuals and families with no children under 18 where: (a) the head is 65 or more years old; (b) the head is less than 65 years old; and (c) individuals.

[34] For example, the effective average tax rate is lower for a family with children than otherwise. Hence, the residual progression of the federal tax schedule in a given state increases with the frequency of families with children living in that state.

[35] These parameters are taken with a lag of three years, so they can reasonable be regarded as exogenous in these regressions.

[36] Calculations carried on at the individual level deliver similar results. These results can be obtained from the author upon request.

Table 2: Descriptive Statistics

	Mean	St. Dev.	Median	Max	Min
s	-0.071	0.016	-0.072	-0.030	-0.116
f	0.108	0.005	0.107	0.116	0.095
Δ^2	1.2e+12	7.9e+11	1.2e+12	3.1e+12	1.7e+11
x	1.017	0.133	1.027	1.404	0.753

A quick look at the table reveals some interesting facts. As expected, the states' index of residual progression is negative for all the states. The more progressive tax is implemented in Delaware, while Washington State has the most regressive tax schedule. This is not surprising. On the one hand, Washington State makes a heavy use of sales and excise taxes, which are regressive. On the other hand, this state does not have an income tax, the only progressive tax schedule. Hence the regressivity of its tax schedule. In contrast, Delaware relays heavily on an income tax (which is marginal-rate progressive) and makes low use of sales and excise taxes. Consequently, the tax schedule of this state is the less regressive among all the states of the US.

The index of federal residual progression is positive for all the states, and as predicted, has a low variance. The maximum value corresponds to the District of Columbia, while the minimum corresponds to Oklahoma. One reason for this is a relatively wealthy population in D.C. ($x_{DC} = 0.917$), and a relatively poor population in Oklahoma ($x_{OK} = 1.122$). Although important, this is not the only reason. There are richer states than D.C. and poorer states than Oklahoma. Some demographic characteristics of these states also influence their index of residual progression. For example, a high percentage of two-adults families in Oklahoma increases the average tax liability in that state, which brings a decrease in the residual progression of the tax schedule. In D.C., the high percentage of families with only one adult has exactly the opposite effect.

The statistics depicted in Table 2 for the remaining two variables are not new. States in the Pacific, Mid Atlantic and New England regions are relatively rich (the minimum x value is obtained by New Jersey), while the poorer states are situated in the East South Central region (Mississippi obtains the maximum value of x). Note that the mean of x is lower than its median. Hence, a majority of states are poorer than the average (this is the case for 29 of the 51 observations). This implies, according to the model, that the population of almost 60 percent of the states supports redistribution through the federal income tax.

With respect to within states inequality, North Dakota (a relatively poor state) presents the smallest income inequality. Surprisingly, the state with the greatest income inequality, Georgia, is also relatively poor; the value of its inequality index ($x_{GA} = 1.052$) is even above the median x.

In some of the estimated specifications, I include several control variables. These variables are: (1) state spending in redistribution (as a ratio of state income) either through the public supply of private goods (PG), or through pure redistribution (PR);[37] (2) a dummy variable that reflects the political affiliation of the governor ($Gov = 1$ if the governor belongs to the Republican Party; $Gov = 0$ otherwise); (3) the share of the state's legislators that belong to the Republican Party (Leg). The main purpose of these variables is to capture exogenous characteristics of each state's residents, not present in the model (like preferences for equality), which may influence the chosen degree of progression of each state's tax schedule.

3.2.2. Results

Table 3 reports the results of the empirical estimation of the model. The data in general support the main hypotheses, but I must note some reservations. On one hand, for all the different specifications used in the estimations, the coefficient of the federal tax schedule is significant and negative. Hence, the results are consistent with the existence of a trade-off between federal and state progressivity. On the other hand (and also robust for all the different specifications), the coefficient of the variance of the states' income distribution is negative and significant as well. While the sign of this coefficient contradicts its theoretical implication derived in Proposition 7, several reasons may bring this result. As mentioned above, abstracting from the idea that all the individuals in the state participate equally in the political process, and allowing for efficiency gains from redistributions, may cause such a negative relation (Bénabou, 2000).

In particular, Column (1) of the table presents the estimated coefficients of equation (3.6). According to this specification, there is empirical evidence of an existent trade-off between state and federal progression. In Column (2) the same specification is utilized but now controlling for several variables that may explain the progressivity of each state's tax schedule. The main results are robust for this different specification.

[37] These two definitions of spending where first stated by Meltzer and Richard (1983), and are also used by Gouveia and Masia (1998).

Table 3: Estimation of the State's Tax Residual Progression

	I	II	III	IV
f	-0.581	-0.334		
	(-14.18)	(-2.21)		
Δ^2	-6.93e-15	-9.04e-15	-6.48e-15	-8.04e-15
	(-2.21)	(-2.67)	(-2.10)	(-2.07)
$\widehat{f}$			-0.588	-0.401
			(-14.70)	(-2.34)
PG		1.168		1.136
		(2.57)		(2.65)
PR		-1.005		-0.946
		(-2.94)		(-2.97)
Gov		-0.009		-0.009
		(-1.93)		(-2.02)
Leg		-0.014		-0.011
		(-1.02)		(-0.75)
$\overline{R}^2$	0.9545	0.9615	0.9573	0.9629

Heteroskedasticity-consistent t-statistics in parenthesis.

A main objection to the previous results is that s and f may be simultaneously determined. In fact, according to the preceding analysis on linear income taxes, s and f are linked through productivity dispersion between the states. In the presence of simultaneity, ordinary least squares estimates are biased. To control for this simultaneity I estimate the model using a two-stage least squares estimation. In the first step f is regressed on x. The predicted value of f, $\widehat{f}$, is then used as a regressor in the second step, where s is the dependent variable. The results of this last regression are reported in column (3). From this column we see that $\widehat{f}$ (the variability in f explained by x) is significant and negative. This reinforces the message that comes out from the previous section: productivity dispersion between states is a significant variable linking the degrees of federal and state progression. In column (4) we obtain similar conclusions when controlling for each state's tastes for equality.

4. Conclusion and Extensions

This paper developed a positive theory of taxation in a federation of states. The results show 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 is a high-productivity individual, even though low-productivity individuals comprise a majority of the population. 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 state regressivity.

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.

Finally, the second part of the paper showed that similar theoretical implications are obtained under nonlinear taxation. These implications are empirically tested using a new data set constructed by the Institute on Taxation and Economic Policy. The empirical estimation confirmed that a significant negative relationship exists between the degree of residual progression of the federal and the states tax schedules. Furthermore, it is found that productivity dispersion between the states is a major force explaining this significant relationship between the progression of the two levels of taxation. Finally, it is also found that the dispersion of the overall income distribution within each state (as measured by its variance) affects negatively the degree of progression of each state's tax schedule. These findings are in concert with the implications of the present model.

Although it delivers new and interesting results, the model is highly stylized. Therefore, this paper developed a first exploration rather than a complete characterization of the subject being studied. As such, the model provides us a challenge to being generalized in several directions.

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's 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.

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. As already mentioned, corner solutions cause the indeterminacy of the states' tax schedule and thus of the federal tax schedule as well.

Finally, the model with nonlinear taxation can be improved upon to better reflect reality. The main direction is to differentiate between income, consumption, and housing expenditures and allow for three different tax schedules. In this framework, individuals would vote over composition of tax schedules. The insurmountable obstacle, so far, is to guarantee the existence of an equilibrium given the multidimensionality of the policy space. 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.

Appendix A: Data Description

I. Main Sources of Data

A. States' residual progression index: Institute on Taxation and Economic Policy (ITEP), 1996. The ITEP has build a microsimulation tax model that allows for a comprehensive analysis of the combined incidence of the major taxes affecting individuals and businesses at the state and local level. The main source of the data for this tax model comes from federal income tax returns (for the filer population) and the Decennial Census (for the non-filer population). Combined, they compose a comprehensive database of approximately 700,000 individual records calibrated to match state by state population, income and other aggregate totals. The economic unit of the data used in this tax model is the family.

B. Federal residual progression index: Congressional Budget Office, *Estimates of Federal Tax Liabilities for Individuals and Families by Income Category and Family Type for 1995 and 1999,* Table 4.

C. Parameters of the state's income distribution (mean, median and variance): U.S. Bureau of the Census, Current Population Survey, March 1992.

D. State government spending by category: U.S. Bureau of the Census, State Government Finances (1995).

E. Political Party of the Governor and share of Republican Legislators in each state: The Book of the States 1995.

II. Definitions of Government Spending

$PG-$ Public supply of private goods includes: higher education, local schools, other education, hospitals, sanitation, other natural resources, non-highway transportation, utilities and liquor stores.

$PR-$ Pure redistribution includes: Aid to Families with Dependent Children, Medicaid, unemployment compensation, supplemental security income, housing and urban renewal, and other insurance.

Appendix B: 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. In this section I solve the general problem for any $p_A \geq \frac{1}{2}$. The results stated in the respective propositions follow directly by setting $p_A = \frac{1}{2}$.

Since $p_A \geq \frac{1}{2}$, I solve 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:

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

$$+\gamma\tfrac{\beta}{\beta+1}\left(\tfrac{1}{\alpha}\right)^{\frac{1}{\beta}}\left[(1-s_A)(1-f)w_A\right]^{\frac{\beta+1}{\beta}}$$

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

where $\gamma = 0$ for low-productivity individuals and $\gamma = 1$ for high-productivity individuals.

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

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

The Lagrangian for the maximization problem above is:

$$\mathcal{L}(f, s_A, s_B, \lambda_A, \lambda_B, \theta_A, \theta_B; x) = s_A(1-f)a\left[\tfrac{(1-s_A)(1-f)}{\alpha}\right]^{\frac{1}{\beta}} +$$

$$f(1-f)^{\frac{1}{\beta}} \textstyle\sum_{i=A,B} p_i n_h^i w_i \left[\tfrac{(1-s_i)w_i}{\alpha}\right]^{\frac{1}{\beta}} + \gamma\tfrac{\beta}{\beta+1}\left(\tfrac{1}{\alpha}\right)^{\frac{1}{\beta}}\left[(1-s_A)(1-f)w_A\right]^{\frac{\beta+1}{\beta}}$$

$$+\textstyle\sum_{i=A,B}\left\{\theta_i s_i - \lambda_i\left[\tfrac{\beta}{\beta+1} - \tfrac{fp_i}{(1-f)(\beta+1)} - s_i\right]\right\}.$$

The corresponding first-order conditions are

1. $\mathcal{L}_f \Rightarrow \left(\frac{1-f}{\alpha}\right)^{\frac{1}{\beta}} \frac{(\beta-\beta f-f)}{\beta(1-f)} \left[ap_A(1-s_A)^{\frac{1}{\beta}} + bp_B(1-s_B)^{\frac{1}{\beta}}\right] -$

$s_A \frac{1+\beta}{\beta} a \left[\frac{(1-s_A)(1-f)}{\alpha}\right]^{\frac{1}{\beta}} - \gamma \left(\frac{1-f}{\alpha}\right)^{\frac{1}{\beta}} [(1-s_A)w_A]^{\frac{\beta+1}{\beta}}$

$+ \sum_{i=A,B} \frac{\lambda_i p_i}{(1-f)^2(\beta+1)} = 0.$

2. $\mathcal{L}_{s_A} \Rightarrow \left[\frac{(1-f)(1-s_A)}{\alpha}\right]^{\frac{1}{\beta}} \frac{(1-f)a}{\beta(1-s_A)}(\beta - \beta s_A - s_A) -$

$f\left[\frac{(1-f)(1-s_A)}{\alpha}\right]^{\frac{1}{\beta}} \frac{ap_A}{\beta(1-s_A)} - \gamma\left(\frac{1-s_A}{\alpha}\right)^{\frac{1}{\beta}} [(1-f)w_A]^{\frac{\beta+1}{\beta}}$

$+\lambda_A + \theta_A = 0$

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

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

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

With two inequality constraints and two non-negative variables, there are 16 possible patterns of equations and inequalities. Let us see which ones offer candidates for optimality.

The first candidate is given by $\lambda_i \geq 0$, and $\theta_i = 0$. This case correspond to the interior solution where states' tax schedules are positive. These tax rates (as a function of f) are given by Lemma 1. Substituting these two equations into the first three FOCs we obtain the following system of three equations:

$$\lambda_A = \left[\frac{1-(1-p_A)f}{\alpha(\beta+1)}\right]^{\frac{1}{\beta}} \gamma a \left(\frac{1-f}{n_h^A}\right) \geq 0, \tag{4.1}$$

$$\lambda_B = \left[\frac{1-(1-p_B)f}{\alpha(\beta+1)}\right]^{\frac{1}{\beta}} \frac{fbp_B(1-f)(\beta+1)}{\beta\left[1-(1-p_B)f\right]} > 0 \tag{4.2}$$

and by replacing λ_i in $\mathcal{L}_f$ according to (4.1) and (4.2) we obtain an implicit function $H(f,x)$ that provides the unique solution $f^*(x)$:

$$H(f,x) = \left[\frac{1-(1-p_A)f}{\alpha(\beta+1)}\right]^{\frac{1}{\beta}} \left[\frac{\gamma}{n_h^A} + \left(\frac{\beta}{\beta+1} - \frac{fp_A}{(\beta+1)(1-f)}\right)\left(\frac{\beta+1}{\beta} - \frac{\gamma}{n_h^A}\right)\right] +$$
$$\left(\frac{1}{\alpha(\beta+1)}\right)^{\frac{1}{\beta}} \left(1 - \frac{f}{\beta(1-f)}\right) \left[p_A\left(1-(1-p_A)f\right)^{\frac{1}{\beta}} + Ip_B\left(1-(1-p_B)f\right)^{\frac{1}{\beta}}\right] +$$

$$
\begin{aligned}
& \left[\frac{1-(1-p_A)f}{\alpha(\beta+1)}\right]^{\frac{1}{\beta}} \frac{\gamma p_A}{n_h^A(\beta+1)(1-f)} + \left[\frac{1-(1-p_B)f}{\alpha(\beta+1)}\right]^{\frac{1-\beta}{\beta}} \frac{\alpha I f p_B^2}{\beta(\beta+1)(1-f)} \\
= \; & 0 \qquad (4.3)
\end{aligned}
$$

The necessary condition

$$
0 \leq f^*(x) \leq \frac{\beta}{\beta + p_A}
$$

delivers the first and second threshold values of x. [38]

Whenever $p_A > 1/2$, there is a range of inequality levels such that s_A is equal to zero while s_B is positive. This is the second relevant range in which $\lambda_A = 0$ and $\theta_A > 0$ while $\lambda_B > 0$ and $\theta_B = 0$. Substituting for s_A and s_B into the first and the third FOCs allow us to solve for $f(x)$ and $\lambda_B(x)$. The relevant range for the inequality level is obtained by checking the following condition for f,

$$
\frac{\beta}{\beta + p_A} < f < \frac{\beta}{\beta + p_B}\text{fl}
$$

which is necessary for this case to be a solution.

Finally, the corner solution is reached when $\lambda_i = 0$ and $\theta_i > 0$, $i = A, B$. In this case

$$
f = \frac{\beta}{\beta + p_B} \text{ and } s_A = s_B = 0.
$$

So far, the proof above demonstrated how f^A is defined to both low and high-productivity individuals. Moreover, the proof stated clearly how the thresholds values are defined. To complete the characterization of the federal tax rate as illustrated in the body of the paper, it remains to show that f^A is increasing in x for all the residents of state A. This task is carried on in

Claim 1: f^A is nondecreasing in x.

Proof: Clearly the claim holds true for $x \leq 1 + \frac{\gamma}{n_A^h(\beta+1)} \equiv x_1$, given that in this range f^A is always equal to zero. The claim is also true for

$$
x \geq \frac{(\beta+1)(n_A^h p_A^2 + \gamma p_B)}{n_A^h p_B \left[\beta(1-p_A) - (p_A - p_B)\right]} \equiv x_2
$$

[38] The bordered Hessians providing the second order sufficient conditions for this and the following cases are cumbersome. The details can be obtained from the author upon request.

since f^A is constant in this range. Therefore, we need to prove that $f^A(x)$ is increasing for $x_1 < x < x_2$. Recall that in this range $f(x)$ is implicitly defined by (4.3). Note that

$$\frac{\partial H}{\partial x} = \left(\frac{1}{\alpha(\beta+1)}\right)^{\frac{1}{\beta}} p_B \left(1-(1-p_B)f\right)^{\frac{1}{\beta}} \left(1-\frac{f}{\beta(1-f)}\right) + \left[\frac{1-(1-p_B)f}{\alpha(\beta+1)}\right]^{\frac{1-\beta}{\beta}} \frac{\alpha f p_B^2}{\beta(\beta+1)(1-f)}$$

is positive for $f \leq \beta/(\beta+p_A)$. Since $\partial H/\partial f$ is negative, the claim follows from the implicit function theorem. ¥

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

$$s = \frac{\beta}{\beta+1} - \frac{f}{2(1-f)(\beta+1)}. \tag{4.4}$$

Substituting (4.4) into (2.8) 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 in the relevant range. ¥

Proof of Proposition 3. From Lemma 2 we know that total redistribution is strictly decreasing in this range. It remains to show that total income taxation is greater than $\beta/(\beta+1)$, the tax rate that corresponds to the highest level of total redistribution. For $x > \underline{x}$ we have that

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

which is greater than $\frac{\beta}{\beta+1}$ if and only if

$$f < \frac{2\beta+1}{2\beta+2}.$$

This will always be the case since the maximum possible value of f in equilibrium, $\frac{2\beta}{2\beta+1}$, is strictly less than $\frac{2\beta+1}{2\beta+2}$. ¥

Proof of Proposition 4. When

$$x \geq \overline{x}$$

the equilibrium income tax schedules are $s_i = 0$ and $f_h^A = 2\beta/(2\beta+1)$. Without states' income taxes, poor 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 $f = \beta/(\beta+1)$. Under this tax rate, federal redistribution is equal to

$$r_f = \frac{\beta}{\beta+1}\left(\frac{1}{\beta+1}\right)^{\frac{1}{\beta}}\left(\frac{a+b}{2}\right),$$

and is greater than the one obtained when states' income taxes are allowed. Since the tax rate is lower in this case, there is a welfare improvement for all the federation's individuals. In fact, this will be case for every $x > x^*$, where x^* is defined by

$$\left[\frac{(1-f)(1-s)}{\alpha}\right]^{\frac{1}{\beta}}\left[f\left(\frac{a+b}{2}\right)+s(1-f)a\right] = \frac{\beta}{\beta+1}\left[\frac{1}{(\beta+1)\alpha}\right]^{\frac{1}{\beta}}\left(\frac{a+b}{2}\right).$$ [39]

Proof of Proposition 5. Under a federal matching grants program, the indirect utility level of low-productivity individuals in state A is

$$\begin{aligned} V_l^A &= s_A(1-f)n_A^h w_A\left[\frac{(1-s_A)(1-f)w_A}{\alpha}\right]^{\frac{1}{\beta}}\left[\frac{1-(1-\delta)p_A}{\delta}\right]+ \\ & f\left(\frac{1-f}{\alpha}\right)^{\frac{1}{\beta}}\sum_{i=A,B} p_i n_i^h w_i\left[(1-s_i)w_i\right]^{\frac{1}{\beta}} - \frac{(1-\delta)}{\delta}r_B \end{aligned} \quad (4.5)$$

[39] Since all the tax rates are a continuous function of x, it implies that total redistribution is a continuous function of x as well. Furthermore, as proven in Lemma 2, total redistribution is strictly decreasing in inequality in this relevant range. Hence, such an x^* exists and is uniquely defined.

obtained by substituting equations (2.3), (2.9), and (2.10) back into (2.2). The implemented state's tax schedule in this state is obtained by maximizing (4.5) over the set of feasible state taxes. The solution to that maximization problem yields[40]

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

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

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

for low-productivity individuals, and

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

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

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

establishing the desired result. ¥

Proof of Proposition 6. $x < \frac{2-n_A^h}{n_A^h}$, implies that

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

[40] The 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.

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 Proposition 7. Both claims follow from Bénabou (2000, p. 122). The second claim follows directly by setting $\lambda = 0$ according to Bénabou's notation. To establish the first claim, notice that $\partial s / \partial f$ has the opposite sign of $\partial s / \partial \Delta^2$. ¥

REFERENCES

Barro, Robert J. and Sala-i-Martin, Xavier. *Economic Growth.* New York: McGraw-Hill, 1995.

Bénabou, Roland. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review,* March 2000, 90(1), pp. 96-129.

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.

Bohn, Henning and Stuart, Charles. "Voting over Non-Linear Taxes in a Stylized Representative Democracy." Mimeo, University of California at Santa Barbara, 2001.

Congressional Budget Office. *Estimates of Federal Tax Liabilities for Individuals and Families by Income Category and Family Type for 1995 and 1999.* Washington D.C., May 1998.

De Donder, Philippe and Hindricks, Jean. "The Politics of Progressive Income Taxation with Incentive Effects." Mimeo, Queen Mary, University of London, 2001.

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.

Feenberg, Daniel R. and Rosen, Harvey S. "State Personal Income and Sales Taxes, 1977-1983," in Harvey S. Rosen, ed., *Studies in State and Local Public Finance.* Chicago, IL: Chicago University Press, 1986, pp. 135-179.

Gordon, Roger H. "An Optimal Taxation Approach to Fiscal Federalism." *Quarterly Journal of Economics,* November 1983, 98(4), pp. 567-586.

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.

Inman, Robert P. and Rubinfeld, Daniel L. "Designing Tax Policy in Federalist Economies: an Overview." *Journal of Public Economics*, June 1996, 60(3), pp. 307-334.

Institute on Taxation and Economic Policy. *Who Pays? A Distributional Analysis of the Tax Systems in all 50 States.* Washington D.C., June *1996.*

Jakobsson, Ulf. "On the Measurement of the Degree of Progression." *Journal of Public Economics,* January-February 1976, 5(1-2), pp. 161-168.

Johnson, William R. "Income Redistribution in a Federal System." *American Economic Review,* June 1988, 78(3), pp. 570-573.

Lambert, Peter J. *The Distribution and Redistribution of Income: A Mathematical Analysis.* Manchester: Manchester University Press, 1993.

Lydall, Harold F. *The Structure of Earnings.* London: Oxford University Press, 1968.

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.

Meltzer, Allan and Richard, Scott F. "Tests of a Rational Theory of the Size of the Government." *Public Choice,* 1983, 41(3), pp. 403-418.

Musgrave, Richard A. and Thin, Tun. "Income Tax Pregression 1929-48." *Journal of Political Economy,* December 1948, 56(6), pp. 498-514.

National Industrial Conference Board, Inc. *State Income Taxes: Analysis of Income Taxes in State Fiscal Systems.* New York, February 1930.

Persson, Torsten and Tabellini, Guido. "Federal Fiscal Constitutions: Risk Sharing and Redistribution." *Journal of Political Economy,* October 1996, 104(5), pp. 979-1009.

Roberts, Kevin W. S. "Voting over Income Tax Schedules." *Journal of Public Economics*, December 1977, 8(3), pp. 329-340.

Rodriguez, Francisco C. "Does Distributional Skewness lead to Redistribution? Evidence from the United States." *Economics and Politics*, July 1999, 11(2), pp. 171-199.

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.

Rosenstone, Steven J. and Hansen, John M. *Mobilization, Participation and Democracy in America.* New York: MacMillan, 1993.

Tax Foundation, Inc. *Facts and Figures on Government Finance.* Washington D.C., 1940-2000.

United States Department of Commerce. *Current Population Survey.* Washington, D.C., 1940-2000.

Wildasin, David E. "Income Redistribution in a Common Labor Market." *American Economic Review,* September 1991, 81(4), pp. 757-774.