

# Fiscal Federalism and Optimal Income Taxes

*Maximiliano Dvorkin*

This paper studies how local policies—specifically, taxes on income with redistributive goals—affect the migration decisions of individuals and, in turn, how these migration decisions affect local and economy-wide tax and redistribution policies. The author develops a model of optimal taxation for a federal system of governments in the tradition of Mirrlees (1971), where taxes can be fully nonlinear but informational asymmetries prevent the equalization of well being across workers due to informational rents. This article extends the large literature on federalism and tax competition by obtaining optimal tax formulas for the federal and state governments. The literature has mainly focused on inefficiencies that arise due to fiscal externalities when governments have access to restricted instruments (for example, allowing only linear taxes). Contrary to previous results in that literature, the author shows here that state governments will provide redistribution through taxes and that, in a symmetric equilibrium, the overall tax schedule that combines the actions of both the federal and state governments is the same as that of a unitary government. This implies that, under the conditions analyzed in the model, there is no reason to restrict income redistribution objectives to the federal government only, as commonly prescribed in the literature. (JEL H2, H71, H77, R23)

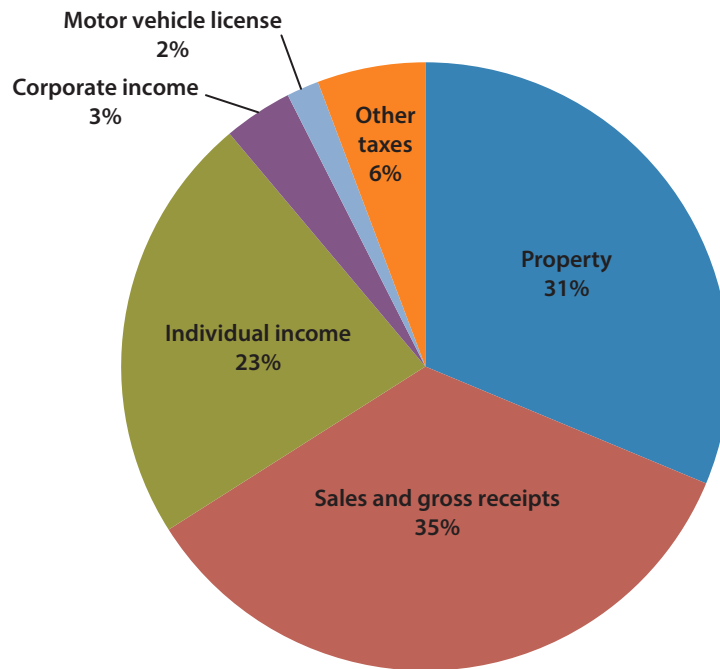
Federal Reserve Bank of St. Louis *Review*, Fourth Quarter 2017, 99(4), pp. 319-36.  
<https://doi.org/10.20955/r.2017.319-336>

## INTRODUCTION

In federal countries, such as the United States, state and local governments play an important role. They provide many different types of local public goods and services and are able to tailor the amount and quality of these goods to the preferences of their constituencies. To provide these services, state and local governments must raise tax revenues from their residents using a menu of different tax instruments. Among these public goods and tax instruments are the provision of social insurance and progressive income taxation, which will result in a redistribution of income. However, this raises important questions. Should local governments redistribute income through taxes and subsidies? Does local tax competition result in an inefficient amount of redistribution? A large literature in public finance discusses taxation

Maximiliano Dvorkin is an economist at the Federal Reserve Bank of St. Louis.

© 2017, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

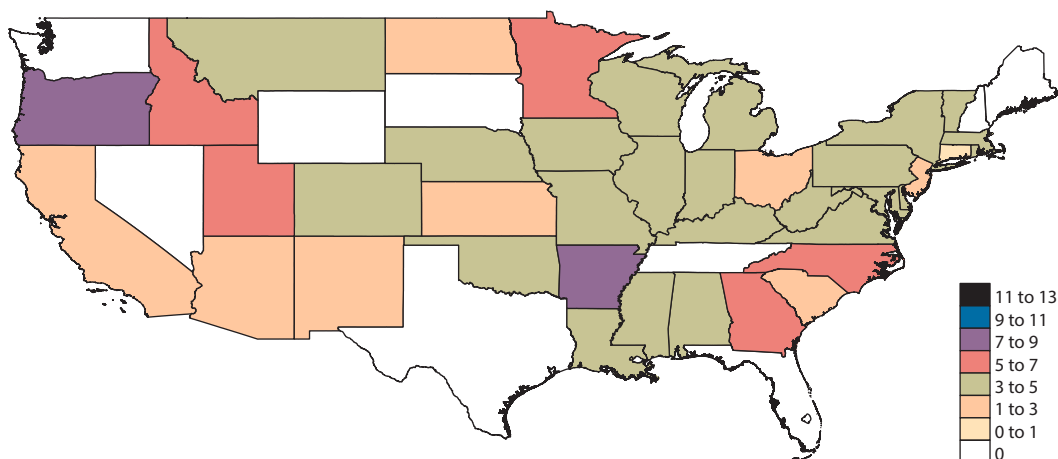
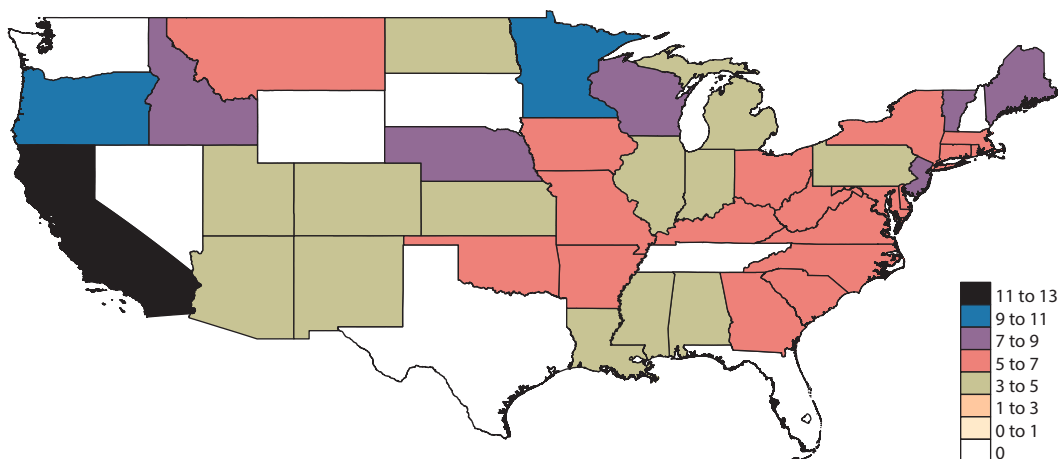
**Figure 1****Share of Different Taxes in State and Local Government Tax Revenues**

SOURCE: Author's calculation using data from the U.S. Census Bureau 2014 Annual Surveys of State and Local Government Finances.

and provision of public goods in economies with multiple and decentralized levels of governments, also known as a federal system.<sup>1</sup> The findings in that literature are that local governments will engage in tax competition, trying to poach workers from other regions by promising lower taxes. This “race to the bottom” will result in an insufficient level of taxation and redistribution. Thus, the literature prescribes that all redistributive objectives and instruments should belong solely to the federal government.

In reality, we observe an important amount of income taxation by state and local governments. For example, in the year 2014, state and local governments collected \$340 billion in income tax revenues. As shown in Figure 1, this represented 23 percent of all tax revenues of state and local governments, being the third largest source of tax income, after sales and property taxes.

At the same time, there are ample differences across states in terms of their income taxation and progressivity. A few states do not tax income at all. In other cases, states tax all incomes at the same rate, referred to as a flat tax. However, in most states, the marginal tax rate on income changes with the level of income. Figures 2 and 3 show an example of this heterogeneity in tax rates across U.S. states.<sup>2</sup> Figure 2 shows the state marginal tax rate on earnings for a single filer with a low yearly income of \$15,000. Similarly, Figure 3 shows the marginal

**Figure 2****Marginal State Tax Rate on Earnings: Single Filer with Income of \$15,000 (2014)**SOURCE: Author's calculation using the NBER TAXSIM database; <http://www.nber.org/taxsim/>.**Figure 3****Marginal State Tax Rate on Earnings: Single Filer with Income of \$1,000,000 (2014)**SOURCE: Author's calculation using the NBER TAXSIM database; <http://www.nber.org/taxsim/>.

tax rate for a single filer with a high income of \$1,000,000. In the first case, the marginal tax rate varies across states, from zero to approximately 6 percent. In the second, it varies from zero to 13 percent. These examples highlight the active role of U.S. states in providing redistribution and social insurance using progressive income taxation.

In this article I study the taxation and redistribution problem of a federal country from a normative perspective. I develop a model of optimal taxation for a federal system of governments in the tradition of Mirrlees (1971), where taxes can be fully nonlinear but informational asymmetries prevent the equalization of well being across workers due to informational rents. Here, I extend the large literature on federalism and tax competition by obtaining optimal tax formulas for the federal and state governments. The literature has mainly focused on inefficiencies that arise due to fiscal externalities when governments have access to restricted instruments, that is, linear taxes. Contrary to previous results, I show here that state governments will provide redistribution through income taxes and that, in a symmetric equilibrium, the overall tax schedule that combines the actions of both the federal and state governments is the same as that of a unitary government. This finding implies that, under the conditions analyzed in this problem, there is no reason to restrict income redistribution objectives to the federal government.

This article is closely related to Gordon and Cullen (2012). I extend their results in three important ways. First, I model explicitly the migration of workers as an extensive-margin decision. Gordon and Cullen (2012) use measures of migration elasticities but do not provide micro-foundations for this elasticity and cannot be linked to the underlying preference or technology parameters. Taxes affect the migration decisions of workers, thus migration elasticities will be affected by taxes. In this work, migration and the resulting tax formulas are functions of the deep parameters of the model. Second, I follow the literature on mechanism design and model the problem as a game of asymmetric information in the tradition of Mirrlees (1971), where governments must provide incentives for workers to work the “right” amount of hours, whereas Gordon and Cullen (2012) use the perturbation approach of Saez (2001). Last, I show that tax competition is more intense the larger is the number of states.

On the technical side, I follow the approach of Diamond (1998) and interpret tax formulas in a similar way. In addition, the effects of combining intensive- and extensive-margin decisions on the labor supply has been studied by Kleven et al. (2009) and my article connects to this work. However, they analyze the case of a single level of government and therefore abstract from tax competition. Introducing many governments that compete to attract workers (in the case of the states) and overlapping tax bases (in the case of the state with the federal government) pushes the model into the literature of mechanism design with many principals, which is known as “common agency” (Bernheim and Whinston, 1986). A well-known result in this literature is that the revelation principle, a tool of utmost importance in mechanism design and principal-agent problems, may not be used to solve the problem (Peters, 2001, and Martimort and Stole, 2002). I discuss the conditions in the model that allow me to use the revelation principle to characterize the optimal tax schedules.

In addition, my model connects to the work by Morelli et al. (2012), which focuses mainly on state taxation and whether citizens prefer independent taxing powers for each state, with states competing to attract highly productive types, or a unified taxing schedule for the nation. They identify this latter case as a federal taxing system. Mirrlees (1982) and many other works on international tax competition are interested in optimal taxation and migration between nations. In this case, the focus is on competition between principals whose policies do not overlap.

This article is organized as follows. I introduce the notation and main assumptions in Section 2. I then review optimal tax formulas for the standard model with a single government: the unitary economy. This section will allow me to compare whether and how the tax formulas in the federal economy differ from the standard case. Section 4 discusses the case of federal and state governments but without migration. The focus in Section 4 is to understand inefficiencies that may arise when two governments engage in redistribution for the same workers. Last, I extend the model to allow for migration and show the effects of tax competition and then conclude. For ease of exposition, all proofs are omitted, but I can provide them upon request.

## MAIN ASSUMPTIONS

The economy consists of a single country. There are two levels of government in the country. The first is a single federal government whose actions and policies affect all territories in the country. The second level consists of many local governments, with no geographic overlap, whose actions and policies affect only their own territory. I will interpret these local governments as states.<sup>3</sup> The economy is populated by an infinite number of workers of measure 1. Workers have different productivities, which are denoted by  $\theta$ . The distribution of productivities  $\theta$  in the population is  $F(\theta)$ . The total number of efficiency units of labor a worker supplies is equal to the total hours she chooses to work times her productivity. Market wages per efficiency unit of labor are assumed exogenous and equal to 1.

Workers derive utility from consumption and disutility from work according to the following function:

$$U(c, l) = c - \theta v(l),$$

where  $U(\cdot)$  is total utility,  $c$  is consumption,  $l$  is hours worked (labor supply), and  $v(\cdot)$  is a differentiable and convex function. It is convenient to rewrite workers' utility depending on their earnings  $y = \theta l$ :

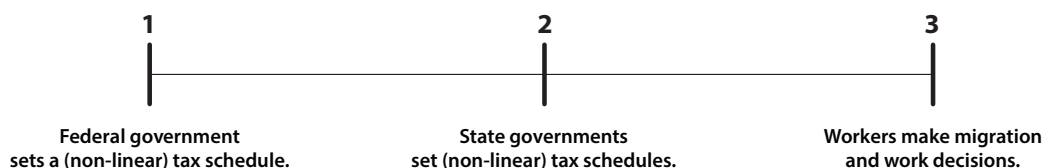
$$(1) \quad U(c, y) = c - \theta v(y/\theta).$$

The economy runs for only one period and therefore workers do not save. The budget constraint of a worker living in state  $i$  is

$$(2) \quad c = y - T_F(y) - T_{S_i}(y),$$

where  $T_F(\cdot)$  represents the tax policy (function) of the federal government and  $T_{S_i}(\cdot)$  is the tax policy of the state government  $i$ .

Workers have preferences over different regions and may decide to migrate. These preferences are represented by a vector  $x$ , in which its  $i$ th element is the extra utility the worker gets if she chooses to live in state  $i$ . This extra utility is additively separable from  $U(\cdot)$ . A worker's moving decision is influenced both by her preferences and by economic conditions in the regions in terms of net income.

**Figure 4****Timing of Events in the Model**

Both levels of government have access to the same tax and transfer instruments, which can be fully nonlinear in income. As in the large literature on optimal taxation starting with Mirrlees (1971), I assume that governments can observe workers' income and consumption, but workers' productivity and labor supply decision are private information known only to the workers. Moreover, governments cannot observe the mobility costs or workers' preferences for locations and therefore cannot condition their tax policies on them. This assumption is a main point of departure of this article from previous literature. As in Rochet and Stole (2002), I restrict the tax schedules set by the different governments to be nonrandom.

Governments have preferences for redistribution represented by social weights  $G_F(\theta)$  for the federal government and  $G_S(\theta)$  for the state governments. In other words, governments have a weighted utilitarian social-welfare function. The assumption is that both  $G$ s are distribution functions and  $F$  is first-order stochastic dominant over the  $G$ s; that is, the utility of workers with lower  $\theta$  accounts more for social welfare than the utility of workers with high  $\theta$ .

In this work, I assume that there are many states and each is small in the sense that its decisions have a negligible impact on the revenues of the federal government and the rest of the states. None the less, states understand that workers are not perfectly mobile across regions and thus have some "market power" over workers.<sup>4</sup>

Timing is as follows. First, the federal government sets a nonlinear income tax schedule for the whole country. In principle, this schedule need not be identical for all territories, and workers in different states may pay different taxes to the federal government. Second, state governments decide the nonlinear tax schedule that will affect the workers that choose to live and work in their territory. Last, workers observe the different tax policies for the federal government and all states and decide where to live and how much to work.<sup>5</sup> A summary of the timing is shown in Figure 4.

Note that the combination of quasilinear utility and exogenous social weights substantially simplifies the problem and will allow me to derive many closed-form expressions.

## OPTIMAL TAXES AND REDISTRIBUTION WITH A SINGLE GOVERNMENT

The case of a single level of government, that is, a unitary government or many fully coordinated governments, serves as a useful benchmark, as it allows me to compare whether

there are gains or losses from having many uncoordinated governments pursuing redistributive objectives. In this section, I present the formulas that implicitly define the optimal tax rates. Since in this case mobility is inconsequential, the results are almost identical to Diamond (1998) and Salanié (2005).

### ***The Problem of the Worker***

As is usual in the literature, I set up the problem as a principal-agent game, where the government is the principal and the worker is the agent. Let  $T_U(\cdot)$  be the tax schedule for the unitary government. Given this tax schedule, and applying the revelation principle, the problem of the worker is to communicate a productivity level  $\tilde{\theta}$  to the government in order to maximize her utility:

$$\begin{aligned} \max_{\tilde{\theta}} U(\tilde{\theta}, \theta) &= c(\tilde{\theta}) - v(y(\tilde{\theta})/\theta) \\ \text{s.t.} \quad c(\tilde{\theta}) &= y(\tilde{\theta}) - T_U(y(\tilde{\theta})). \end{aligned}$$

And the optimality condition for the agent's problem is

$$(3) \quad v'(y(\tilde{\theta})/\theta) = \theta(1 - T'_U(y(\tilde{\theta}))).$$

This condition will be sufficient for a global maximum if the second-order condition of the agent holds as well. It can be shown that if  $v$  is a strictly increasing and convex function, Equation (3) is necessary and sufficient for a global maximum if  $y(\cdot)$  is nondecreasing.<sup>6</sup>

If  $y'(\theta) > 0$  for all  $\theta$ , the equilibrium is fully separating. In case it does not hold everywhere and  $y'(\theta) = 0$  for some types, there will be an equilibrium with bunching.<sup>7</sup>

### ***The Problem of the Unitary Government***

The unitary government designs a contract specifying labor earnings and consumption for each productivity level that the worker communicates, to maximize social welfare.<sup>8</sup> The government needs to provide incentives for workers to reveal their true productivity level in the communication game:

$$\begin{aligned} (4) \quad \max_{y(\theta), c(\theta)} \quad & \int_{\underline{\theta}}^{\bar{\theta}} [c(\theta) - v(y(\theta)/\theta)] g_U(\theta) d\theta \\ \text{s.t.} \quad & \int_{\underline{\theta}}^{\bar{\theta}} T_U(y(\theta)) f(\theta) d\theta = 0 \\ & v'(y(\theta)/\theta) = \theta(1 - T'_U(y(\theta))) \\ & c(\theta) = y(\theta) - T_U(y(\theta)) \\ & y'(\theta) \geq 0, \end{aligned}$$

where  $g_U(\theta)$  is the weight the unitary government assigns to workers of productivity  $\theta$  in social welfare. The first equation in (4) is social welfare, the second is the government balanced budget constraint and requires that resources are not wasted. The third equation combined with the last equation form the (global) incentive-compatibility or truth-telling constraint, which requires that agents report their type truthfully to the government or principal.<sup>9</sup> The fourth equation is the workers' budget constraint.

The problem of the government can be summarized as

$$(5) \quad \max_{l(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta l(\theta) - T_U(\theta l(\theta)) - v(l(\theta)) - \frac{G_U(\theta) - F(\theta)}{f(\theta)} v'(l(\theta)) \frac{l(\theta)}{\theta} \right] f(\theta) d\theta.$$

And the optimal tax schedule is implicitly defined by

$$(6) \quad \frac{T'_U(\theta)}{(1 - T'_U(\theta))} = \left( 1 + \frac{1}{e} \right) \left( \frac{G_U(\theta) - F(\theta)}{1 - F(\theta)} \right) \frac{1 - F(\theta)}{\theta f(\theta)},$$

where  $e = \frac{v'(l)}{lv''(l)}$  is the labor supply elasticity. This equation resembles that of Diamond (1998), with the exception that I am using a weighted utilitarian criterion rather than a social-welfare function. The left-hand side of this equation states that the slope of the tax schedule or marginal tax rate,  $T'(\cdot)$ , for a productivity level  $\theta$  depends on three main factors: the labor supply elasticity, the (right) tail ratio of the distribution (or inverse hazard rate), and the redistribution objectives of the government.

When the labor supply is more responsive to changes in net wages (more elastic), taxes have a more distortionary effect. This implies that the marginal tax rate should be lower in this case. The effect of the tail of the distribution can be seen as follows. When we increase the marginal tax rate at productivity level  $\theta$ , we collect extra taxes on more productive individuals, who are  $(1 - F(\theta))$  in number. On the other hand, such a rate increase has a disincentive effect on the individuals whose productivity is close to  $\theta$ . The number of these individuals is proportional to  $f(\theta)$ , and the disincentive effect is more detrimental, for a given labor supply elasticity, when they are more productive. Finally, regarding the redistribution objectives of the government, if  $G(\theta) = F(\theta)$ , then the principal is a classic utilitarian and values all workers equally. In this case, the optimal marginal tax rate is zero and the government does not distort the work decision. If there are budget needs, the government should finance them only via lump-sum taxes. However, when  $F(\theta)$  is first-order stochastic dominant over  $G(\theta)$ , the government has a lower valuation of higher utilities (higher types) and wants to redistribute income. The marginal tax rate will be higher the higher this redistribution motive.

## A FEDERAL ECONOMY WITH NO MIGRATION

I start the analysis of a federal economy by assuming there is no migration of workers across states and derive the optimal tax schedules in the model. I write the problem ignoring the location preferences of workers since they are additively separable and therefore do not

distort the labor supply decision. Given the timing of the model, I require all actions to be subgame perfect and therefore solve for the equilibrium of the game by backward induction.

Because the worker (agent) now interacts with many governments (principals), the model falls in the literature of common agency (Bernheim and Whinston, 1986). The standard version of the revelation principle may be invalid in common agency models (Peters, 2001, and Martimort and Stole, 2002). However, given the sequential nature of the game and the fact that the agent must “contract” with both principals (intrinsic common agency) and report the same information to both, a version of the revelation principle holds.<sup>10</sup> The reason is that conditional on the tax schedule of the other government, the problem becomes exactly as that of a single agent and single principal, but the agent has a modified (net) income.

### ***The Problem of the Agent***

Given a tax schedule from the federal government  $T_F(y(\theta))$  and the states  $T_S(y(\theta))$ , the problem of the agent is to communicate a type  $\tilde{\theta}$  (the same to both governments) in order to maximize his utility:

$$\begin{aligned} \max_{\tilde{\theta}} u(\tilde{\theta}, \theta) &= c(\tilde{\theta}) - v(y(\tilde{\theta})/\theta) \\ \text{s.t.} \quad c(\tilde{\theta}) &= y(\tilde{\theta}) - T_F(y(\tilde{\theta})) - T_S(y(\tilde{\theta})). \end{aligned}$$

And the optimality of the agent’s problem is

$$(7) \quad v'(y(\tilde{\theta})/\theta) = \theta(1 - T'_F(y(\tilde{\theta})) - T'_S(y(\tilde{\theta}))).$$

This condition will be sufficient for a global maximum if the second-order condition of the agent holds as well, which will happen if  $y'(\theta) \geq 0$ . It is clear that equation (7) is very similar to equation (3), with the exception that now both tax schedules, federal and state, affect the intensive-margin decision of the worker.

### ***The Problem of the State***

As mentioned before, the federal government “moves first,” so the state takes the federal tax schedule ( $T_F(y(\theta))$ ) as given and designs the state tax schedule  $T_S(y(\theta))$  to maximize the welfare of its own citizens. The problem can be stated as

$$\begin{aligned} \max_{y(\theta), c(\theta)} \quad & \int_{\underline{\theta}}^{\bar{\theta}} [c(\theta) - v(y(\theta)\theta)] g_S(\theta) d\theta \\ \text{s.t.} \quad & \int_{\underline{\theta}}^{\bar{\theta}} T_S(y(\theta)) f(\theta) d\theta = 0 \\ & v'(y(\theta)/\theta) = \theta(1 - T'_F(y(\theta)) - T'_S(y(\theta))) \\ & y'(\theta) \geq 0. \end{aligned}$$

The problem of the state is fairly similar to the problem of the unitary government in (4) with one exception: The income (and marginal income) that the state can affect for redistribution is the income net of federal taxes. In other words, if we define  $\hat{y}(\theta) = y(\theta) - T_F(\theta)$ , then the problem of the state is the same as the standard optimal taxation model.

The problem of the state can be summarized as

$$(9) \quad \max_{l(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta l(\theta) - T_F(\theta l(\theta)) - v(l(\theta)) - \frac{G_S(\theta) - F(\theta)}{f(\theta)} v'(l(\theta)) \frac{l(\theta)}{\theta} \right] f(\theta) d\theta.$$

And the optimal state tax schedule is implicitly defined as

$$(10) \quad \frac{T'_S(\theta)}{(1 - T'_F(\theta) - T'_S(\theta))} = \left( 1 + \frac{1}{e} \right) \left( \frac{G_S(\theta) - F(\theta)}{1 - F(\theta)} \right) \frac{1 - F(\theta)}{\theta f(\theta)},$$

where  $e = \frac{v'(l)}{lv''(l)}$  is the labor supply elasticity.

Equation (10) is almost identical to equation (6). However, the federal tax schedule affects the optimal state tax schedule, which shows the interaction between both principals. Given the sequential timing assumption, the state can adapt its policies to what the federal government is doing. Since all terms on the right-hand side are positive and the denominator on the left-hand side cannot be negative, marginal state taxes are everywhere positive.<sup>11</sup>

Note also that if  $T'_F(\theta) > 0$ , then marginal state taxes are lower than the case of  $T'_F(\theta) = 0$ . The intuition is that, with positive federal marginal taxes, it is more costly for states to provide redistribution because workers with high productivity cannot keep all of the informational rents the state provides. Other than that, the interpretation of this formula is similar to the previous case and therefore omitted.

### The Problem of the Federal Government

The problem of the federal government is to maximize welfare:

$$(11) \quad \begin{aligned} & \max_{y(\theta), c(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} [c(\theta) - v(y(\theta)\theta)] g_F(\theta) d\theta \\ & \text{s.t.} \quad \int_{\underline{\theta}}^{\bar{\theta}} T_F(y(\theta)) f(\theta) d\theta = 0 \\ & \quad \{u(\theta), y(\theta)\} \text{ solve the state's problem conditional on } T_F \\ & \quad \{u(\theta), y(\theta)\} \text{ solve the agent's problem conditional on } T_F, T_S, \end{aligned}$$

where  $g_F(\theta)$  is the weight assigned by the federal government to agents with productivity  $\theta$ . The restrictions the federal government faces are a balanced budget and the optimal behavior from the state government and the agent.

The problem of the federal government can be summarized as

$$(12) \quad \max_{l(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \left[ U_S(\theta) - \frac{G_F(\theta) - G_S(\theta)}{f(\theta)} \right] f(\theta) d\theta,$$

where

$$(13) \quad U_s(\theta) = \theta l(\theta) - v(l(\theta)) - \frac{G(\theta) - F(\theta)}{f(\theta)} v'(l(\theta)) \frac{l(\theta)}{\theta},$$

and the optimal federal tax policy is implicitly determined as

$$(14) \quad \frac{T'_F(\theta)}{(1 - T'_F(\theta) - T'_S(\theta))} = \left(1 + \frac{1}{e}\right) \left( \frac{G_F(\theta) - G_S(\theta)}{1 - F(\theta)} \right) \frac{1 - F(\theta)}{\theta f(\theta)}.$$

In this case, equation (14) is once again very similar to that of the unitary government with one main exception: the redistribution objectives. The formula says that the tax schedule of the federal government will not depend solely on its own redistribution motives but those of the state as well.

One standard result in the literature of optimal taxation is the nonnegativity of marginal tax rates (Mirrlees, 1971). In the case of a federal country, equation (14) shows that it is possible for federal taxes to have a negative marginal tax rate if the redistribution objectives of the state are much stronger than those of the federal government.

As discussed in Keen (1998) and Keen and Kotsogiannis (2002), because both the federal and state governments share the same tax base—that is, the income of their citizens—their uncoordinated actions may generate a vertical fiscal externality. The intuition for this case can easily be understood in the counterfactual case in which the federal government sets the tax schedule to that of the unitary government. If both the federal and state governments could coordinate and commit, then this tax schedule would be optimal under the same redistribution objectives. However, after the federal government sets these taxes, incomes are not fully equalized for all workers. The state has the option not to tax, but additional taxes on high-income individuals that are transferred to low-income individuals have a first-order benefit in social welfare and small costs in terms of disincentives to work, so the state will decide to implement a tax. However, the state government does not internalize the tax's first-order costs in terms of lost revenues for the federal government, as high-income workers now work less. Because of this effect, and in the case analyzed here of no migration of individuals, if both the federal and all state governments have the same social weights, then the optimal marginal tax schedule for the federal government is  $T'_F = 0$ .

The intuition is simple. When the objectives of the state and the nation are perfectly aligned, there is no conflict between governments; therefore, the federal government prefers the state to provide all the redistribution and avoids the costs of the vertical fiscal externality. In case some federal taxes need to be raised to finance federal public goods, these taxes must be lump sum.

More generally when  $g_F \neq g_S$ , the federal government chooses a policy such that, when combined with the state's policy, the objective of the federal government is maximized. That is, when  $g_F \neq g_S$ , the *aggregate* tax schedule satisfies

$$(15) \quad \frac{T'_S(\theta) + T'_F(\theta)}{1 - T'_S(\theta) - T'_F(\theta)} = \left(1 + \frac{1}{e}\right) \frac{G_F(\theta) - F(\theta)}{1 - F(\theta)} \frac{1 - F(\theta)}{\theta f(\theta)}.$$

Equation (23) resembles that of the unitary government in equation (6). Therefore, under the assumption of no mobility of agents across states and identical distribution of types in the states, the existence of more than one tax authority does not create an inefficiency relative to the unitary government. Moreover, in the non-symmetric case, if the federal government can tax individuals in different states in a different way, there is also no efficiency cost and the second best can be achieved.

It is worth highlighting the important differences between the case of fully nonlinear taxes as analyzed here and the case of restricted tax instruments—for example, linear taxes. In the latter, a vertical fiscal externality is unavoidable unless lump-sum transfers are available as an instrument for the federal government. In other words, the results presented in this section suggest that the vertical fiscal externality analyzed in the fiscal federalism literature is the result of restricted tax instruments and do not necessarily arise in more general settings.

## ALLOWING FOR MIGRATION BETWEEN STATES

Now I allow agents to move from their initial state to a different one in response to tax incentives and preferences. As in Piaser (2007) and Morelli et al. (2012), there is an additively separable location preference term that affects the utility of living in a state. The difference with these previous works is that this term can be negative, which I interpret as a disutility of living in that region. The consequences of these idiosyncratic preferences for the states are that individuals may migrate to another state even if the tax schedule is less favorable than the state in which they are currently living.

Piaser (2007) and Morelli et al. (2012) are based on the models of simultaneous common agency developed by Rochet and Stole (2002), which I also follow here. The main modeling difference is that I add an additional principal, the federal government, which would choose the tax policy in a previous stage.

Recall that individuals first choose whether to migrate or not (an extensive-margin decision) and then how much to work and consume (an intensive-margin decision). Because the model is static, individuals can migrate only once. For the simultaneous choice of tax policies for the states, I restrict the problem to deterministic mechanisms. With no overlapping state principals and no direct externalities between states due to exclusivity (individuals cannot live in more than one state at the same time), as in Rochet and Stole (2002), the revelation principle can still be applied to the problem between the state and the agent. As before, the revelation principle will also hold for the problem between the nation and the agent.

### *The Problem of the Agent*

The problem of the agent now involves a location decision. Let's call  $x_i$  the additive preference term for living in state  $i$  for a worker of type  $\theta$  whose initial state is  $j$ .  $x_i$  is independent of the state the agent is initially living in and independent of the agent's type. For convenience, I assume that  $x_i$  is distributed type I extreme value. This distributional assumption is standard in the literature of discrete choice models with random utility and has recently been used by Kennan and Walker (2011) to model migration decisions of workers across states.

For the worker, the utility that a worker that initially is living in state  $i$  will get if she moves to state  $j$  is

$$(16) \quad u(\theta, \tilde{\theta}, i, j) = c_j(\tilde{\theta}) - v(y_j(\tilde{\theta})/\theta) + \sigma x_j - M(i, j),$$

where  $\sigma$  is a positive constant,  $M(i, j)$  is the cost of moving from state  $i$  to state  $j$ , and  $M(i, i) = 0$ . Note that, before making the migration decision, the worker observes the full vector of preferences  $x$ . The worker will choose to live in state  $j$  if

$$(17) \quad u(\theta, \tilde{\theta}, i, j) \geq u(\theta, \tilde{\theta}, i, k); \quad \forall k \neq j.$$

After the migration decision has taken place and because of the additive separability of the preference for the state, the problem of the worker is to communicate her type to the federal and state governments. The problem of the worker is similar to the previous case, and the optimality condition of the agent is the same as in equation (7).

The assumption that each element of  $x$  is i.i.d. (standardized) type I extreme value is useful to get a closed-form expression for problem (17). The share of workers in state  $i$  that will move to state  $j$  is<sup>12</sup>

$$(18) \quad S(\theta, i, j) = \frac{\exp\left(\left(c_j(\tilde{\theta}) - v(y_j(\tilde{\theta})/\theta) - M(i, j)\right)/\sigma\right)}{\sum_k \exp\left(\left(c_k(\tilde{\theta}) - v(y_k(\tilde{\theta})/\theta) - M(i, k)\right)/\sigma\right)}.$$

### The Problem of the States

In the standard optimal tax design problem, the government faces the classic trade-off between redistribution and providing incentives to work. With mobility, this trade-off becomes more intense: A generous redistribution will attract workers with low productivity to the state, reducing the amount each of these workers would get if migration were not allowed, and high marginal taxes not only will induce workers to work less, but also to move out, shrinking the tax base and the state budget.

Let  $N$  be the total number of states and  $\gamma_i(\theta) = \sum_j S(\theta, j, i) \frac{f_j(\theta)}{f_i(\theta)}$ . This equation represents the fraction of all workers of type  $\theta$  that will decide to live in state  $i$  relative to the initial measure in that state.

The problem of the state can then be written as follows:

$$(19) \quad \begin{aligned} \max_{y(\theta), c(\theta)} \quad & \int_{\underline{\theta}}^{\bar{\theta}} u(\theta) \gamma_i(\theta) g_{s_i}(\theta) d\theta \\ \text{s.t.} \quad & \int_{\underline{\theta}}^{\bar{\theta}} T_s(y(\theta)) \gamma_i(\theta) f_i(\theta) d\theta = 0 \\ & v'(y(\theta)/\theta) = \theta(1 - T'_F(y(\theta)) - T'_S(y(\theta))) \\ & y'(\theta) \geq 0, \end{aligned}$$

where  $\gamma(\theta)$  and  $S(\theta, i, j)$  are defined as before. This problem is similar to the case without mobility with the exception of the factor  $\gamma_i(\theta)$ .

Note that under some restrictive conditions, (i)  $f_i(\theta) = f_j(\theta)$  for all  $j$ ; (ii) mobility costs  $M(i, j)$  are the same for all  $i \neq j$ ; and (iii) in a symmetric equilibrium, we have that  $\gamma_i(\theta) = 1$  for all  $\theta$  in equilibrium. In this case, the optimal state tax schedule is implicitly defined as

$$(20) \quad \frac{T_{s'}(\theta)}{1 - T_{F'}(\theta) - T_{s'}(\theta)} = \left(1 + \frac{1}{e(\theta)}\right) \frac{1 - F(\theta)}{\theta f(\theta)} \left[ \frac{G(\theta) - F(\theta)}{(1 - F(\theta))} \right. \\ \left. - \frac{N-1}{\sigma N \lambda (1 - F(\theta))} \int_{\theta}^{\bar{\theta}} [u(x) - E_{g_s}(u)] g_s(x) dx \right. \\ \left. - \frac{N-1}{\sigma N (1 - F(\theta))} \int_{\theta}^{\bar{\theta}} T_s(x) f(x) dx \right],$$

where  $\lambda = 1 + \left(\frac{N-1}{\sigma N}\right) E_{g_s}[u(\theta)]$  and  $e(\theta) = \frac{v'(l(\theta))}{l(\theta)v''(l(\theta))}$  is the labor supply elasticity.

The main difference between state taxes in an economy with migration (equation (20)) and one without migration (equation (10)) are the last two terms inside the brackets in (20). These new terms reflect the effects of tax competition in a world where migration is possible. Since both  $T_s$  and  $u$  are increasing functions in a fully separating equilibrium, the last two terms inside the square brackets are positive, and the marginal tax schedule of the state is lower than in the case without mobility.

The intuition for this result is that now states face extra costs of redistribution. The term  $\frac{G-F}{1-F}$  implies that the government would like to redistribute away from higher types. However, since taxes affect workers with productivity  $\theta$  or higher, some of them will react to higher taxes by migrating to another state. This effect has a costs in terms of utility (the second term in the brackets), as social welfare is reduced by the utility of workers that leave and tax revenues (the third terms in bracket), since taxes cannot be collected on workers that leave. Higher taxes will have an (increasing) effect on revenues and redistribution (higher social welfare) from the intensive-margin decision of workers already in the state, but these gains are second order. Clearly a reduction of tax revenues will have the opposite effect, providing incentives to states to lower taxes. This scenario captures the “race to the bottom” that is usually discussed in the literature of tax competition and shows the effects of the horizontal fiscal externality (Keen, 1998).

These effects depend on parameter  $\sigma$ , which captures how important location preferences are relative to economic conditions. Low values of  $\sigma$  imply a lower attachment to a state and more mobility. In this case, tax competition between states is stronger and marginal state taxes are low. On the other hand, high values of  $\sigma$  imply strong preferences for living in a state and mobility due to tax purposes is low. In the limit as  $\sigma$  goes to infinity, the state tax schedule approaches the case with no mobility.<sup>13</sup>

The number of states  $N$  also plays a role. With many states in an economy, the returns from tax competition are larger. The reason is that, in a symmetric equilibrium, a single state

has  $1/N$  workers. By lowering taxes, the state can attract workers from all other  $N - 1$  states and the response at the extensive margin is larger the larger is the mass of workers at risk of immigrating to the state.

### The Problem of the Federal Government

Tax competition induces a race to the bottom, which affects the ability of local governments to provide an adequate level of redistribution. The question is to what extent a federal government can alleviate this issue. I will focus on the case discussed before, where the three restrictive conditions are met:  $f_i(\theta) = f_j(\theta)$  for all  $j$ ,  $M(i, j)$  is the same for all  $i \neq j$ , and the equilibrium is symmetric. These conditions imply that the federal government anticipates a symmetric equilibrium in the state policies as described before. The problem of the federal government is

$$\begin{aligned}
 (21) \quad & \max_{y(\theta), c(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} u(\theta) g_F(\theta) d\theta \\
 & s.t. \int_{\underline{\theta}}^{\bar{\theta}} T_F(y(\theta)) f(\theta) d\theta = 0 \\
 & \{u(\theta), l(\theta)\} \text{ solve the state problem conditional on } T_F \\
 & \{u(\theta), l(\theta)\} \text{ solve the agent's problem conditional on } T_F, T_{S_i},
 \end{aligned}$$

where, as before,  $g_F(\theta)$  is the weight assigned by the federal government to agents with productivity  $\theta$ . The restrictions the federal government faces are a balanced budget and the optimal behavior from the state governments and the agent. In this case, the optimal tax schedule for the federal government is implicitly defined as

$$\begin{aligned}
 (22) \quad & \frac{T'_S(\theta)}{1 - T'_F(\theta) - T'_S(\theta)} = \left(1 + \frac{1}{e(\theta)}\right) \frac{1 - F(\theta)}{\theta f(\theta)} \left[ \frac{G_F(\theta) - G_S(\theta)}{(1 - F(\theta))} \right. \\
 & \quad + \frac{N - 1}{\sigma N \lambda (1 - F(\theta))} \int_{\underline{\theta}}^{\bar{\theta}} [u(x) - E_{g_s}(u)] g_S(x) dx \\
 & \quad \left. + \frac{N - 1}{\sigma N (1 - F(\theta))} \int_{\underline{\theta}}^{\bar{\theta}} T_S(x) f(x) dx \right],
 \end{aligned}$$

where  $\lambda = 1 + \left(\frac{N - 1}{\sigma N}\right) E_{g_s}[u(\theta)]$  and  $e(\theta) = \frac{v'(l(\theta))}{l(\theta)v''(l(\theta))}$  is the labor supply elasticity. Moreover, the aggregate tax schedule satisfies

$$(23) \quad \frac{T'_S(\theta) + T'_F(\theta)}{1 - T'_S(\theta) - T'_F(\theta)} = \left(1 + \frac{1}{e}\right) \frac{G_F(\theta) - F(\theta)}{1 - F(\theta)} \frac{1 - F(\theta)}{\theta f(\theta)}.$$

As discussed previously, the second and third terms inside the brackets are always positive, so even in the case of  $G_F = G_S$ , federal taxes will be positive. The proposition and lemma show that the federal government understands the underprovision of redistribution by the state governments due to fiscal competition and the race to the bottom and is able to undo it

and provide the “right” amount of redistribution. In the symmetric case, no inefficiencies arise due to the lack of coordination among different levels of governments.

In addition, these results show that marginal state taxes will not, in general, be zero and there is a role for states to provide redistribution. However, state tax competition becomes more intense as the number of states increases and therefore state redistribution and the progressivity of the state tax code is lower with a larger number of states.

## CONCLUSIONS

In this paper, I developed a model of optimal taxation for a federal system of governments in the tradition of Mirrlees (1971). I obtained optimal tax formulas for the federal and state governments in an economy with and without migration. My analysis extends a large literature on federalism and tax competition that has discussed horizontal and vertical fiscal externalities in the context of restricted tax instruments (mainly linear taxes). Contrary to previous results in that literature, I show here that state governments can provide redistribution through taxes and that, in a symmetric equilibrium, the overall tax schedule that combines the actions of both the federal and state governments is the same as that of a unitary government.

I extend the results of Gordon and Cullen (2012) to an economy with many governments and with a fully microfounded migration decision at the extensive margin. In this paper, I considered only a symmetric equilibrium with identical states. Natural extensions are to consider heterogeneous preferences in terms of the provision of local public goods and amenities, heterogeneity in the initial distribution of workers, and collective choices by direct voting and sorting. ■

## NOTES

- <sup>1</sup> Oates (1999) provides a survey of this literature.
- <sup>2</sup> To compute these examples I use the National Bureau of Economic Research (NBER) TAXSIM tax calculator (Feenberg and Coutts, 1993). In all cases I assume a worker with only wage income filing individually and no dependents. More information about TAXSIM can be found at <http://www.nber.org/taxsim/>.
- <sup>3</sup> Cities may be a third level of government, but I abstract from them in this work.
- <sup>4</sup> This set-up is analogous to the problem of a monopolistic competitor in standard microeconomic analysis of markets. Monopolistic competitors are small and have no impact on economy-wide aggregates, yet they face a downward-sloping demand for their products.
- <sup>5</sup> It is worth highlighting that the federal government acts as a Stackelberg leader in the game. However, since I am assuming that states are small and have no direct impact on the other states or the federal government's tax revenues, the timing assumption can be relaxed, with many of the results still holding.
- <sup>6</sup> In this case, local incentive compatibility (or truth-telling) constraints imply that global incentive constraints hold.
- <sup>7</sup> It is worth noting that we need only local conditions (first and second order), which hold under the Spence-Mirrlees single-crossing condition (by the assumptions in the utility function) and  $y'(\theta) > 0$ . This greatly simplifies the principal's problem. However, these conditions need not hold always and we should proceed with caution.
- <sup>8</sup> Specifying earnings and consumption also determines taxes for that productivity level.
- <sup>9</sup> This is an application of the revelation principle.
- <sup>10</sup> See the appendix in Martimort (2006).
- <sup>11</sup> The denominator cannot be negative by the optimality condition of the agent.
- <sup>12</sup> This is a well-established property of the extreme value distribution.
- <sup>13</sup> Note however that in this case there will be mobility due purely to preferences. This effect will be so strong that state tax policies will have no role in the location decisions when  $\sigma$  is very high.

## REFERENCES

- Bernheim, B. Douglas and Whinston, Michael D. "Common Agency." *Econometrica*, 1986, 54(4), pp. 923-42; <https://doi.org/10.2307/1912844>.
- 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 and Coutts, Elisabeth. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management*, 1993, 12(1), pp. 189-94; <https://doi.org/10.2307/3325474>.
- Gordon, Roger H. and Cullen, Julie B. "Income Redistribution in a Federal System of Governments." *Journal of Public Economics*, 2012, 96(11), pp. 1100-09; <https://doi.org/10.1016/j.jpubeco.2011.07.001>.
- Keen, Michael. "Vertical Tax Externalities in the Theory of Fiscal Federalism." *International Monetary Fund Economic Review*, September 1998, 45(3), pp. 454-85; <https://doi.org/10.2307/3867412>.
- Keen, Michael and Kotsogiannis, Christos. "Does Federalism Lead to Excessively High Taxes?" *American Economic Review*, 2002, 92(1), pp. 363-70; <https://doi.org/10.1257/000282802760015784>.
- Kennan, John and Walker, James R. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica*, January 2011, 79(1), pp. 211-51; <https://doi.org/10.3982/ECTA4657>.
- Kleven, Henrik J.; Kreiner, Claus T. and Saez, Emmanuel. "The Optimal Income Taxation of Couples." *Econometrica*, 2009, 77(2), pp. 537-60; <https://doi.org/10.3982/ECTA7343>.
- Martimort, David. "Multi-contracting Mechanism Design," in Richard Blundell, Whitney K. Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications*. Volume 1. Cambridge: Cambridge University Press, 2006, pp. 57-101; <https://doi.org/10.1017/cbo9781139052269.004>.

- Martimort, David and Stole, Lars. "The Revelation and Delegation Principles in Common Agency Games." *Econometrica*, July 2002, 70(4), pp. 1659-73; <https://doi.org/10.1111/1468-0262.t01-1-00345>.
- Mirrlees, James A. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*, 1971, 38(2), pp. 175-208; <https://doi.org/10.2307/2296779>.
- Mirrlees, James A. "Migration and Optimal Income Taxes." *Journal of Public Economics*, 1982, 18(3), pp. 319-41.
- Morelli, Massimo; Yang, Huanxing and Ye, Lixin. "Competitive Nonlinear Taxation and Constitutional Choice." *American Economic Journal: Microeconomics*, 2012, 4(1), pp. 142-75; [https://doi.org/10.1016/0047-2727\(82\)90035-4](https://doi.org/10.1016/0047-2727(82)90035-4).
- Oates, Wallace E. "An Essay on Fiscal Federalism." *Journal of Economic Literature*, September 1999, 37(3), pp. 1120-49; <https://doi.org/10.1257/jel.37.3.1120>.
- Peters, Michael. "Common Agency and the Revelation Principle." *Econometrica*, September 2001, 69(5), pp. 1349-72; <https://doi.org/10.1111/1468-0262.00245>.
- Piaser, Gwenaél. "Labor Mobility and Income Tax Competition," in Colin Read and Greg N. Gregoriou, eds., *International Taxation Handbook: Policy, Practice, Standards and Regulations*. Burlington, MA: CIMA Publishing Elsevier, 2007, pp. 73-92.
- Rochet, Jean-Charles and Stole, Lars A. "Nonlinear Pricing with Random Participation." *Review of Economic Studies*, January 2002, 69(1), pp. 277-311; <https://doi.org/10.1111/1467-937X.00206>.
- Saez, Emmanuel. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies*, 2001, 68(1), pp. 205-29; <https://doi.org/10.1111/1467-937X.00166>.
- Salanié, Bernard. *The Economics of Contracts: A Primer*. Second Ed. Cambridge, MA: MIT Press, 2005.
- Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 1956, 64(5), pp. 416-24; <https://doi.org/10.1086/257839>.