Competitive Nonlinear Taxation and Constitutional Choice*

Massimo Morelli,† Huanxing Yang,‡ and Lixin Ye§

April 2008

Abstract

In an economy where agents have heterogeneous abilities and horizontally differentiated location preferences, we compare a unified nonlinear optimal taxation schedule with the equilibrium taxation schedules that would be chosen by two competing tax authorities if the same economy were divided into two States. This is the first analysis of competitive nonlinear income taxation, with unobserved heterogeneity in both vertical (ability) and horizontal (location preference) dimensions. We show that with competitive taxation the overall level of progressivity and redistribution is lower, and the welfare comparison is favorable to a unified system. If citizens could choose by majority rule (constitutional stage) whether to have unified taxation institutions or competitive tax authorities, the choice would depend on the initial distribution of abilities (incomes) and the intensity of location preferences. The rich should always be in favor of competing authorities and local governments, whereas the poor should always be in favor of unified taxation. The preferences of the middle class depend on the initial conditions.

Keywords: Competitive nonlinear taxation, Location preferences, Inequality, Integration, Type preferences over institutions.

*We thank Stefania Albanesi and Bernard Salanie for very useful comments and suggestions on our preliminary notes and results. The usual disclaimer applies.

†Dept. of Political Science and Dept. of Economics, Columbia University. mm3331@columbia.edu.

‡Dept. of Economics, The Ohio State University. huanxing@econ.ohio-state.edu.

§Dept. of Economics, The Ohio State University. lixinye@econ.ohio-state.edu.

1 Introduction

Redistribution is one of the most common goals of Governments. It is self evident that there is an enormous variation in terms of the amount of redistribution pursued and achieved by different Governments.[1] As reported by Piketty and Saez (2006), such differences in progressivity levels have very significant impact on the long run distribution of accumulated wealths – even greater than the effect of World Wars through capital destruction.

Obviously the variation in progressivity of redistribution systems must be due to a combination of ideological and social differences, but they can also be related to (1) the institutional structure chosen at the constitutional stage and (2) the initial conditions at the time of the constitutional choice. This paper is exactly motivated by these considerations and aims to provide a clear analysis of the impact of key fiscal institutional choices (different under different initial conditions) on the difference in progressivity. In other words, the analysis will clarify the *effects* of different tax regimes on redistribution as well as the *endogenous determination* of those institutions themselves, as a function of crucial initial conditions in terms of inequality, relative class power, poverty, and location preferences.

The constitutional choice of which "taxation regime" to select (centralized versus decentralized, city taxes versus State taxes, European taxes versus national taxes etc.) may affect the location decision and distribution of disposable income of consumers and producers, and may in turn be affected by the perceived mobility and by the initial conditions in terms of relative power of the various classes. In the case of the European Union, the increased mobility of citizens and the recent expansion of the Union clearly have effects on the taxation systems of the various States, and in turn the new conditions in terms of distribution of incomes and classes affects the likelihood of further integration steps.

We are used to think that the level of progressivity of a tax system is mainly a *political choice*, reflecting the ideology and the preferences of the class(es) holding power. On the other hand, we are used to think of the choice "Federal versus State taxes," "City versus State taxes," or "property taxes versus centralized funding of schools" as mainly due to *efficiency or freedom to choose* considerations.

[1] Ireland and Sweden are very well known for the high degree of progressivity of their tax system and for the generous transfers, whereas Switzerland is at the opposite extreme. For the available evidence across the OECD countries, see reference [17].

This paper challenges this view, demonstrating that even if taxes are always chosen "optimally" on the basis of standard utilitarian criteria, a centralized taxation system leads to higher progressivity for any distribution of types and preferences.

We consider a competitive nonlinear taxation framework with both vertically and horizontally differentiated agents. More specifically, we consider a standard Salop circular model in which different States compete for different agents (citizens, workers, or consumers) along two dimensions. The vertical dimension captures the agents' heterogeneity in terms of their innate abilities or productivities. The horizontal dimension captures the agents' heterogeneity in terms of their location preferences, including their tastes for different cultures, landscapes, food, political systems, weather conditions, etc. We consider both the unified taxation system and the independent taxation system. Under the unified taxation system, the Federation's objective is to choose an optimal tax schedule to maximize the total utilities of all the citizens in the economy. Under the independent taxation system, each State's objective is to choose a tax schedule to maximize the total utilities of all the citizens choosing to live in the State, given all the other States' tax schedules.

In our base model we consider the case in which agents have three vertical types, type H (the rich), type M (the middle class), and type L (the poor). Under the independent authority regime, a taxation authority has to take into account not only the resource constraints and incentive compatibility constraints of a standard taxation designer, but also the additional individual rationality constraint derived from location preferences. In this independent taxation regime, the tax for high types is lower and the subsidy for low type is lower accordingly. Moreover, under the independent regime the total output and consumption are higher, but the total social welfare is lower, regardless of the preferences of the middle class. Intuitively, with competition each independent tax authority tries to attract more high type citizen-workers, since this increases a city's tax revenue to subsidize the low type. This competition effect reduces the tax to the high type, which means that the subsidy to the low type decreases accordingly.

The representatives of the interests of low productivity types (the poor) should always be in favor of a unified taxation regime. On the other hand, the representatives of the high productivity types (the rich) should prefer the independent regime. Hence the constitutional choice between the two regimes can always be thought of as determined by the preferences of the middle class (excluding the trivial cases in which one of the two extreme types has the absolute majority at the constitutional stage).

While it may be intuitive that, other things equal, the closer the ability of the middle type to that of the high type, the more likely that the middle type will prefer the independent taxation (and vice versa), the following relationship identified by our computations would have been very difficult to guess without the guidance of our model: the greater the percentage of the middle type in the total population, the more likely that the middle class will prefer independent taxation. The intuition has to do with two countervailing incentives for the middle type. The middle type's first incentive is to try to milk the rich, which makes the middle type prefer the unified regime. The middle type's second incentive is to avoid being milked by the poor, which makes the middle type prefer the independent regime. Other things equal, the higher the ability and/or the dimension of the middle class, the stronger the second effect and the weaker is the first. As a result, it is more likely that the interest of the middle type will align with that of the high type to prefer independent taxation.

In the case of a continuum of types no analytical result is feasible for the independent taxation regime, but the simulations have confirmed the same qualitative results of the three-type model. In particular, we demonstrate that the median type, responsible for constitutional choice, is more likely to choose the unified system if (1) the intensity of location preferences is lower, (2) the population of the poor or unproductive types is lower, or (3) the middle class is not too large.

One difference between our three-type model and the continuous-type model, is that in the continuum case the median type is generically different from the type who is indifferent between the various constitutional choices, and they vary at different rates when the parameters change. This has significant implications, for example, for the analysis of the impact of inequality on the constitutional choice.

Our results fit our intuition about the situation within the European Union, where things seemed at some point mature for a new European Constitution that would concentrate a larger fraction of policy decisions in Brussels, but such a preference for unification of policy making has reversed itself after the enlargement of the Union to include a set of poorer countries that have altered the distribution of income in the Union.

Literature Review

Since the seminar work of Mirrlees (1971), there is a growing literature on optimal taxation and competitive taxation. In the tax competition literature it is well established that capital tax competition leads to lower taxes and lower efficiency when tax revenue is used for public good provision, in contrast with the Tiebout hypothesis.[2] However, this debate in the literature is mostly about capital taxation and commodity taxation, but does not extend to the basic and fundamental case of purely redistributive governments. Moreover, a reasonable evaluation of the pros and cons of tax competition must take into account location preferences as well as the fact that productivity and abilities are not fully observable, and hence optimal taxes must be nonlinear. Our paper studies the trade-off between centralized or unified taxation and independent taxation in a framework in which the latter system has to involve a non-cooperative game between two mechanism designers facing a two-dimensional adverse selection problem.

Huber (1999) studies a two-country model with unobserved heterogeneity and hence non-linear taxes, but there are only two types and there is no labor mobility. The tax competition dimension is introduced by adding capital taxation with perfect mobility of capital. The main point of the paper is that capital tax affects the self-selection constraint of high type workers. There is no competition in income tax with mobile labor. Moreover, to be able to say something about progressivity of income tax and degree of inequality, it is important to have at least three types.

Oates (1977) and Ladd and Doolittle (1982) point to migration as the central issue in the normative evaluation of which level of government should undertake redistribution. The connection between mobility and redistribution is studied in Epple and Romer (1991), but their analysis of redistribution involves two exogenous communities competing for mobile tax payers of *observable* types. They take the existence of two communities as given, whereas we question the reason of their existence as separate taxation units. We enrich their framework in two main ways. First, we allow for unobservable heterogeneities, and taxes are decided competitively by the governments' designers. Secondly, taxation systems (unified versus State independent taxation) are endogenous.

The reason we prefer to use the optimal taxation approach is similar to the motivation in Gordon (1983): "Assuming that each government does in fact act in the best interest of it's own citizens, will

[2] See Wilson (1999) for a survey. The famous Tiebout hypothesis, in favor of independent policy-making with perfect mobility, was expressed in Tiebout (1956). A standard reference for the first opposing view is Oates (1977).

the collection of units of government together act in the best interests of all their citizens?" Gordon (1983) focuses, however, on the characterization of externalities of one State decisions over the other, but the taxes are commodity taxes and tax schedules are chosen independently. We focus on income taxation allowing instead for labor mobility and strategic interaction of taxation authorities.

Ticchi and Vindigni (2007) argue that countries with high inequality tend to select majoritarian constitutions whereas countries with lower inequality tend to select consensual democratic institutions. Since the latter institutions redistribute more, the relationship between inequality and redistribution is non existent or negative. Our framework allows for a nuanced look at the potential role of the initial conditions for the choice of taxation systems and hence for future redistributive policies.

A paper closely related to ours is Simula and Trannoy (2006), who study how allowing agent migration affects the optimal nonlinear income tax schedule in a Mirrleesian economy. The outside option in their model is assumed to be increasing in agents' types. They show that mobility significantly alters the closed-economy results, as a "curse" of the middle-skilled agents is identified: the marginal tax rate is negative at the top, and the average tax rate is decreasing near the top; consequently some middle-skilled individuals may end up paying higher tax than more productive individuals. Our analysis in the current paper is more general, as the "outside option" in Simula and Trannoy is endogenously determined in our competitive environment and we endogenize the choice of the taxation system. Interestingly, the curse of the middle class is absent in our three-type model, and in our continuous-type model as well when the location preference parameter is not too small. We only identify negative tax rates near the top in the continuous-type model when the intensity of location preferences is extremely small, and is quantitatively insignificant.

The paper is organized as follows. In section 2 we analyze the three-type model under both the unified and independent taxation regimes. Section 3 analyzes the case of continuum distribution of ability. Section 4 extends our analysis to the finite n-state case. Section 5 provides concluding remarks with some directions for future research and extensions.

2 Three-type Model

Consider two States in a potential Federation, since this is the minimal situation in which we can compare the progressivity of competitive State taxation versus that of a unified Federal tax.[3] Citizens (a continuum) differ in two dimensions. Along what we call the "horizontal" dimension, different citizens have different location preferences (tastes for culture, political system, language, friends, landscape, and weather conditions etc.) about the two States. We treat this horizontal differentiation by adopting a spatial framework. Specifically, citizens are uniformly distributed on a circle with unit length, and the two States are located at the two extreme points of a diameter of the circle. Let d_i denote the distance between a citizen on the circle and State i, $i = 1, 2$.[4] As a "vertical" dimension, citizens have different productivities or abilities θ. There are three (vertical) types of citizens (or workers): type H (the "rich"), type M (the middle type), and type L (the "poor"), with abilities θ_H, θ_M and θ_L, respectively ($\theta_H > \theta_M > \theta_L$). The corresponding proportions of three types are μ_H, μ_M and μ_L, respectively. We assume that the distribution of abilities is independent from location preferences. A θ-type individual working l units of time (or exerting an effort l) leads to production or gross income $Q = \theta l$.

Each State decides on a tax schedule $T(Q)$. Equivalently we assume that each State offers a menu of contracts, which is a collection of the consumption and production pairs with the form (C, Q), where $C = Q - T$. (A positive T will be called tax, and a negative one a subsidy). Given the menus of contracts offered by both States, citizens decide which contract to accept (or equivalently which State to work or stay with). Suppose an agent of type (θ, d_1) chooses to stay in State 1 and exerts an effort l, leading to gross income $Q = \theta l$ and consumption C. We assume that the utility for such an agent is given by

$$U(C, Q; \theta, d_1) = u(C) - l - kd_1 = u(C) - \frac{Q}{\theta} - kd_1$$

where $u(\cdot)$ is strictly increasing, strictly concave and twice continuously differentiable, and k, $k > 0$, is the intensity of the horizontal differentiation between the two States, or the weight of location preferences in the utility function. The smaller the k, the less horizontally differentiated the two

[3] Of course the analysis would apply unchanged to two cities whose provinces or counties together constitute a State, hence comparing the properties of centralized State level taxation versus decentralized city level taxation.

[4] Since $d_1 + d_2 = 1/2$, d_1 alone completely characterizes the agent's location type.

States, or the more intense the competition between the two States as people put less weight on their location preferences. In other words, k represents a cultural or personal cost for people to adjust to the life in one State, per unit of distance from the individual culture or personal knowledge of institutions.

Given a tax schedule $T(Q)$, the agent's problem is to choose Q to maximize his utility $u(Q - T(Q)) - Q/\theta - kd_1$. It is obvious that the single crossing property only holds along vertical dimension. The implication is that the tax authorities can only design tax schedules to sort agents along the vertical dimension.

It is well known that in the environment of competitive mechanism design, it is no longer without loss of generality to restrict attention to direct contracts.[5] To sidestep this problem, as in Rochet and Stole (2002) and Yang and Ye (2008) we restrict attention to deterministic contracts.[6] Since the preferences of a citizen with vertical type θ over the available consumption-production pairs conditional on staying with a State are independent of her horizontal type d_1, in what follows it is without loss of generality to consider direct contracts of the form $\{C(\theta), Q(\theta)\}_{\theta \in \{\theta_H, \theta_M, \theta_L\}}$. The tax amount incurred by type-θ citizen is then given by the tax function $T(\theta) = Q(\theta) - C(\theta)$. For brevity of exposition, from now on we will often refer to vertical types as the types, especially when there is no confusion in the context.

This basically completes a description of the model with independent taxation. For the model of unified taxation, our main goal is to lay down a benchmark with which we can identify the effect of competition on the taxation schedule. As such we assume that in the unified taxation case, all

[5] As demonstrated in a series of examples in Martimort and Stole (1997) and Peck (1997), equilibrium outcomes in indirect mechanisms may not be supported when sellers are restricted to using direct mechanisms where buyers report only their private types. Moreover, as demonstrated by Martimort and Stole (1997), an equilibrium in such direct mechanisms may not be robust to the possibility that sellers might deviate to more complicated mechanisms. The reason for such failures, as pointed out by McAfee (1993) and Katz (1991), is that in competition with nonlinear pricing the offers made by other firms may also be private information of the consumers when they make their purchase decisions, which means that this private information can also potentially be used when firms set up their revelation mechanisms.

[6] See Rochet and Stole (2002) for a discussion on the restrictions resulting from focusing on deterministic contracts. More general approaches in restoring the "without loss of generality" implication of the revelation principle in the environment of competitive nonlinear pricing have been proposed and developed by, for example, Epstein and Peters (1999), Peters (2001), and Page and Monteiro (2003).

the modeling elements are the same as in the independent taxation model, except that the two tax schedules are now designed by a Federal authority.[7] The objective of the Federal authority is to maximize the citizens' total utilities from the two States by choosing the menu of contracts for each State.

Note that in autarky $Q = C$, and the optimal consumption $C^*(\theta)$ is characterized by

$$u'(C^*) = 1/\theta \tag{1}$$

2.1 Unified taxation

Under unified taxation, the tax schedules are designed by a central authority, the Federation. The Federation's objective is to maximize the citizens' total utilities from the two States. Since citizens are uniformly distributed along the horizontal dimension, we focus on the symmetric solution in which each State offers the same menu of contracts and the resulting "market" shares are symmetric.[8] We can thus drop the State index to write $\{C_i(\theta), Q_i(\theta)\} = \{C(\theta), Q(\theta)\}$, $i = 1, 2$.

The Federation's objective is to set the pairs (C_H^U, Q_H^U), (C_M^U, Q_M^U) and (C_L^U, Q_L^U) to maximize the total social welfare

$$\max \mu_H \left[u(C_H) - \frac{Q_H}{\theta_H} \right] + \mu_M \left[u(C_M) - \frac{Q_M}{\theta_M} \right] + \mu_L \left[u(C_L) - \frac{Q_L}{\theta_L} \right]$$

subject to the binding resource constraint

$$\mu_H(Q_H - C_H) + \mu_M(Q_M - C_M) + \mu_L(Q_L - C_L) = 0 \quad \textit{(RC)}$$

The incentive compatibility (IC) conditions can be formulated as:

$$\begin{aligned} u(C_H) - \frac{Q_H}{\theta_H} &\geq u(C_M) - \frac{Q_M}{\theta_H} \quad \textit{(DIC-H)} \\ u(C_M) - \frac{Q_M}{\theta_M} &\geq u(C_L) - \frac{Q_L}{\theta_M} \quad \textit{(DIC-M)} \end{aligned}$$

The above DICs must be binding at the optimum. The IR conditions are not relevant, since everyone is required to participate. Let the multipliers of (DIC-H), (DIC-M) and (RC) be λ_H, λ_M

[7] So our benchmark is analogous to a multi-product monopoly in the IO literature.

[8] We focus on the symmetric solution here for ease of comparison with the independent case, where we will focus on symmetric equilibrium in which each state offers the same menu of contracts. While a formal proof is not attempted here, we conjecture that symmetric solution is optimal for the Federation.

and λ_R respectively. The first order conditions can be written as follows:

$$\begin{aligned}
\frac{\partial L}{\partial Q_H} = -\frac{\mu_H}{\theta_H} - \frac{\lambda_H}{\theta_H} + \mu_H \lambda_R &= 0 \\
\frac{\partial L}{Q_M} = -\frac{\mu_M}{\theta_M} + \frac{\lambda_H}{\theta_H} - \frac{\lambda_M}{\theta_M} + \mu_M \lambda_R &= 0 \\
\frac{\partial L}{\partial Q_L} = -\frac{\mu_L}{\theta_L} + \frac{\lambda_M}{\theta_M} + \mu_L \lambda_R &= 0 \\
\frac{\partial L}{\partial C_H} = \mu_H u'(C_H) + \lambda_H u'(C_H) - \mu_H \lambda_R &= 0 \\
\frac{\partial L}{\partial C_M} = \mu_M u'(C_M) - \lambda_H u'(C_M) + \lambda_M u'(C_M) - \mu_M \lambda_R &= 0 \\
\frac{\partial L}{\partial C_L} = \mu_L u'(C_L) - \lambda_M u'(C_L) - \mu_L \lambda_R &= 0
\end{aligned}$$

From the above equations, we have

$$\begin{aligned}
u'(C_H^U) &= \frac{1}{\theta_H} \\
\lambda_R &= \frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L} \\
\lambda_H &= \mu_H(\theta_H \lambda_R - 1); \;\; \lambda_M = \mu_L \theta_M (\frac{1}{\theta_L} - \lambda_R) \\
u'(C_L^U) &= \frac{\lambda_R}{(1 - \frac{\theta_M}{\theta_L}) + \theta_M \lambda_R}; \;\; u'(C_M^U) = \frac{\mu_M \lambda_R}{\mu_M - \lambda_H + \lambda_M}
\end{aligned}$$

It can be verified that $u'(C_M^U) > 1/\theta_M$ and $u'(C_L^U) > 1/\theta_L$ and $u'(C_M^U) < u'(C_L^U)$. Therefore, compared to the autarky case there is no distortion of consumption for type H, but the consumptions of type M and type L are both distorted downward (due to the concavity of $u(\cdot)$). Moreover, $C_H > C_M > C_L$.

Lemma 1 $T_H > T_M > T_L$.

Proof. Suppose $T_H \leq T_M$. That is, $Q_H - C_H \leq Q_M - C_M$. By the binding DIC-H,

$$\begin{aligned}
u(C_H) - u(C_M) = \frac{Q_H - Q_M}{\theta_H} &\leq \frac{C_H - C_M}{\theta_H} \\
\Rightarrow \;\; u(C_H) - \frac{C_H}{\theta_H} &\leq u(C_M) - \frac{C_M}{\theta_H}.
\end{aligned}$$

But this contradicts the fact that $C_H = \arg\max_C \{u(C) - \frac{C}{\theta_H}\}$ ($u'(C_H) = 1/\theta_H$) and $C_M < C_H$. Therefore, we must have $T_H > T_M$. Similarly, suppose $T_M \leq T_L$, that is, $Q_M - C_M \leq Q_L - C_L$. By the binding DIC-M,

$$\begin{aligned} u(C_M) - u(C_L) &= \frac{Q_M - Q_L}{\theta_M} \leq \frac{C_M - C_L}{\theta_M} \\ \Rightarrow \quad u(C_M) - \frac{C_M}{\theta_M} &\leq u(C_L) - \frac{C_L}{\theta_M}. \end{aligned}$$

By the properties of $u(C)$, the function $u(C) - \frac{C}{\theta_M}$ is strictly concave, which means that $u(C) - \frac{C}{\theta_M}$ is strictly increasing in C for $C \leq C_M^*$. Since $C_L < C_M < C_M^*$, we have $u(C_M) - \frac{C_M}{\theta_M} > u(C_L) - \frac{C_L}{\theta_M}$. A contradiction. Thus we must have $T_M > T_L$. ■

Given $T_H > T_M > T_L$, by RC we must have $T_H > 0$. This is because if $T_H \leq 0$, then by the lemma both T_M and T_L are strictly negative, and RC will be violated. Similarly, we must have $T_L < 0$. The sign of T_M is ambiguous and depends on parameter values.

2.2 Independent taxation

Under the independent taxation regime, each State chooses its taxation schedule simultaneously and independently. Our solution concept is Bertrand-Nash equilibrium: given the other State's tax schedule, each State chooses its own tax schedule to maximize its citizens' total utilities, where the set of citizens of each State is endogenously determined. We focus on symmetric equilibria in which both States choose the same taxation schedule.[9] Denote $v_j \equiv u(C_j) - Q_j/\theta_j$ to be the rent provision to type-$(\theta_j, 0)$ citizen, $j = H, M, L$. Suppose the other State's taxation rule leads to rent provisions v_j^*, $j = H, M, L$. Then the horizontal type x_j who is indifferent between the two States is determined by:

$$v_j(\theta) - kx_j = v_j^*(\theta) - k\left(\frac{1}{2} - x_j\right) \Rightarrow x_j = \frac{1}{4} + \frac{1}{2k}\left(v_j - v_j^*\right).$$

Actually, x_j is the type-θ_j "market share" for State 1: all the types (θ_j, d_1) with $d_1 < x_j$ choose to live in State 1 and those with $d_1 > x_j$ choose to live in State 2.[10]

[9] We have assumed a uniform distribution on the unit-length circle and that the fraction of the various ability types is constant across locations, hence the two States are identical if they split the circle equally (as they do in equilibrium). Symmetric equilibria would obviously be unjustified if the two States had some asymmetric initial conditions.

[10] More precisely, x_j is one half of the type-θ "market share" for State 1, given our circular model.

The State in question has the following programming problem:

$$\max \mu_H(x_H v_H - \frac{1}{2}kx_H^2) + \mu_M(x_M v_M - \frac{1}{2}kx_M^2) + \mu_L(x_L v_L - \frac{1}{2}kx_L^2)$$
$$x_H = \frac{1}{4} + \frac{1}{2k}(v_H - v_H^*);\; x_M = \frac{1}{4} + \frac{1}{2k}(v_M - v_M^*);x_L = \frac{1}{4} + \frac{1}{2k}(v_L - v_L^*)$$
$$u(C_H) - \frac{Q_H}{\theta_H} = u(C_M) - \frac{Q_M}{\theta_H};\; u(C_M) - \frac{Q_M}{\theta_M} = u(C_L) - \frac{Q_L}{\theta_M}$$
$$\mu_H x_H(Q_H - C_H) + \mu_M x_M(Q_M - C_M) + \mu_L x_L(Q_L - C_L) = 0.$$

Let λ_H and λ_M be the multiplier of DIC-H and DIC-M respectively, and let λ_R be the multiplier of RC. We first derive the first order conditions, then impose symmetry. To save notation, we define

$$\beta_H = \frac{1}{8} + \frac{v_H}{2k},\; \beta_M = \frac{1}{8} + \frac{v_M}{2k},\; \beta_L = \frac{1}{8} + \frac{v_L}{2k}$$

In the symmetric equilibrium, $v_i = v_i^*$, $i = H, M, L$. Thus the FOCs can be simplified into:

$$\begin{aligned}
-\frac{\mu_H}{\theta_H}\beta_H - \frac{\lambda_H}{\theta_H} + \mu_H\lambda_R\left[\frac{1}{4} - \frac{T_H}{2k\theta_H}\right] &= 0\\
-\frac{\mu_M}{\theta_M}\beta_M + \frac{\lambda_H}{\theta_H} - \frac{\lambda_M}{\theta_M} + \mu_M\lambda_R\left[\frac{1}{4} - \frac{T_M}{2k\theta_M}\right] &= 0\\
-\frac{\mu_L}{\theta_L}\beta_L + \frac{\lambda_M}{\theta_M} + \mu_L\lambda_R\left[\frac{1}{4} - \frac{T_L}{2k\theta_L}\right] &= 0\\
\mu_H\beta_H u'(C_H) + \lambda_H u'(C_H) + \mu_H\lambda_R\left[-\frac{1}{4} + T_H\frac{u'(C_H)}{2k}\right] &= 0\\
\mu_M\beta_M u'(C_M) - \lambda_H u'(C_M) + \lambda_M u'(C_M) + \mu_M\lambda_R\left[-\frac{1}{4} + T_M\frac{u'(C_M)}{2k}\right] &= 0\\
\mu_L\beta_L u'(C_L) - \lambda_M u'(C_L) + \mu_L\lambda_R\left[-\frac{1}{4} + T_L\frac{u'(C_L)}{2k}\right] &= 0
\end{aligned}$$

From the above equations, we obtain

$$\begin{aligned}
u'(C_H) &= 1/\theta_H,\\
\lambda_H &= \mu_H\left[-\beta_H + \theta_H\lambda_R\left(\frac{1}{4} - \frac{T_H}{2k\theta_H}\right)\right]\\
\lambda_M &= \mu_L\left[\frac{\theta_M}{\theta_L}\beta_L - \theta_M\lambda_R\left(\frac{1}{4} - \frac{T_L}{2k\theta_L}\right)\right]\\
\lambda_R &= \frac{\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M} + \mu_L\frac{\beta_L}{\theta_L}}{\frac{1}{4} - [T_H\frac{\mu_H}{2k\theta_H} + T_M\frac{\mu_M}{2k\theta_M} + T_L\frac{\mu_L}{2k\theta_L}]}\\
u'(C_M) &= \frac{\frac{1}{4}\mu_M\lambda_R}{\mu_M\beta_M - \lambda_H + \lambda_M + \mu_M\lambda_R\frac{Q_M - C_M}{2k}}\\
u'(C_L) &= \frac{\mu_L\lambda_R/4}{\mu_L\beta_L - \lambda_M + \mu_L\lambda_R T_L/(2k)}
\end{aligned}$$

Again, it can be verified that $u'(C_M) > 1/\theta_M$, $u'(C_L) > 1/\theta_L$, and $u'(C_L) > u'(C_M)$, hence $C_L^I < C_M^I < C_H^I$. Moreover, like in Lemma 1, we have $T_H^I > T_M^I > T_L^I$. As a result, $T_H^I > 0$, $T_L^I < 0$ and the sign of T_M^I is ambiguous.

From the FOCs, it is easy to verify that

$$u'(C_H^I) = \frac{1}{\theta_H}.$$

Thus $C_H^I = C_H^U$, and it involves no distortion.

LEMMA 2 *$u'(C_L^I) < u'(C_L^U)$, and $u'(C_M^I) < u'(C_M^U)$. Thus, $C_L^I > C_L^U$ and $C_M^I > C_M^U$.*

The proof is rather tedious, and is therefore relegated to the appendix.

2.3 Comparison

Even though the consumption level of high types is un-distorted, the same in both systems, the effects on Q's differ, and we obtain the following general result:

PROPOSITION 1 *$T_H^I < T_H^U$. That is, type H pays a lower tax in the independent State regime.*

Proof. By RC, DIC-H, and DIC-M, we have

$$\begin{aligned} Q_H &= \mu_H C_H + \mu_M C_M + (1-\mu_H)\theta_H[u(C_H) - u(C_M)] \\ &\quad + \mu_L[C_L + \theta_M(u(C_M) - u(C_L))] \end{aligned} \tag{2}$$

Define $\Delta Q_i = Q_i^I - Q_i^U$, $\Delta C_i = C_i^I - C_i^U$, and $\Delta u(C_i) = u(C_i^I) - u(C_i^U)$ where $i = H, M, L$. We have $\Delta C_H = 0$. In addition, $\Delta C_M > 0$ and $\Delta C_L > 0$ by Lemma 2. Then from (2), we have

$$\Delta Q_H = \mu_M \Delta C_M + \mu_L \Delta C_L - [(1-\mu_H)\theta_H - \mu_L\theta_M]\,\Delta u(C_M) - \mu_L\theta_M \Delta u(C_L)$$

By concavity of $u(\cdot)$, we have $\Delta u(C_i) \geq u'(C_i^I)\Delta C_i$. We thus have

$$\begin{aligned} \Delta Q_H &\leq \mu_M \Delta C_M + \mu_L \Delta C_L - [(1-\mu_H)\theta_H - \mu_L\theta_M]\,u'(C_M^I)\Delta C_M - \mu_L\theta_M u'(C_L^I)\Delta C_L \\ &< \mu_M \Delta C_M + \mu_L \Delta C_L - [(1-\mu_H)\theta_H - \mu_L\theta_M]\,\frac{1}{\theta_M}\Delta C_M - \mu_L\theta_M \frac{1}{\theta_L}\Delta C_L \\ &< (1-\mu_H)\frac{\theta_M - \theta_H}{\theta_M}\Delta C_M + \mu_L\left(\frac{\theta_L - \theta_M}{\theta_L}\right)\Delta C_L < 0 = \Delta C_H \end{aligned}$$

Thus $T_H^I < T_H^U$. ■

Proposition 2 $T_L^I > T_L^U$. *That is, L type gets less subsidy under independent taxation. Moreover,* $v_L^I < v_L^U$, *in spite of the higher consumption.*

Proof. Suppose $T_L^I \le T_L^U$. Then $Q_L^I - Q_L^U \le C_L^I - C_L^U$.

$$\begin{aligned} v_L^I - v_L^U &= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_L} \\ &\ge u(C_L^I) - u(C_L^U) - \frac{C_L^I - C_L^U}{\theta_L} \\ &= u'(C_L^*)(C_L^I - C_L^U) - \frac{C_L^I - C_L^U}{\theta_L} > 0. \end{aligned}$$

The first inequality is due to the fact that $Q_L^I - Q_L^U \le C_L^I - C_L^U$. The second equality follows from the intermediate value theorem, where $C_L^* \in (C_L^U, C_L^I)$. The last inequality holds since $u'(C_L^*) > u'(C_L^I) > 1/\theta_L$. Thus we have $v_L^I > v_L^U$. Next we compare v_M^I and v_M^U. By the binding DIC-M, we have

$$\begin{aligned} v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\ &= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\ &\ge u(C_L^I) - u(C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} \\ &= u'(C_L^*)(C_L^I - C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} > 0, \end{aligned}$$

where the last inequality follows from the fact that $u'(C_L^*) > u'(C_L^I) > 1/\theta_L > 1/\theta_M$. Thus $v_M^I > v_M^U$. From our results above, we have $C_H^I = C_H^U$, and $Q_H^I < Q_H^U$, hence $v_H^I > v_H^U$. However, the tax schedules under independent taxation $\{(Q_H^I, C_H^I), (Q_M^I, C_M^I), (Q_L^I, C_L^I)\}$ satisfy all the constraints under unified taxation, thus it is a feasible solution as well. However, the fact that $v_j^I > v_j^U$ for all $j = H, M, L$ contradicts the fact that the tax schedules $\{(Q_H^U, C_H^U), (Q_M^U, C_M^U), (Q_L^U, C_L^U)\}$ are the optimal solution for unified taxation. Therefore, we must have $0 > T_L^I > T_L^U$. Given that $T_L^I > T_L^U$ and $C_L^I > C_L^U$, we have $Q_L^I > Q_L^U$. Now suppose $v_L^I \ge v_L^U$. This implies that

$$\begin{aligned} v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\ &= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\ &> u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_L} \\ &= v_L^I - v_L^U \ge 0. \end{aligned}$$

Thus $v_j^I \geq v_j^U$ for all $j = H, M, L$ and $v_j^I > v_j^U$ for some j. But this again leads to a contradiction that $\{(Q_H^I, C_H^I), (Q_M^I, C_M^I), (Q_L^I, C_L^I)\}$ is feasible under unified taxation but the optimal solution is $\{(Q_H^U, C_H^U), (Q_M^U, C_M^U), (Q_L^U, C_L^U)\}$. Therefore, we must have $v_L^I < v_L^U$. ■

Since the tax schedules under independent taxation $\{(Q_H^I, C_H^I), (Q_M^I, C_M^I), (Q_L^I, C_L^I)\}$ are also feasible under unified taxation, we have the following immediate corollary:

COROLLARY 1 *Equilibrium welfare is always greater under the unified taxation than under the independent taxation regime.*

Even if the unified taxation system is welfare superior, it is clear that if the taxation system is chosen by majority rule at the constitutional stage, and if $\mu_i < 1/2$ for $i = H, L$, then the independent taxation regime can be chosen if and only if it yields higher equilibrium utility for the middle class. It is impossible to obtain general analytical results on the preferences of the middle type as a function of relative productivities (distribution of θs) and income distribution (distribution of μs). However, the computations we now turn to, provide interesting results.

2.4 Computational results

Our computations show that given $\theta_H = 2$ and $\theta_L = 1$, and given any percentage of each type, there exists a cutoff $\theta_M^* \in (1, 2)$ such that type M prefers the independent taxation system if and only if his type is higher than θ_M^*. Our computations also show that given $\theta_M \in (1, 2)$ and $\mu_H = \mu_L = (1 - \mu_M)/2$,[11] there exists a cutoff μ_M^* such that type M prefers the independent taxation regime if and only if $\mu_M > \mu_M^*$. Intuitively, as θ_M or μ_M increases, type M's interest aligns more with that of type H.

This result has an important implication in terms of welfare. Assume $\theta_H = 2, \theta_L = 1$ and $\theta_M = 1.51$. We can compute μ_M^* again by keeping $\mu_H = \mu_L$. We can then compare the utilitarian welfare of a federation with $\mu_M^* - \epsilon$ with that of a competitive taxation regime obtained with $\mu_M^* + \epsilon$. Even though the average θ is *higher* in the second case, welfare is higher in the former Federation, for ϵ small enough (by Corollary 1 above). This means that

Remark: A country with "better" initial conditions (higher productivity, or higher average θ here) may end up with lower welfare because of a suboptimal constitutional choice due to majority decision

[11] If μ_M increases, μ_H and μ_L decrease by the same amount.

making at the constitutional stage.[12]

The implications of poverty are very clear: Letting $k = 0.5$ and $\mu_M = 0.4$, Figure 1 shows that the cutoff of middle type ability is decreasing in the percentage of type L. As the population of the low type increases, the middle type becomes more eager to prefer independent tax regime.

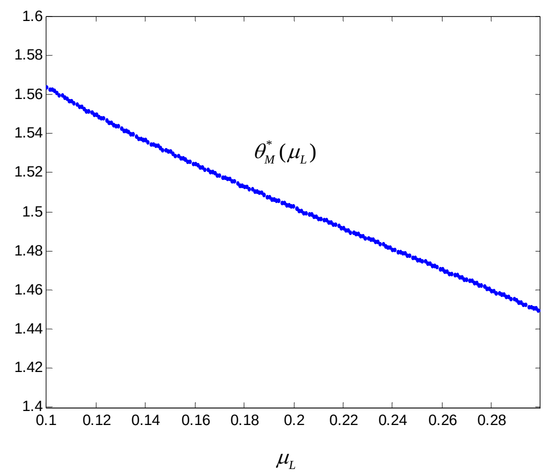

Figure1: Constitutional choice as a function of μ_L

Our computations also point out some less intuitive relationships between initial conditions and constitutional preferences by the middle type. For example, both θ_M^* and μ_M^* are decreasing in k, the intensity of horizontal differentiation. This suggests that when k decreases, for a given θ_M or μ_M, the middle type is more likely to prefer unified taxation system. The schedules $\theta_M^*(k)$ and $\mu_M^*(k)$ are shown in Figure 2, where $\theta_M^*(k)$ is plotted under the parameter values $\mu_H = \mu_M = \mu_L = 1/3$, and $\mu_M^*(k)$ is plotted by keeping $\mu_H = \mu_L$.

[12] For example, when $k = 0.5$, our computation shows that $\mu_M^* = .2630$. When $\mu_M = .2628$, the unified taxation regime is chosen and the total welfare is .3763; when $\mu_M = .2632$, the independent taxation regime is chosen and the total welfare is .3624. So even if the average productivity (average θ) increases, the welfare may go down due to the institutional shift from unified to independent.

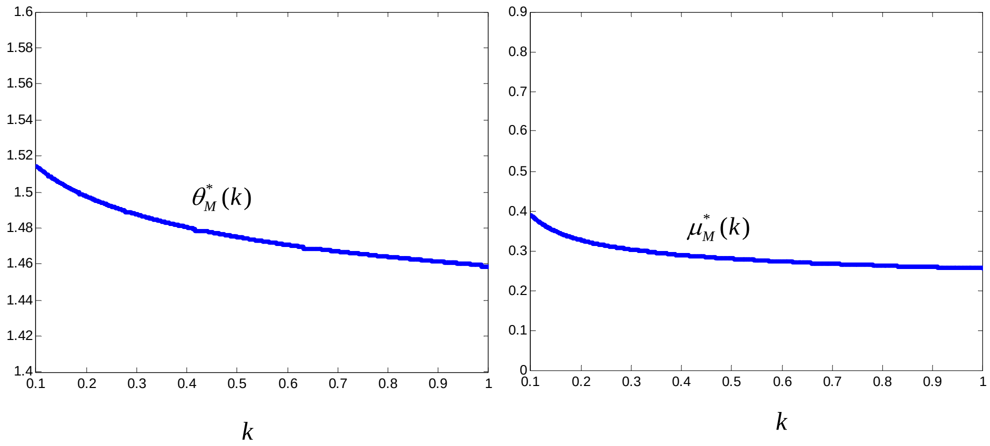

Figure 2: Constitutional choice as a function of k

When k goes down, the rich extract higher rents from the competition between States for their tax base, and hence the middle class that was previously indifferent now prefers the more efficient unified solution. The intuition for $\mu^*_M(k)$ is similar: when k goes down the share of consumption of the rich goes up, and hence the previously indifferent type would now prefer the unified regime. Indifference could be restored only if the productivity of the median type or it's numerosity went up, so that the interests of the middle class type converged to those of the rich given the utilitarian taxation designers.

In a picture with μ_M on horizontal axis and θ^*_M on vertical axis, our computations show that θ^*_M decreases as μ_M increases (while the other two types decrease symmetrically at the same time). Increasing μ_M in this way reduces inequality but also reduces total productivity when $\theta_M < 1.5$. If θ_M is less than the mean, the reduced total productivity makes the fear of being milked by the poor increase even if there are less poor agents, because that reduction is perfectly offset by an equal reduction in the number of rich.[13]

It is difficult to design a comparative statics exercise in the three type model to isolate the effect of inequality, since, as shown above, any change in the productivity distribution has also other confounding effects. We will be able to say something clearer about the role of initial inequality

[13] The pattern between θ^*_M and μ_M is a fortiori decreasing when the increase in μ_M is balanced by a reduction in μ_H only, without touching the percentage of the poor. Type M is more worried about being milked by the poor, which leads to a lower cutoff of θ^*_M.

when studying the case of a continuum of ability types.

In summary:

Remark: Weaker horizontal preferences (lower k) would push towards unification of fiscal policy in the region, but the middle class is likely to go for that only if the poor are not too poor and not too many, or if there is a sufficiently large fraction of high income earners.

This set of results fits our intuition about the situation within the European Union, where things seemed at some point mature for a new European Constitution that would concentrate a larger fraction of policy decisions in Brussels, but such a preference for unification of policy making has reversed itself after the enlargement of the Union to include a set of poorer countries that have altered the distribution of income in the Union in the opposite direction.[14]

3 Continuum of types

In this section we extend our analysis to the continuous type case. Specifically, in the vertical dimension worker-consumers are distributed on $[\underline{\theta}, \overline{\theta}]$ with density function $f(\theta)$, where $f(\theta)$ is continuous, strictly positive everywhere in the support. All the other assumptions are the same as those in the previous discrete type model.

As in the discrete type model, citizens can only be sorted in the vertical dimension. Thus, offering a tax schedule $T(Q)$ is equivalent to offering a menu of consumption and production pairs $\{C(\theta), Q(\theta)\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$. Define the tax function $T(\theta) = Q(\theta) - C(\theta)$. In the autarkic economy (no tax), a citizen's optimal consumption or gross income is given by

$$C(\theta, d) = Q(\theta, d), \textit{where} \;\; u'(C) = 1/\theta.$$

Note that the optimal consumption or gross income does not depend on d_i in autarky, and each citizen should live in the State with a lower d_i. Again we will consider two taxation rules: unified

[14] The decisions about taxation reforms may well depend on the voting system in the union: in fact, if two rich countries accept a third poorer country in the union, perhaps for reasons of economies of scale in a larger market, the "popular vote" would be more likely than earlier to be in favor of unified tax system, whereas the "electoral college" would be more likely than before to oppose it, since the median voters of the two richer countries would be against supporting also the poorer or more poor people of the new country added to the union. All these issues are vaguely related to the findings of this paper, but, strictly speaking, the model cannot be directly applied to this European Union expansion issue, since the assumption of symmetric initial conditions and distributions obviously would be violated.

rule and independent rule. Under either taxation rule, incentive compatibility has to hold. Define $V(\theta, \widehat{\theta}, d_i)$ to be the utility of a citizen with type (θ, d_i) who chooses to live in State i and reports vertical type $\widehat{\theta}$. Incentive compatibility requires

$$V(\theta, \theta, d_i) \geq V(\theta, \widehat{\theta}, d_i) \quad \forall (\theta, \widehat{\theta}) \in [\underline{\theta}, \overline{\theta}]^2$$

Given that the term kd_i enters the utility function in an additive way, the IC is equivalent to

$$V(\theta, \theta, 0) \geq V(\theta, \widehat{\theta}, 0) \quad \forall\ (\theta, \widehat{\theta}) \in [\underline{\theta}, \overline{\theta}]^2$$

That is, if IC is satisfied in the vertical dimension for $d_i = 0$, then it is satisfied for any d_i. As in the three-type case, let $v(\theta)$ denote the rent provision to type-$(\theta, 0)$ citizen. Then

$$v(\theta) = V(\theta, \theta, 0) = u(C) - \frac{Q}{\theta}.$$

By standard approach, the IC conditions are equivalent to the following two conditions:

$$v'(\theta) = \frac{Q(\theta)}{\theta^2} = \frac{1}{\theta}[u(C(\theta)) - v(\theta)] \tag{3}$$

$$Q'(\theta) \geq 0 \tag{4}$$

Note that by (3) given $v(\theta)$, $Q(\theta)$ is uniquely determined and so is $C(\theta)$. For convenience, we will work with the rent provision contract $v(\theta)$. In terms of $Q(\theta)$, (3) can be rewritten as

$$Q'(\theta) = \theta u'(C)C' \tag{5}$$

Since everyone is required to participate in one of the tax systems, the individual rationality constraint only concerns which State to live in. Given the two States' rent provisions $v_1(\theta)$ and $v_2(\theta)$, the horizontal type $x_1^*(\theta)$ who is indifferent between the two States is determined by:

$$v_1(\theta) - kx_1^* = v_2(\theta) - k\left(\frac{1}{2} - x_1^*\right) \Rightarrow x_1^*(\theta) = \frac{1}{4} + \frac{1}{2k}\left[v_1(\theta) - v_2(\theta)\right]$$

Again, $x_1^*(\theta)$ determines the borders, with all the types (θ, d_1) such that $d_1 < x_1^*(\theta)$ living in State 1 and $d_1 > x_1^*(\theta)$ living in State 2.

3.1 Unified taxation

Under the unified taxation regime, the central authority maximizes the total sum of utilities. We concentrate on symmetric taxation rules, in the sense that the two States should have the same taxation policy. As a result, the central authority just sets a tax schedule to maximize two States' sum of utility with $x_1^*(\theta) = \frac{1}{4}$. Mathematically, this can be formulated as an optimal control problem:

$$\begin{aligned}
\max \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\frac{1}{4}} [v(\theta) - kd_1]dd_1 f(\theta)d\theta &= \int_{\underline{\theta}}^{\overline{\theta}} \left[\frac{1}{4}v(\theta) - \frac{k}{32}\right] f(\theta)d\theta \\
s.t.\ \ v'(\theta) &= \frac{1}{\theta}[u(C(\theta)) - v(\theta)] \\
Q'(\theta) &\geq 0 \\
\int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{4}[Q(\theta) - C(\theta)]f(\theta)d\theta &= \int_{\underline{\theta}}^{\overline{\theta}} \frac{1}{4}[\theta(u(C) - v) - C]f(\theta)d\theta \geq 0
\end{aligned}$$

The last constraint is the resource or budget constraint, which means that a State cannot consume more than its total gross income.

We first consider a benchmark case where citizens' θ is observable. In this scenario, the programming problem is the above one without the first two (IC) constraints. The solution to this problem is trivial: everyone gets the same consumption and only the highest type $\overline{\theta}$ works. The first property comes from the concavity of the utility function, and the second one follows the fact that the highest type incurs the least disutility in producing a given quantity of output.

Now let us come back to the programming problem with θ unobservable. We first ignore the monotonicity constraint on $Q(\theta)$ to consider the relaxed program (and we will justify this approach if the solution of $Q(\theta)$ is indeed monotone). Following the standard technique of optimal control with integral constraints, we define another state variable $J(\theta)$ with

$$J'(\theta) = \frac{1}{4}[Q(\theta) - C(\theta)]f(\theta) = \frac{1}{4}[\theta(u(C) - v) - C]f(\theta)$$

Now the budget constraint can be translated into $J(\overline{\theta}) \geq 0$ and $J(\underline{\theta}) = 0$. The Hamiltonian of the problem is:

$$H = \left[\frac{1}{4}v(\theta) - \frac{k}{32}\right] f(\theta) + \lambda\frac{1}{\theta}\left[u(C(\theta)) - v(\theta)\right] + \mu\frac{1}{4}[\theta(u(C) - v) - C]f(\theta)$$

where λ and μ are the two costate variables. The optimality conditions are as follows:

$$\frac{\partial H}{\partial C} = \frac{\lambda}{\theta}u'(C) + \frac{1}{4}\mu[\theta u'(C) - 1]f = 0 \tag{6}$$

$$\lambda' = -\frac{\partial H}{\partial v} = -\frac{1}{4}f + \frac{\lambda}{\theta} + \frac{1}{4}\mu\theta f \tag{7}$$

$$\mu' = -\frac{\partial H}{\partial J} = 0 \tag{8}$$

From (8), μ is a constant. Moreover, the budget constraint should be binding at the optimum, since $u(C)$ is strictly increasing. Therefore, μ should be a positive constant. From (6) and (7) we can get rid of λ, which yields

$$-\frac{u''(C)}{[u'(C)]^2}C' = 2 - \frac{1}{\mu\theta} + \frac{f'}{f}\left[\theta - \frac{1}{u'(C)}\right] \tag{9}$$

The second order differential equation (9) is very hard to analyze for general utility functions and general density functions. In what follows we consider a specific setting, in which vertical type θ is uniformly distributed over interval $[0.4, 1.4]$, and $u(C) = 2\sqrt{C}$. In this setting, (9) becomes

$$-\frac{1}{2}\mu\theta C^{-\frac{1}{2}}C' + 2\mu\theta - 1 = 0$$

Solving for the above differential equation, we have

$$C(\theta) = \left[2\theta - \frac{1}{\mu}\log\theta + a\right]^2 \tag{10}$$

where a is some constant yet to be determined.

We also have two transversality conditions: $\lambda(\underline{\theta}) = \lambda(\overline{\theta}) = 0$. Combining with (6), we have $C(0.4) = 0.16$ and $C(1.4) = 1.96$. Substituting these two conditions into (10), we have

$$\mu = \log(1.4) - \log(0.4) = 1.2528 \ \textit{and} \ \ a = \frac{\log(0.4)}{\log(1.4) - \log(0.4)} - 0.4 = -1.1314$$

Therefore, the optimal consumption schedule is:

$$C(\theta) = \left[2\theta - 1.1314 - \frac{\log\theta}{1.2528}\right]^2 \tag{11}$$

From (5) and (11), a differential equation about $Q(\theta)$ is obtained.

$$Q'(\theta) = 4\theta - \frac{2}{1.2528} \Rightarrow Q(\theta) = 2\theta^2 - 1.5964\theta + b \tag{12}$$

where b is some constant. We can use the binding budget constraint to determine b. Specifically, b is implicitly defined by the following equation:

$$\int_{0.4}^{1.4} [Q(\theta) - C(\theta)]d\theta = 0$$

Finally, we need to complete the consistency check to verify the monotonicity condition of $Q(\theta)$. This is not an issue since $Q'(\theta) = 4\theta - \frac{2}{1.2528} \geq 4(0.4) - \frac{2}{1.2528} > 0$. As in the three-type model, the optimal solution does not depend on k, the degree of horizontal differentiation. This is because there is basically no competition between the States in luring higher ability citizens when tax schedules are set by a central authority.

The optimal tax function $T(\theta) = Q(\theta) - C(\theta)$ can be obtained from (11) and (12). Specifically,

$$\begin{aligned} T(\theta) &= -2\theta^2 - \left(4a + \frac{2}{\mu}\right)\theta + b - a^2 + \left(\frac{2a}{\mu} + \frac{4}{\mu}\theta\right)\log\theta - \left(\frac{\log\theta}{\mu}\right)^2 \\ T'(\theta) &= -4a - 4\theta + \frac{2}{\mu} + \frac{2a}{\mu}/\theta + \frac{4}{\mu}\log\theta - \frac{2\log\theta}{\theta\mu^2} \\ &= 4\left(\frac{1}{2\theta\mu} - 1\right)\left(\theta + a - \frac{\log\theta}{\mu}\right) \end{aligned} \tag{13}$$

It can be easily verified that $T'(0.4) = T'(1.4) = 0$ and $T'(\theta) > 0$ for any $\theta \in (0.4, 1.4)$. Thus the optimal tax schedule $T(Q)$ exhibits *perfect sorting*, as $T'(Q) = T'(\theta)/Q'(\theta) > 0$ for $\theta \in (.4, 1.4)$. Combining with the binding budget constraints, we conclude that there must be a cutoff type $\widehat{\theta}_u \in (0.4, 1.4)$ such that $T(\theta) < 0$ for $\theta \in [0.4, \widehat{\theta}_u)$ and $T(\theta) > 0$ for $\theta \in (\widehat{\theta}_u, 1.4]$. That is, all the types below $\widehat{\theta}_u$ receive subsidy and all types above $\widehat{\theta}_u$ pay positive tax.

3.2 Independent Taxation

Under the independent taxation regime, each State chooses its taxation schedule simultaneously and independently. Again we focus on symmetric equilibria, in which the two States choose the same taxation schedule. Suppose State 2's rent provision contract is given by $v^*(\theta)$. Then if State 1 offers rent provision contract $v(\theta)$, it can be easily verified that the type-θ "market share" for State 1 is given by (twice of) $\eta(\theta) = \frac{1}{4} + \frac{1}{2k}[v(\theta) - v^*(\theta)]$. Now State 1's maximization problem can be formulated as the following optimal control problem:

$$\begin{aligned} \max \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\eta(\theta)} [v(\theta) - kd_1]dd_1 f(\theta)d\theta &= \int_{\underline{\theta}}^{\overline{\theta}} \left[v(\theta)\eta(\theta) - \frac{k}{2}\eta(\theta)^2\right] f(\theta)d\theta \\ \text{s.t. } v'(\theta) &= \frac{1}{\theta}[u(C) - v] \\ Q'(\theta) &\geq 0 \\ J'(\theta) &= [\theta(u(C) - v) - C]\eta f(\theta) \\ J(\underline{\theta}) &= 0, \ J(\overline{\theta}) \geq 0 \end{aligned}$$

where $J(\theta) = \int_{\underline{\theta}}^{\overline{\theta}} [\theta(u(C) - v) - C]\eta(\theta)f(\theta)d\theta$ is the state variable associated with the budget constraint.

Benchmark with θ observable If θ were observable, then the programming problem is the above one without the first two (IC) constraints. Without the transition equation, we can do pointwise maximization. Let μ be the multiplier of the resource constraint. After imposing symmetry, the first order conditions that characterize the equilibrium can be written as:

$$[\frac{1}{8}+\frac{1}{2k}v(\theta)+\frac{\mu}{2k}(Q-C)]u'(C) = \frac{\mu}{4} \tag{14}$$
$$[\frac{1}{8}+\frac{1}{2k}v(\theta)+\frac{\mu}{2k}(Q-C)]\frac{1}{\theta} = \frac{\mu}{4} \tag{15}$$

From (14) and (15), we get $u'(C) = 1/\theta$. This implies that the consumptions are not distorted (using autarky as the benchmark case). After determining the consumption schedule, the schedule $Q(\theta)$ is implicitly defined by (15) and the resource constraint.

Unobservable θ Turning back to the programming problem with θ unobservable, we again drop the monotonicity constraint $Q'(\theta) \geq 0$ and define the Hamiltonian:

$$H = \left(v\eta - \frac{k}{2}\eta^2\right) f + \frac{\lambda}{\theta}[u(C)-v] + \mu\eta[\theta(u(C)-v)-C]f$$

The optimality conditions (after imposing symmetric conditions in equilibrium) are:

$$\begin{aligned}
\frac{\partial H}{\partial C} &= \frac{\lambda}{\theta}u'(C) + \mu\eta[\theta u'(C)-1]f = 0 & (16)\\
\lambda'(\theta) &= -\frac{\partial H}{\partial v} = -\left[\frac{v}{2k}+\frac{\eta}{2}\right]f + \frac{\lambda}{\theta} - \frac{\mu}{2k}[\theta(u(C)-v)-C]f + \mu\eta\theta f & (17)\\
\mu'(\theta) &= -\frac{\partial H}{\partial J} = 0 \Rightarrow \mu \textit{ is a constant}
\end{aligned}$$

where $\eta = \frac{1}{4}$. Further simplifying the above conditions, we have

$$\begin{aligned}
\lambda &= \mu\eta[\frac{\theta}{u'(C)} - \theta^2]f \\
-\frac{u''(C)}{[u'(C)]^2}C' &= 2 - \frac{k+4v+[\theta(u(C)-v)-C]}{2k\mu\theta} + \frac{f'}{f}\left[\theta - \frac{1}{u'(C)}\right] & (18)\\
v'(\theta) &= \frac{1}{\theta}[u(C)-v] & (19)
\end{aligned}$$

For the specific setting we consider above, (18) becomes

$$C' = 4\sqrt{C} - \frac{k+4v+[\theta(u(C)-v)-C]}{2k\mu\theta}\sqrt{C} \tag{20}$$

We also have two transversality conditions

$$\lambda(0.4) = \lambda(1.4) = 0 \Rightarrow C(0.4) = 0.16 \textit{ and } C(1.4) = 1.96$$

(20) and (19) form a first order differential equation system with three unknowns $(C(\theta), v(\theta), \mu)$, along with three boundary conditions (the third is from the resource constraint).

Since this differential equation system does not have analytical solution, we resort to computations.

3.3 Computational results

The numerical results given different values of k all show that the equilibrium under independent taxation exhibits perfect sorting. For $k = 0.5$, Figure 3 compares the tax schedules under two taxation systems. It is evident that the tax schedule under independent rule is flatter. Generally speaking, higher types are taxed less and lower types get less subsidy under the independent system.

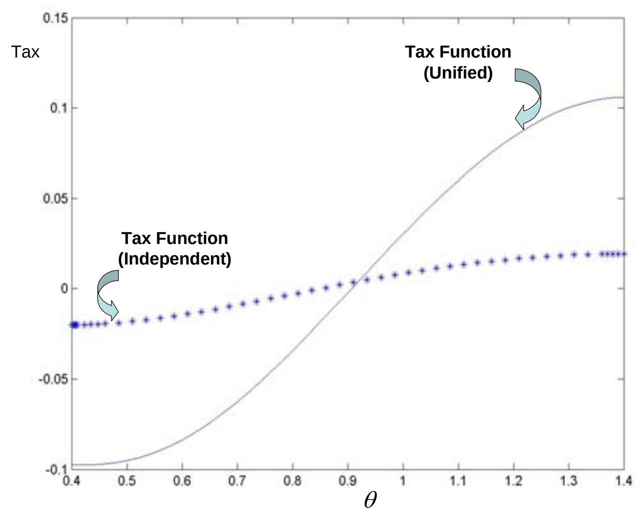

Figure 3: Tax functions comparison ($k = 0.5$)

This pattern is robust with respect to other values of k. Actually, as k decreases, the tax function under the independent rule becomes flatter, as shown in Figure 4. So our model predicts that *the difference in progressivity will be smaller when the intensity of location preferences is smaller.*[15] In

[15] Note that the tax function under the unified rule is independent of k.

the limit as $k \to 0$, $T(\theta) = 0$ under the independent rule.[16]

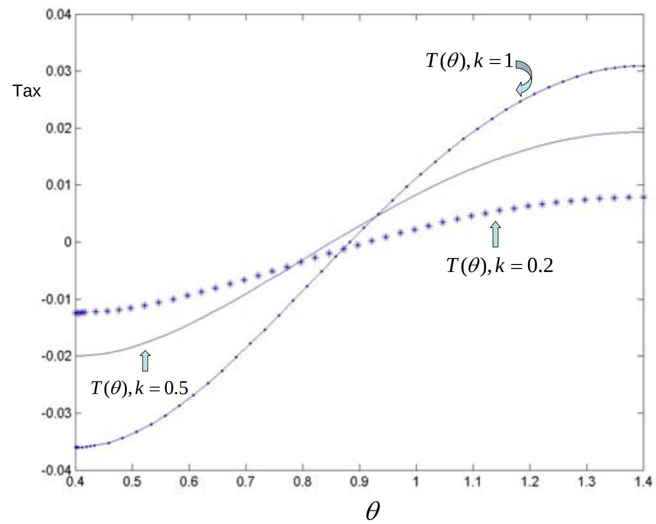

Figure 4: Tax functions under different k's

The following figure compares the tax schedules under independent taxation with θ observable and θ unobservable ($k = 0.5$). As we can see from the figure, the tax schedule is flatter with θ being unobservable. This makes sense, since unobservable θ makes redistribution more difficult as high types can mimic low types. Another observation from this figure is that even with observable θ, competition will significantly flatten the tax schedule: each state's incentive to steal high types from the other state leads to lower taxes to high types.

[16] This is similar to Bertrand competition outcome. First of all, positive tax cannot be sustained over any interval of types; otherwise the other State can always benefit by attracting all the types from this interval by reducing the tax by an arbitrarily small (ϵ) amount. Second of all, given that the other State offers zero tax rates, the State in question cannot provide subsidy to attract higher type citizens without imposing some positive tax over some interval of types (the resource constraint).

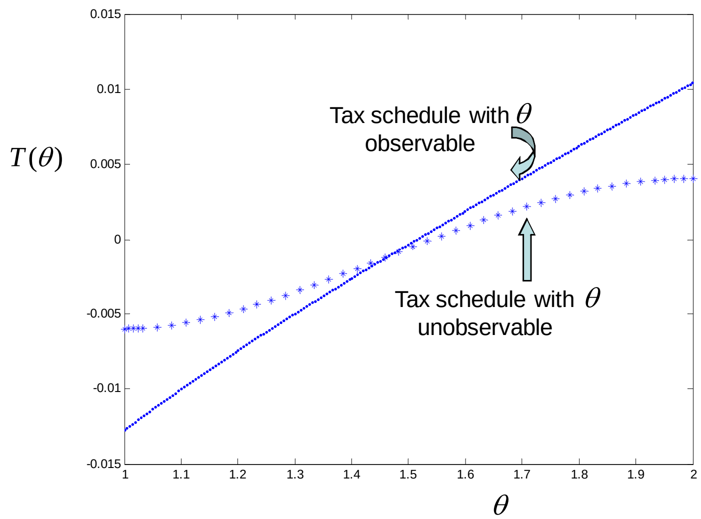

Figure 5: Tax Schedules: θ observable vs. θ unobservable

In Simula and Trannoy (2006), a "curse" of middle-skilled workers is identified, in the sense that the marginal tax rate is negative at the top and the average tax rate is decreasing over some interval close to the top. Such a curse does not occur in our three-type model.[17] In this model with a continuum of types, such a curse continues to be absent when k is not too small. We observe negative tax rate close to the top only when k is sufficiently small. Even in such cases, our computation shows that the effect is quantitatively insignificant, either because the interval over which the negative marginal tax rate occurs is very small, or the magnitude of the negativity is insignificant. Taking these together, our result suggests that Simula and Trannoy's finding may not be robust in a more general analysis when outside options are endogenously determined.[18]

Figure 6 compares the utility schedules under the two tax regimes for $k = 0.5$. As we can see, there is a unique cutoff type below which all types prefer the unified regime and above which all types prefer the independent regime. This pattern holds for various k's as well. For $k = 0.5$, the cutoff (indifferent) type is about 0.836, which is below the median 0.9. So under majority voting the independent regime will be selected.

[17]In our three-type model, our computation shows that as $k \to 0$, all taxes go to zero. But the average tax rate $(T(Q)/Q)$ is always higher for type H than that for type M, no matter how small k is.

[18]The difference between our finding and theirs may also be due to the specific functional form assumed for the moving cost in their model.

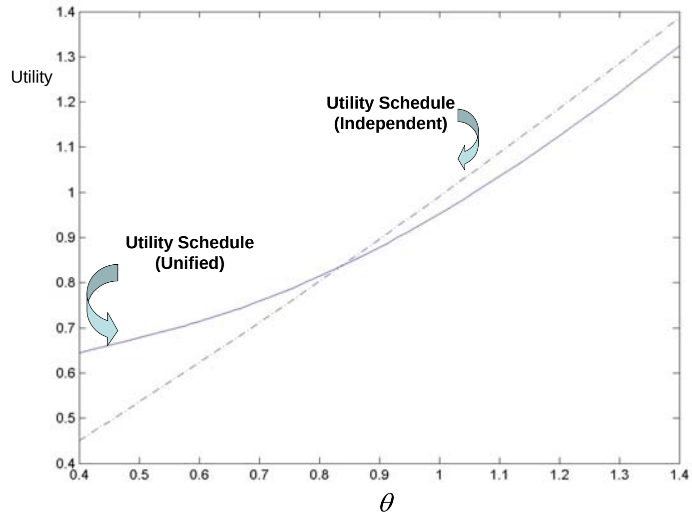

Figure 6: Utility schedules comparison ($k = 0.5$)

All the above results are consistent with those in the discrete type model: (1) Independent system has less progressive tax; (2) The smaller the horizontal differentiation between two States, the fiercer competition leads to less progressive tax in the independent system; (3) Higher types are better off in the independent system, lower types are better off in the unified system, and the median type may make an inefficient constitutional choice.

We are also interested in how changes in the (type) income distribution affect the constitutional choice. Fix $k = 0.5$. Consider a family of distributions with support $[1, 2]$. Specifically, the density function is given by

$$f_a(\theta) = \frac{1}{3(1 - a/2)}(3 - a\theta).$$

with $a \in [0, 1.5)$. As a increases, the distribution is tilted toward more poor people, and $a = 0$ is the uniform distribution. The median type θ_m is given below.

$$\theta_m(a) = \frac{6 - \sqrt{36 - 36a + 10a^2}}{2a}$$

The simulation results are reported in the following table. (θ^* is the cutoff type indifferent between the two tax regimes, and θ_m is the median type):

	$a = 0$	$a = 1$	$a = 1.3$	$a = 1.4$
θ^*	1.463	1.4205	1.3705	1.362
θ_m	1.5	1.4189	1.3578	1.3284

For the four as we computed, all the results exhibit perfect sorting. Two observations are worth noting:

First, as a increases (more poor around), the *indifferent type* monotonically decreases. This is consistent with the results from the three-type model and the intuition is similar: more poor people means more taxes for higher types in the unified regime, while in the independent regime the solution is closer to autarky. Therefore, the indifference type will decrease.

On the other hand, if a is sufficiently large ($a > 1$), the *median type* prefers the unified regime. Having more poor people makes the unified system more likely. This seems to contradict the results from the three-type case. However, they can be reconciled. In the three-type case, when we increase the poor the median type remains the middle type. In the continuous case, the median type decreases as the distribution is tilted toward the poor.[19] The interesting finding is that

Remark: As poverty increases, the *indifference type* decreases slower than the *median type* does. As a result, the percentage of people favoring the unified system gradually increases. Hence we should have the optimal unified system chosen when poverty is very high or very low, whereas in the intermediate region the constitutional choice is suboptimal.

In the next comparative statics exercise, we study how the degree of inequality affects constitutional choice by examining a mean preserving spread. Fix $k = 0.5$. Consider a family of distributions with support $[1, 2]$. Specifically,

$$f_a(\theta) = \frac{1}{3 - a/12}[3 - a(1.5 - \theta)^2]$$

with $a \in [0, 12)$. As a increases, the distribution becomes more concentrated around the mean (median) 1.5 (equality increases) and $a = 0$ is the uniform distribution. The simulation results are reported in the following table. (θ^* is once again the cutoff type indifferent between two tax regimes):

[19] In the three-type case, our numerical exercises always assume that the median voter is type M, and we still have the freedom to vary θ_M. However, in the continuous type case, the median voter is endogenously determined. When we change the distribution, the median voter will change correspondingly.

	$a = 0$	$a = 8$	$a = 10$	$a = 11$
θ^*	1.463	1.446	1.4345	1.4185

The table shows the pattern that

Remark: As the income distribution becomes more equal, the indifference type decreases.[20]

With respect to the horizontal differentiation parameter k, our result shows that

Remark: The cutoff type θ^* is not very sensitive to changes in k, even though they significantly affect the progressivity of tax functions in the independent system.

4 Extension to n States

In this section we extend our analysis to any finite n number of States. Specifically, assume that on the horizontal dimension there are n States ($n \geq 2$), the locations of which evenly split the unit-length circle. Each State offers vertically differentiated contracts $(C(\theta), Q(\theta))$. Again, we consider two taxation regimes. In the unified regime, the federation offers optimal tax schedules to maximize the total sum of utilities. In the independent taxation regime, each State's objective is to maximize the total sum of utilities among the residents in its own State, by offering its menu of contract $(C(\theta), Q(\theta))$ given other States' menus of contracts. Again, for the unified model we look for symmetric optimal menu, and for the independent model we look for symmetric Bertrand-Nash equilibria in which each State offers the same menu of contracts.[21] An n-tuple rent provision contract profile $(v^*, ..., v^*)$ constitutes a symmetric equilibrium if, given that all other States each offer v^*, the best response of the State in question is also to choose $v = v^*$.

For the unified taxation model, it is straightforward to see that the Federation's optimization program does not involve n, hence the solution does not depend on the number of States. For the independent taxation model, if we define $k' = k/n$ as the normalized degree of horizontal differentiation, then by inspection, in terms of k' the first order conditions in the three-type model and the differential equation system in the continuous-type model are exactly the same as their counterparts

[20] Given that the median is 1.5 in this example, the independent tax system will be endogenously chosen by majority voting. This example does suggest one potential explanation for why countries like Sweden show more independence nowadays.

[21] As a direct consequence each state is effectively competing with two adjacent states, a common feature implied by the Salop model.

in the two-State case (where $k' = k/2$). This implies that the analysis of the n-State case can be translated into the analysis of the two-State case through normalizing k by n, and in terms of k' the solution to the n-State model is the same as the solution to the two-State model. Thus all the results from the two-State model carry over to the n-State competitive model. In particular, the effect of an increase in n (while holding k fixed) on the equilibrium is exactly the same as the effect of a decrease in k on the two-State equilibrium.

One interesting implication from our n-State model is that

Remark: The preference ordering of the middle type (in the three-type model) may exhibit non-monotonic features: consider a situation in which at most two States can be formed, and $\theta_M = \theta_M^* + \epsilon$, so that the middle type prefers independent regime with two separate tax authorities. Now suppose that it becomes feasible to have three States. Since we know that $\theta_M^*(n)$ is increasing in n, for ϵ small enough it must be true that the middle type now prefers the unified system to having three States. However, the utility under unified regime does not depend on n, hence the preference ordering of the middle type is as follows: the independent system with two States is preferred to the unified system with two States. But the unified system with three States is preferred to the independent system with three-States.

5 Concluding remarks

This paper has provided the first analysis of nonlinear income taxation when strategic authorities compete for resident citizens along two dimensions, the vertical productivity dimension and the horizontal location dimension. Every agent's productivity and location preferences are private information and we have explored the relative importance of these two dimensions for the degree of progressivity of the tax system, comparing the competitive nonlinear taxation game with the unified optimal taxation benchmark of Mirrlees (1971). Moreover, the model has allowed us to discuss the incentives of different classes of agents to advocate for different systems at the constitutional stage.

In the model with three productivity types, we have shown analytically that the independent taxation system yields lower progressivity than in the unified case. We have studied how the preferences of the median type, responsible for the constitutional choice, vary with the initial conditions on the parameters, showing that the median type is more likely to choose the unified system (1) when location choices have a lower weight in the utility function with respect to consumption (k is smaller), or when the number of States is larger; (2) when the distribution of income has a lower percentage of poor or unproductive types; (3) when the middle class is not too large.

In this paper we have studied only the case in which location preferences are unobservable. Beside the fact that unobservable location preferences is a realistic assumption, the study of the observable location preferences case is problematic: in this case the taxation strategy chosen by each State has to depend also on the horizontal location. This implies that, for example, the preferences at the constitutional stage are in the two-dimensional space, and the identity of the median voter changes in an unclear way when the parameters of the model vary. This extension will be very interesting for the political economy implications of our analysis.

Perhaps a more important extension of this model will be the consideration of asymmetric initial conditions. Making the density function of agents on the circle be type dependent and with asymmetric mass in the neighborhood of the various States, allowing therefore for States that are more frequently preferred by high types than other States, will enable us to study the interaction of such initial asymmetries with the relative progressivity implications of different systems as well as with the constitutional choice with respect to the symmetric benchmark. Tracing the impact of different initial conditions on constitutional choice will also allow us to start a dynamic analysis of persistence of inequality differences across countries due to the different institutions that have different feedbacks on inequality.

Finally, a very important part of our future research on this topic will be an empirical study of the relationship between progressivity and centralization of the tax system. It is extremely difficult to evaluate the progressivity of a government. Some studies exist about the progressivity of tax schedules, i.e. measuring the marginal tax rate on the various income brackets and comparing the derived schedule with the benchmark proportional overall tax schedule.[22] Such data on progressivity across countries[23] does not however include the variation across countries in terms of transfers and public provision of private goods. Public economists usually separate the study of the tax system

[22] The proportional overall benchmark tax system is one in which, when considering all types of taxes, the distribution of after tax disposable income is the same as that before tax.

[23] See OECD data at
http://www.oecd.org/document/60/0,3343,en_2649_37427_1942460_1_1_1_37427,00.html.

from the expenditure side on purpose. But, by doing so, the overall impact of a government on inequality is harder to establish. To give a simple example, if country A has a proportional overall tax system but targets the tax revenue to the poor with targeted transfers or public provision of the private goods, they would have had otherwise to pay for housing, education, or health, etc. Such a country would have an overall progressive impact, and would certainly be more progressive than a country B whose government opted for a progressive tax system but didn't provide any of the social goods mentioned above and the poor would have to get them on the market (or remain uninsured, uneducated, or homeless). A serious empirical investigation will require the construction of an enriched data set about the overall progressivity of the system, something that currently does not exist.[24]

[24] Another serious obstacle could be endogeneity. For example, our model suggests that countries with less inequality may choose independent regimes, but independent regimes do not reduce inequality as much as a unified system does. Hence, if we see that in a country like Sweden with very low inequality there is a large share of decentralized taxes, this observation would be consistent with the theory if we focus on the constitutional incentives, but inconsistent if we reversed the causality and considered the taxation system as the independent variable. A third problem with the empirical analysis relates to the other main variable of interest, namely the degree of decentralization of taxation agencies. At the theoretical level, we would need to allow for an *overlap* of the national authorities and the local authorities in competition with one another, and then evaluate the impact of changing the share of tax revenue that the overall nation wants from the two levels. Then, it will be still necessary to integrate the existing data with some new collection, because to the best of our knowledge there is no systematic study of such a degree of centralization. The OECD data set only distinguishes between tax rates at the central level from the tax rates at the subcentral level, but this is too little information.

Appendix

Proof of Lemma 2: First we show that $u'(C_L^I) < u'(C_L^U)$.

$$\begin{aligned}
u'(C_L^I) &= \frac{\mu_L \lambda_R/4}{\mu_L\beta_L - \lambda_M + \mu_L\lambda_R T_L/(2k)} \\
&= \frac{1/4}{\frac{\beta_L}{\lambda_R}\frac{\theta_L-\theta_M}{\theta_L} + \frac{1}{4}\theta_M + T_L\frac{1}{2k}\frac{\theta_L-\theta_M}{\theta_L}}
\end{aligned}$$

Define operation $\sim$ such that $A \sim B$ means that B has the same sign as A. Then

$$\begin{aligned}
u'(C_L^U) - u'(C_L^I) &= \frac{1}{\frac{\theta_L-\theta_M}{\lambda_R^U\theta_L} + \theta_M} - \frac{1/4}{\frac{\beta_L}{\lambda_R}\frac{\theta_L-\theta_M}{\theta_L} + \frac{1}{4}\theta_M + T_L\frac{1}{2k}\frac{\theta_L-\theta_M}{\theta_L}} \\
&\sim \left[\frac{\beta_L}{\lambda_R}\frac{\theta_L-\theta_M}{\theta_L} + \frac{1}{4}\theta_M + T_L\frac{1}{2k}\frac{\theta_L-\theta_M}{\theta_L}\right] - \frac{1}{4}\left[\frac{\theta_L-\theta_M}{\lambda_R^U\theta_L} + \theta_M\right] \\
&= \frac{\theta_L-\theta_M}{\theta_L}\left[\frac{\frac{1}{4}}{\lambda_R^U} - \frac{\beta_L}{\lambda_R} - \frac{T_L}{2k}\right] \\
&\sim \frac{1}{4}\frac{1}{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}} - \beta_L\frac{\frac{1}{4} - \left(T_H\frac{\mu_H}{2k\theta_H} + T_M\frac{\mu_M}{2k\theta_M} + T_L\frac{\mu_L}{2k\theta_L}\right)}{\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M} + \mu_L\frac{\beta_L}{\theta_L}} - \frac{T_L}{2k} \\
&\sim \frac{1}{4}\left[\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M} + \mu_L\frac{\beta_L}{\theta_L}\right] \\
&\quad -\beta_L\left[\frac{1}{4} - \left(T_H\frac{\mu_H}{2k\theta_H} + T_M\frac{\mu_M}{2k\theta_M} + T_L\frac{\mu_L}{2k\theta_L}\right)\right]\left(\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}\right) \\
&\quad -\frac{T_L}{2k}\left(\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}\right)\left(\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M} + \mu_L\frac{\beta_L}{\theta_L}\right) \\
&= \frac{1}{4}\left[(\beta_H-\beta_L)\frac{\mu_H}{\theta_H} + (\beta_M-\beta_L)\frac{\mu_M}{\theta_M}\right] + \frac{1}{2k}\left(\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}\right)\cdot \\
&\quad \left[T_H\frac{\mu_H}{\theta_H}\beta_L + T_M\frac{\mu_M}{\theta_M}\beta_L - T_L\left(\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M}\right)\right] \\
&\sim T_H\frac{\mu_H}{\theta_H}\beta_L + T_M\frac{\mu_M}{\theta_M}\beta_L - T_L\left(\mu_H\frac{\beta_H}{\theta_H} + \mu_M\frac{\beta_M}{\theta_M}\right)
\end{aligned}$$

If $T_M \geq 0$, then $u'(C_L^U) - u'(C_L^I) > 0$. If $T_M < 0$, we can rewrite the last equation above into

$$\frac{\mu_H}{\theta_H}\left[T_H\beta_L - T_L\beta_H\right] + \frac{\mu_M}{\theta_M}\left[T_M\beta_L - T_L\beta_M\right]$$

which is also positive since $T_M < 0$ implies that $|T_M| < |T_L|$ (by Lemma 1). Next we show that $u'(C_M^I) < u'(C_M^U)$.

$$
\begin{aligned}
u'(C_M^U) &= \frac{\mu_M \lambda_R^U}{\mu_M - \lambda_H + \lambda_M} = \frac{1}{\frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R^U} - \frac{\mu_H}{\mu_M}\theta_H - \frac{\mu_L}{\mu_M}\theta_M} \\
u'(C_M^I) &= \frac{\frac{1}{4}\mu_M \lambda_R}{\mu_M \beta_M - \lambda_H + \lambda_M + \mu_M \lambda_R \frac{T_M}{2k}} \\
&= \frac{\frac{1}{4}}{\frac{\beta_M+\frac{\mu_H}{\mu_M}\beta_H+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\beta_L}{\lambda_R} - \frac{1}{4}\frac{\mu_H}{\mu_M}\theta_H - \frac{1}{4}\frac{\mu_L}{\mu_M}\theta_M + \frac{\mu_H}{\mu_M}\frac{T_H}{2k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{2k} + \frac{T_M}{2k}}
\end{aligned}
$$

$$
\begin{aligned}
u'(C_M^U) - u'(C_M^I) &\sim \frac{\beta_M+\frac{\mu_H}{\mu_M}\beta_H+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\beta_L}{\lambda_R} - \frac{1}{4}\frac{\mu_H}{\mu_M}\theta_H - \frac{1}{4}\frac{\mu_L}{\mu_M}\theta_M + \frac{\mu_H}{\mu_M}\frac{T_H}{2k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{2k} + \frac{T_M}{2k} \\
&\quad -\frac{1}{4}\frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R^U} + \frac{1}{4}\frac{\mu_H}{\mu_M}\theta_H + \frac{1}{4}\frac{\mu_L}{\mu_M}\theta_M \\
&\sim \left(\beta_M + \frac{\mu_H}{\mu_M}\beta_H + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\beta_L\right)\lambda_R^U - \frac{1}{4}\left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)\lambda_R \\
&\quad + \left(\frac{\mu_H}{\mu_M}\frac{T_H}{2k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{2k} + \frac{T_M}{2k}\right)\lambda_R^U \cdot \lambda_R \\
&\sim \left(\beta_M + \frac{\mu_H}{\mu_M}\beta_H + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\beta_L\right)\left(\frac{\mu_H}{\theta_H}+\frac{\mu_M}{\theta_M}+\frac{\mu_L}{\theta_L}\right) \times \\
&\quad \left[\frac{1}{4} - \left(T_H\frac{\mu_H}{2k\theta_H} + T_M\frac{\mu_M}{2k\theta_M} + T_L\frac{\mu_L}{2k\theta_L}\right)\right] \\
&\quad -\frac{1}{4}\left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)\left(\mu_H\frac{\beta_H}{\theta_H}+\mu_M\frac{\beta_M}{\theta_M}+\mu_L\frac{\beta_L}{\theta_L}\right) \\
&\quad + \left(\frac{\mu_H}{\mu_M}\frac{T_H}{2k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{2k} + \frac{T_M}{2k}\right)\left(\frac{\mu_H}{\theta_H}+\frac{\mu_M}{\theta_M}+\frac{\mu_L}{\theta_L}\right)\left(\mu_H\frac{\beta_H}{\theta_H}+\mu_M\frac{\beta_M}{\theta_M}+\mu_L\frac{\beta_L}{\theta_L}\right) \\
&= A + B
\end{aligned}
$$

where

$$
\begin{aligned}
A &= \mu_H\left(\frac{1}{\theta_M}-\frac{1}{\theta_H}\right)(\beta_H-\beta_M) + \frac{\mu_H}{\mu_M}\frac{\mu_L}{\theta_L}\left(1-\frac{\theta_M}{\theta_H}\right)(\beta_H-\beta_L) \\
B &= \mu_H\left(\frac{1}{\theta_M}-\frac{1}{\theta_H}\right)\left(\frac{T_H}{2k}\beta_M - \frac{T_M}{2k}\beta_H\right) + \frac{\mu_H}{\mu_M}\frac{\mu_L}{\theta_L}\left(1-\frac{\theta_M}{\theta_H}\right)\left(\frac{T_H}{2k}\beta_L - \frac{T_L}{2k}\beta_H\right)
\end{aligned}
$$

$A > 0$, and $B > 0$ too if $T_M \leq 0$. If $T_M > 0$, then the resource constraint implies that $T_M < |\frac{\mu_L}{\mu_M}T_L|$, in which case

$$
B \sim \frac{1}{2k}\mu_H\beta_H\left(\frac{1}{\theta_M}-\frac{1}{\theta_H}\right)\left[-T_M - \frac{\theta_M}{\theta_L}\frac{\mu_L}{\mu_M}T_L\right] > 0
$$

■

References

[1] Alesina, Alberto, Ignazio Angeloni, and Federico Etro (2005): "International Unions," *American Economic Review* 95 (3): 602-615.

[2] Alesina, A. and E. Spolaore (2003): *The size of Nations*, MIT Press, Cambridge.

[3] Alesina, A., E. Spolaore, and R. Wacziarg (2000): "Economic Integration and Political Disintegration," *American Economic Review* 90 (5): 1276-1296.

[4] Austen-Smith, David (2000): "Redistributing Income Under Proportional Representation," *Journal of Political Economy* 108 (6): 1235-69.

[5] Besley, Timothy and Anne Case (2003): "Political Institutions and Policy Choices: Evidence from the United States," *Journal of Economic Literature* 41(1): 7-73.

[6] Deininger, Klaus and Squire, Lyn (1996): "A New Data Set Measuring Income Inequality," *World Bank Economic Review* 10: 565-591.

[7] Epple, Dennis, and Thomas Romer (1991): "Mobility and Redistribution," *The Journal of Political Economy* 99 (4): 828-58.

[8] Gordon, Roger H. (1983): "An Optimal Taxation Approach to Fiscal Federalism," *The Quarterly Journal of Economics* 98(4): 567-86.

[9] Huber, Bernd (1999): "Tax competition and tax coordination in an optimum income tax model," *Journal of Public Economics* 71: 441-58.

[10] Iversen, T. and Soskice, D. (2006): "Electoral institutions, parties and the politics of class: Why some democracies distribute more than others," *American Political Science Review* 100(2): 165-181.

[11] Katz, Michael (1991): "Game Playing Agents: Unobservable Contracts as Precommitments," *Rand Journal of Economics* 22(3): 307-329.

[12] Ladd, Helen F. and Fred C. Doolittle (1982): "Which Level of Government Should Assist the Poor?" *National Tax Journal* 35 (3): 323-36.

[13] Martimort, David and Lars Stole (1997): "Communication Spaces, Equilibria Sets and the Revelation Principle Under Common Agency," unpublished working paper.

[14] McAfee, Preston (1993): "Mechanism design by competing sellers," *Econometrica* 61: 1281-1312.

[15] Mirrlees, James (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies* 38(2): 175-208.

[16] Oates, Wallace E. (1977): "An Economist's Perspective on Fiscal Federalism," In *The Political Economy of Fiscal Federalism*, edited by Wallace E. Oates. Lexington, Mass.: Heath.

[17] OECD (2006): Revenue Statistics 1965-2005, Centre for Tax Policy and Administration.

[18] Peck, James (1997): "A Note on Competing Mechanisms and the Revelation Principle," unpublished working paper.

[19] Persson, T. and G. Tabellini (2003): *The Economic Effects of Constitutions*, MIT University Press, Cambridge.

[20] Piketty, Thomas and Emmanuel Saez (2006): "The Evolution of Top Incomes: A Historical and International Perspective," *American Economic Review*, May 96 (2): 200.

[21] Piketty, Thomas and Emmanuel Saez (2007): "How Progressive is the U.S. Federal Tax System? A Historical and International Perspective," *Journal of Economic Perspectives* 21(1): 3-24.

[22] Rochet, Jean-Charles and Lars Stole (2002): "Nonlinear Pricing with Random Participation," *Review of Economic Studies* 69(1): 277-311.

[23] Roemer, John E.; Aaberge, Rolf; Colombino, Ugo; Fritzell, Johan; Jenkins, Stephen P.; Lefranc, Arnaud; Marx, Ive; Page, Marianne; Pommer, Evert; Ruiz-Castillo, Javier; Jesus San Segundo, Maria; Tranaes, Torben; Trannoy, Alain; Wagner, Gert G.; Zubiri, Ignacio (2003): "To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?" *Journal of Public Economics* 87 (3-4): 539-565.

[24] Simula, Laurent and Alain Trannoy (2006): "Optimal Non-linear Income Tax when Highly Skilled Individuals Votes with Their Feet," Working Paper.

[25] Ticchi, Davide and Andrea Vindigni (2007): "Endogenous Constitutions," Unpublished Working Paper, Princeton University.

[26] Wilson, John Douglas (1999): "Theories of Tax Competition," *National Tax Journal*, Vol. 52, 269-304.

[27] Yang, Huanxing and Lixin Ye (2008): "Nonlinear Pricing, Market Coverage, and Competition," *Theoretical Economics* 3: 123-153.