Competitive Nonlinear Taxation and Constitutional Choice[*]

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

First draft May 2008, this draft May 2009

Abstract

In an economy where agents are characterized by different productivities (vertical types) and different abilities to move (horizontal types), we compare a unified nonlinear optimal taxation schedule with the equilibrium taxation schedule that would be chosen by two competing tax authorities if the same economy were divided into two States. The overall level of progressivity and redistribution is unambiguously lower under competitive taxation than under unified taxation; the "rich" are always in favor of competing authorities and local governments, whereas the "poor" are always in favor of unified taxation. The constitutional choice between fiscal regimes depends on the preferences of the middle class, which in turn depend on the initial conditions in terms of the distribution of abilities (incomes), the relative power of the various classes, and mobility costs. In particular, as mobility increases, it becomes increasingly likely that a reform in the direction of unification of fiscal policies in a federation will receive majority support, while an increase in poverty can have the opposite effect.

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

[†]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 the institutional structure chosen at the constitutional stage and the initial conditions at that time. This paper is motivated by these considerations and aims to provide a clear analysis of the impact of key fiscal institutional choices, made under different initial conditions, on the difference in progressivity. In other words, our 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 mobility.

The constitutional choice of which "taxation regime" to select (centralized versus decentralized, State taxes versus City 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 affect 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 "State versus Federal taxes," "City versus State taxes," or "property taxes versus centralized funding of schools" as mainly due to *efficiency or freedom to choose* considerations. 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.

---

[1] Sweden is well known for the high degree of progressivity of the tax system and for the generous transfers, whereas Switzerland is at the opposite extreme. For the available evidence across the OECD countries, see reference [22].

We consider a competitive nonlinear taxation framework with both vertically and horizontally differentiated agents. More specifically, we consider a Hotelling type model in which two 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 abilities to move from one State to the other, or equivalently, their location preferences, reflecting their tastes for different cultures, landscapes, food, political systems, weather conditions, etc.

Under a unified taxation system, the Federation's objective is to choose an optimal tax schedule to maximize the average utility 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 average utility of all the citizens choosing to live in the State, given the other States' tax schedules. At the constitutional stage, the representatives of the various types or classes of citizens evaluate the two regimes on the basis of the solutions of these maximization programs.[2]

In the 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 optimal taxation designer, but also the additional individual rationality constraint derived from location preferences. We show that in this independent taxation regime the tax for the high type is lower and the subsidy for the 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 (so as to raise its 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

---

[2] An alternative assumption that would lead to the same relative evaluation of the two regimes would be to assume that fiscal constitutional designers expect that fiscal policies will be decided by a two-party electoral competition with probabilistic voting (see, e.g., Coughlan, 1992 for a well known implementation result of this kind).

trivial cases in which one of the two extreme types has the absolute majority at the constitutional stage). Even though a unified regime always yields higher average utility, we can show that a country with better initial conditions (higher average productivity) may end up with lower welfare because of a suboptimal majority decision at the constitutional stage.

One of our clearest findings is that, as mobility increases, it becomes increasingly likely that the decisive middle class will prefer to have (or to switch to) a unified system. The intuition for this robust result is as follows: under any taxation regime the middle class "benefits" from the presence of richer citizens who pay more taxes (or even pay them indirectly a transfer) and "suffer" from having to support the poor through the tax system; under a unified system these two contrasting effects do not depend on mobility costs, but in the independent system they do: as mobility costs go down, competition for the rich reduces the "benefits" mentioned above, while the need to support the poor remains roughly unchanged, hence the previously indifferent middle type likes the unified system more in relative terms. Our computations also show that the greater the size of the middle class, the more likely it is that the preferences of such a decisive class will be in favor of independent taxation, as the support of the poor is more spread out. Finally, our computations show that the larger the population of the poor, the more likely that the middle type will prefer independent taxation, as the fear to support the poor increases.

Since our analysis about constitutional choice (the preferences of the middle type) mostly relies on computations, for robustness check we extend our analysis to the continuous type model, which can be regarded as the limiting case of many finite types. With a continuum of types, the tax schedule chosen under each regime is characterized by a second-order differential equation with two boundary values. By focusing on the case where the vertical types are distributed uniformly, we are able to show that under independent taxation, the higher the mobility, the higher the consumption for all but the highest and lowest types; the rich (types sufficiently close to the highest type) pay lower tax, and the poor (the types sufficiently close to the lowest type) receive lower subsidy under competition; there exists a cutoff type θ^* so that all types above θ^* are better off, and all types below are worse off with competition. Our computations have also confirm most of the findings from the three type model regarding the preferences of the median type, who is now responsible for the constitutional choice.

It is important to remark that when we talk about constitutional choice we always think of it as being made by the same people who are then going to be subject to the regime they choose, the

opposite extreme with respect to a choice made behind a veil of ignorance. Thus, we have in mind situations like the choice to adopt or not a new constitution with more integrated fiscal policy in the European Union, where preferences for such a potential reform are likely to be affected by self interest considerations by the citizens who would be asked to ratify it. Our analysis in this research provides a number of considerations and interpretations regarding such situations in the European Union. As barriers to labor mobility fall and mobility costs go down, a first effect based on our analysis is a reduction in redistribution if independent taxation systems remain; but the second effect from our analysis is to make the median type more and more likely to prefer the unified system, hence the downward trend of progressivity could at some point be reversed by a spontaneous constitutional reform towards a unified government. However, expansion to include more poor countries shifts those preferences of the median type back, away from unification of fiscal policy. So the expansion decision is something that favors the rich, because they eliminate for the near future the possibility that the median voter will require a unification of fiscal policy in Europe.

Related Literature

Our paper contributes to the literature on optimal income taxation with mobile labor and competition. A general view from this literature is that the ability of individuals to move from one jurisdiction to another imposes additional constraints on the amount of redistribution that each jurisdiction can undertake (see, for example, Wilson, 1980, 1992; Mirrlees, 1982; Bhagwati and Hamada, 1982; Leite-Monteiro, 1997; Hindriks, 1999; and Osmundsen, 1999). More recently, Simula and Trannoy (2006), Wilson (2006), and Krause (2007) study how allowing agent migration affects the optimal nonlinear income tax schedule of a State, taking the other States as outside options with exogenous value. In particular, Simula and Trannoy (2006) 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. In our model, by endogenizing the outside option, we show that such a "curse" of the middle-type agents disappears. Hamilton and Pestieau (2005) provide a general equilibrium analysis of tax competition among a large number of small countries. They consider two skilled types and only one type can move.

To the best of our knowledge, Piaser (2007) and Brett and Weymark (2008) are the only papers that model the strategic interaction between tax authorities as we do. Piaser (2007) analyzes competitive nonlinear taxation between two governments with two types of workers. In order to

analyze the effect of competition on the progressivity of income taxes and say something about the relationship between constitutional choice and the degree of inequality, it is necessary to have at least three types, which we do in our model.

Brett and Weymark (2008) analyze strategic nonlinear tax competition between two governments with a finite number of types of agents. Unlike in our model, they assume perfect mobility so agents are only differentiated along the vertical dimension. They show that there do not exist equilibria in which either the highest type pay positive taxes, or the lowest type receive positive subsidies, which is an illustration of the "race-to-the-bottom" proposition in the context of tax competition with perfect mobility. This result is consistent with ours when the mobility cost parameter $k \to 0$.

The effect of mobility and competition on progressivity has also been analyzed in contexts other than income taxation. For example, 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.[3] Huber (1999) studies a two-country model with two types but without 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 and there is no competition in income tax with mobile labor.

The connection between mobility and redistribution is also studied in Epple and Romer (1991) in the context of local property taxes. Basically they develop a general equilibrium framework in which the population of each local jurisdiction is endogenously determined. Tax rates and redistribution levels are chosen by majority vote of local residents. Voters anticipate changes in housing prices and migration that will occur in response to changes in the local tax rate and level of redistribution.

Finally, 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.

The paper is organized as follows. In section 2 we analyze our base model with three ability types under both the unified and independent taxation regimes. Section 3 analyzes the case of a

---

[3] 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).

continuum of abilities. Section 4 provides concluding remarks with some directions for future research and extensions.

2   The Base Model

Consider two States in a potential Federation – the minimal situation in which we can compare the progressivity of competitive State taxation versus that of a unified Federal tax.[4] Citizens (a continuum) differ in two dimensions. Along what we call the "horizontal" dimension, different citizens have different home States and different moving costs. Specifically, each State i, $i = 1, 2$, has a total measure of 1 original citizens attached to it. The State that a citizen is initially attached to is called her home State. In each State, citizens have heterogeneous moving costs if they move from their home State to the other State.

Let x be the citizen's horizontal type. A citizen of type x incurs a moving cost of $k(1 - x)$, $k > 0$, if and only if she moves from her home State to the other State. So x measures the degree of flexibility of an agent: the smaller is x, the larger is the moving cost, or the greater the attachment to the home State (which can be interpreted as being located closer to the base of her home State in a Hotelling-type model). k is the common factor affecting the moving cost for all the citizens, regardless of their horizontal types; it can also be interpreted as the weight of location preferences in the utility function. The smaller is k, the smaller is the moving cost (given x), or the more intense the competition between the two States, as people put less weight on their location preferences. While x represents a personal cost in adjusting to life in a new State, k can be interpreted as some common component of adjustment cost.

Citizens also differ in the "vertical" dimension. That is, they have different productivities or abilities θ, which is referred to as the vertical type. In the base model we consider three vertical types: 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 the three types in the population are μ_H, μ_M and μ_L, respectively.

We assume that production exhibits constant-returns-to-scale, so that a θ-type individual working l units of time (or exerting an effort l) leads to total production or gross income $Q = \theta l$.

---

[4]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 against decentralized city level taxation.

We assume that neither the vertical type θ nor the horizontal type x is observable to the tax authority. Each State i decides on a tax schedule $T_i(Q)$. Equivalently, each State can offer a menu of contracts, which is a collection of consumption and production pairs (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 first decide on their State of residence (whether to stay in their home State or to move to the other State), and then choose their labor supply (effort level), or equivalently, choose a contract (C, Q) from their State of residence. Suppose an agent of type (θ, x) chooses to stay in her home State and a labor supply (effort) l, leading to gross income $Q = \theta l$ and consumption C; then the utility for such an agent is given by

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

where $u(\cdot)$ is strictly increasing, strictly concave and twice continuously differentiable. Similarly, if an agent of type (θ, x) chooses to move to the other State and chooses contract (C, Q) offered there, then her utility is given by $u(C) - \frac{Q}{\theta} - k(1 - x)$.

Expressing the above game in terms of Q, we can say that given a tax schedule profile $T_1(Q)$ and $T_2(Q)$, agents choose their State of residence and then Q, to maximize $u(Q - T(Q)) - Q/\theta$. It is obvious that the single crossing property only holds along the 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 (Martimort and Stole, 1997 and Peck, 1997). To sidestep this problem, we restrict attention to deterministic contracts.[5] 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 x, 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 simply 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

---

[5] See Rochet and Stole (2002) for a discussion on the restrictions resulting from focusing on deterministic contracts. More general approaches to restore 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), Martimort and Stole (2002), and Page and Monteiro (2003).

of unified taxation, all 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.

As a benchmark, in autarky economy without taxes $(Q = C)$, the optimal consumption $C^*(\theta)$ is characterized by

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

The optimal consumption or gross income does not depend on x in autarky, and each citizen should live in her own home State. Moreover, it is easily verified that $C^*(\theta)$ is strictly increasing in θ.

2.1 Unified Taxation

Under unified taxation, we solve for the tax schedule that maximizes the average utility of all citizens in the Federation. The tax schedule that maximizes such a social welfare function with weights equal to the proportions (percentages) of the vertical types, μ_H, μ_M, and μ_L, is also the income tax system that would be chosen in standard two-party electoral competition models with probabilistic voting (see, e.g., Coughlan, 1986, 1992; and Lindbeck and Weibull, 1987).[6]

Since the two States are identical in terms of the original composition of the population, we focus on the symmetric solution in which each State offers the same menu of contracts and the resulting "market shares" are symmetric.[7]

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 weighted average utility

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

subject to the binding resource constraint

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

and the incentive compatibility (IC) constraints, which basically require that no type has incentive

---

[6] In the continuous type model, the corresponding weighted function will be the density function for the vertical types.

[7] 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.

to mimic any of the other types. With three types there will be 6 inequality conditions:

$$u(C_H) - \frac{Q_H}{\theta_H} \geq u(C_M) - \frac{Q_M}{\theta_H} \quad (DIC_{HM})$$

$$u(C_M) - \frac{Q_M}{\theta_M} \geq u(C_L) - \frac{Q_L}{\theta_M} \quad (DIC_{ML})$$

$$u(C_H) - \frac{Q_H}{\theta_H} \geq u(C_L) - \frac{Q_L}{\theta_H} \quad (DIC_{HL}) \tag{2}$$

$$u(C_M) - \frac{Q_M}{\theta_M} \geq u(C_H) - \frac{Q_H}{\theta_M} \quad (UIC_{MH})$$

$$u(C_L) - \frac{Q_L}{\theta_L} \geq u(C_M) - \frac{Q_M}{\theta_L} \quad (UIC_{LM})$$

$$u(C_L) - \frac{Q_L}{\theta_L} \geq u(C_H) - \frac{Q_H}{\theta_L} \quad (UIC_{LH})$$

Working with all 6 inequalities can be quite tedious. It turns out that with a monotonicity constraint $Q_H \geq Q_M \geq Q_L$ (which implies $C_H \geq C_M \geq C_L$), only the two local DIC's bind:

LEMMA 1 *The set of IC constraints under unified taxation is equivalent to the monotonicity constraint* $Q_H \geq Q_M \geq Q_L$, *and the following two local downward IC conditions:*

$$u(C_H) - \frac{Q_H}{\theta_H} = u(C_M) - \frac{Q_M}{\theta_H}, \quad (DIC\text{-}H) \tag{3}$$

$$u(C_M) - \frac{Q_M}{\theta_M} = u(C_L) - \frac{Q_L}{\theta_M} \quad (DIC\text{-}M)$$

Proof. See Appendix. ∎

We will solve the relaxed program by ignoring the monotonicity constraint (we shall do the consistency check after we have obtained the solutions). For the Lagrangian let the multipliers of (DIC-H), (DIC-M) and (RC) be λ_H, λ_M and λ_R respectively. The first order conditions can be written as follows:

$$\frac{\partial L}{\partial Q_H} = -\frac{\mu_H}{\theta_H} - \frac{\lambda_H}{\theta_H} + \mu_H \lambda_R = 0$$

$$\frac{\partial L}{Q_M} = -\frac{\mu_M}{\theta_M} + \frac{\lambda_H}{\theta_H} - \frac{\lambda_M}{\theta_M} + \mu_M \lambda_R = 0$$

$$\frac{\partial L}{\partial Q_L} = -\frac{\mu_L}{\theta_L} + \frac{\lambda_M}{\theta_M} + \mu_L \lambda_R = 0$$

$$\frac{\partial L}{\partial C_H} = \mu_H u'(C_H) + \lambda_H u'(C_H) - \mu_H \lambda_R = 0$$

$$\frac{\partial L}{\partial C_M} = \mu_M u'(C_M) - \lambda_H u'(C_M) + \lambda_M u'(C_M) - \mu_M \lambda_R = 0$$

$$\frac{\partial L}{\partial C_L} = \mu_L u'(C_L) - \lambda_M u'(C_L) - \mu_L \lambda_R = 0$$

10

From the above equations, we can obtain

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

$$\lambda_R = \frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}; \ \lambda_H = \mu_H(\theta_H \lambda_R - 1); \ \lambda_M = \mu_L \theta_M \left(\frac{1}{\theta_L} - \lambda_R\right)$$

$$u'(C_L^U) = \frac{\lambda_R}{(1 - \frac{\theta_M}{\theta_L}) + \theta_M \lambda_R}; \ u'(C_M^U) = \frac{\mu_M \lambda_R}{\mu_M - \lambda_H + \lambda_M}$$

First of all, it is clear that the solution does not depend on k, the mobility parameter, as a direct consequence of our focus on symmetric solution. Second, it can be verified that $u'(C_M^U) > 1/\theta_M$ and $u'(C_L^U) > 1/\theta_L$. Thus $C_M^U < C_M^*$, and $C_L^U < C_L^*$ (due to the concavity of $u(\cdot)$), i.e., 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. Moreover, since $C_M^U < C_M^* < C_H^* = C_H^U$, type M and type H never pool in the optimal solution.

LEMMA 2 *In the optimal solution under unified taxation, $T_H > T_M \geqslant T_L$.*

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

$$u(C_H) - u(C_M) = \frac{Q_H - Q_M}{\theta_H} \leq \frac{C_H - C_M}{\theta_H}$$

$$\Rightarrow \ u(C_H) - \frac{C_H}{\theta_H} \leq u(C_M) - \frac{C_M}{\theta_H}.$$

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 $T_M < T_L$, that is, $Q_M - C_M < Q_L - C_L$. By the binding DIC-M,

$$u(C_M) - u(C_L) = \frac{Q_M - Q_L}{\theta_M} < \frac{C_M - C_L}{\theta_M}$$

$$\Rightarrow \ u(C_M) - \frac{C_M}{\theta_M} < u(C_L) - \frac{C_L}{\theta_M}.$$

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 $C_L \leqslant C_M$, we have $u(C_M) - \frac{C_M}{\theta_M} \geqslant u(C_L) - \frac{C_L}{\theta_M}$. A contradiction. Thus we must have $T_M \geqslant T_L$. ∎

Given $T_H > T_M \geqslant T_L$, by (RC) we must have $T_H > 0$: 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. So under a unified regime, while the rich always pay taxes and the poor receive subsidies, the middle class may pay taxes or receive subsidies.

2.2 Independent Taxation

Under the independent taxation regime, each State chooses its taxation schedule simultaneously and independently to maximize the weighted average utility of the citizens residing in its own State, given the other State's taxation schedule. Given that the two States are identical, we focus on symmetric Nash equilibria in which both States choose the same taxation schedule.

Since everyone is required to participate in one of the tax systems, the individual rationality constraint only concerns which State to live in. Let $v_j \equiv u(C_j) - Q_j/\theta_j$ be the rent provision to type θ_j citizen who accepts contract (Q_j, C_j). Suppose the other State's taxation rule leads to rent provisions v_j^*, $j = H, M, L$. Then a citizen with vertical type θ_j and horizontal type x will stay with her home State if and only if

$$v_j \geqslant v_j^* - k\,(1-x) \text{ or } x \leq \min\left\{1 + \frac{v_j - v_j^*}{k}, 1\right\}.$$

When $v_j \geq v_j^*$, all the type-θ_j citizens in the State in question will stay with their home State, and all the types (θ_j, x) where $x \geq 1 - \left(v_j - v_j^*\right)/k$ in the other State will move to the State in question. Therefore for vertical type θ_j, the total measure of horizontal types that will reside in the State in question will be $1 + \left(v_j - v_j^*\right)/k$.[8] For this reason, x_j defined below can be regarded as the "market share" of type θ_j, $j = H, M, L$, for the State in question:

$$x_j = 1 + \frac{v_j - v_j^*}{k}. \tag{4}$$

The objective of the State in question is to maximize $\mu_H v_H + \mu_M v_M + \mu_L v_L$, subject to the appropriate resource constraint and the incentive compatibility constraints. The resource constraint is given by

$$\mu_H x_H (Q_H - C_H) + \mu_M x_M (Q_M - C_M) + \mu_L x_L (Q_L - C_L) = 0,$$

where x_j's are given by (4).

It turns out that the IC constraints under independent taxation are much more involved than in the unified taxation case.

Like in the first two steps in the proof of Lemma 1, the 6 IC's (2) can be reduced to 4 local IC's $(\text{DIC}_{HM}, \text{DIC}_{ML}, \text{UIC}_{MH}, \text{and UIC}_{LM})$ plus the monotonicity constraint $Q_H \geq Q_M \geq Q_L$.

---

[8] Apparently this expression also applies when $v_j < v_j^*$.

We then argue that UIC's cannot bind so these two constraints can be dropped. Given each State's objective function, each State has incentive to redistribute as much as possible.[9] But this is restricted by the DIC's. With independent taxation, each State tries to steal the high types from the other State. The purpose of this move is not to attract high types *per se,* but to increase its total tax revenue from high types. Given that redistribution is only restricted by DIC's, UIC's should not bind in equilibrium.

LEMMA 3 *Under independent taxation, the UIC's are inactive.*

Proof. See Appendix. ∎

Given that UIC's can be dropped, the State in question has the following programming problem:

$$\max \mu_H v_H + \mu_M v_M + \mu_L v_L$$
$$u(C_H) - \frac{Q_H}{\theta_H} \geq u(C_M) - \frac{Q_M}{\theta_H}; \ u(C_M) - \frac{Q_M}{\theta_M} \geq u(C_L) - \frac{Q_L}{\theta_M}$$
$$\mu_H x_H (Q_H - C_H) + \mu_M x_M (Q_M - C_M) + \mu_L x_L (Q_L - C_L) = 0$$
$$Q_H \geq Q_M \geq Q_L$$

where x_j's are given by (4).

Unlike in the unified taxation case, under independent taxation the DIC's may not bind simultaneously.[10] One or both DIC's may not bind since two States are competing for higher type agents under independent taxation. The rent provision for H type now depends on two forces: competition in the horizontal dimension and self-selection (sorting) in the vertical dimension. If competition is strong on the horizontal dimension, then H type will secure high rent anyway, which makes sorting in the vertical dimension automatically satisfied and the DIC's not binding. Hence we need to cover multiple cases.

Case 1: Both DIC's bind. 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

---

[9] In the complete information benchmark, it is easily seen that given the concavity of the utility function, the solution would have only the high type working, redistributing income to the other types.

[10] The argument showing that the DIC's must bind under unified taxation does not work here. To see this, suppose in a candidate symmetric equilibrium DIC(H) does not bind. Now if State 1 increases Q_H and decreases Q_M by the same amount, this might lead to budget deficit for State 1, as some H type will move to State 2 and some M type will move to State 1. Under unified taxation, the central authority can change the tax schedules of two States *simultaneously,* but this is not feasible under independent taxation.

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

$$-\frac{\mu_H}{\theta_H} - \frac{\lambda_H}{\theta_H} + \mu_H \lambda_R \left[1 - \frac{T_H}{k\theta_H}\right] = 0$$

$$-\frac{\mu_M}{\theta_M} + \frac{\lambda_H}{\theta_H} - \frac{\lambda_M}{\theta_M} + \mu_M \lambda_R \left[1 - \frac{T_M}{k\theta_M}\right] = 0$$

$$-\frac{\mu_L}{\theta_L} + \frac{\lambda_M}{\theta_M} + \mu_L \lambda_R \left[1 - \frac{T_L}{k\theta_L}\right] = 0$$

$$\mu_H u'(C_H) + \lambda_H u'(C_H) + \mu_H \lambda_R \left[-1 + T_H \frac{u'(C_H)}{k}\right] = 0$$

$$\mu_M u'(C_M) - \lambda_H u'(C_M) + \lambda_M u'(C_M) + \mu_M \lambda_R \left[-1 + T_M \frac{u'(C_M)}{k}\right] = 0$$

$$\mu_L u'(C_L) - \lambda_M u'(C_L) + \mu_L \lambda_R \left[-1 + T_L \frac{u'(C_L)}{k}\right] = 0$$

From the above equations, we obtain

$$
\begin{aligned}
u'(C_H) &= 1/\theta_H, \\
\lambda_H &= \mu_H \left[-1 + \theta_H \lambda_R \left(1 - \frac{T_H}{k\theta_H}\right)\right] \\
\lambda_M &= \mu_L \left[\frac{\theta_M}{\theta_L} - \theta_M \lambda_R \left(1 - \frac{T_L}{k\theta_L}\right)\right] \\
\lambda_R &= \frac{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}}{1 - [T_H \frac{\mu_H}{k\theta_H} + T_M \frac{\mu_M}{k\theta_M} + T_L \frac{\mu_L}{k\theta_L}]} \\
u'(C_M) &= \frac{\mu_M \lambda_R}{\mu_M - \lambda_H + \lambda_M + \mu_M \lambda_R T_M/k} \\
u'(C_L) &= \frac{\mu_L \lambda_R}{\mu_L - \lambda_M + \mu_L \lambda_R T_L/k}
\end{aligned}
\tag{5}
$$

As in the unified taxation case, it can be verified that $u'(C_M) > 1/\theta_M$, and $u'(C_L) > 1/\theta_L$. Therefore, compared to the autarky case there is no distortion at the top, but the consumptions of type M and type L are both distorted downward. Moreover, following exactly the arguments paralleling those in the proof of Lemma 2, we have $T_H^I > T_M^I \geq T_L^I$. As a result, $T_H^I > 0$, $T_L^I < 0$ and the sign of T_M^I is ambiguous.

Case 2: Neither DIC binds. If neither DIC binds, we have $\lambda_H = \lambda_M = 0$. From the first order conditions, we get

$$
\begin{aligned}
u'(C_H^I) &= \frac{1}{\theta_H} < u'(C_M^I) = \frac{1}{\theta_M} < u'(C_L^I) = \frac{1}{\theta_L}; \\
T_H &= k\mu_M(\theta_H - \theta_M) + k\mu_L(\theta_H - \theta_L); \\
T_M &= T_H - k(\theta_H - \theta_M); \quad T_L = T_M - k(\theta_M - \theta_L).
\end{aligned}
$$

Clearly, consumption is no longer distorted: $C_j^I = C_j^*$, $j = L, M, H$. Moreover, $T_H > T_M > T_L$.[11] Now the DIC's can be rewritten as:

$$u(C_H^*) - \frac{C_H^*}{\theta_H} - \frac{T_H}{\theta_H} \geq u(C_M^*) - \frac{C_M^*}{\theta_H} - \frac{T_M}{\theta_H};$$
$$u(C_M^*) - \frac{C_M^*}{\theta_M} - \frac{T_M}{\theta_M} \geq u(C_L^*) - \frac{C_L^*}{\theta_M} - \frac{T_L}{\theta_H}.$$

From the above inequalities we can see that if k is small enough, the difference between T_H and T_M and that between T_M and T_L will be sufficiently small. As a result, the DIC's will not bind as k is sufficiently small. In the limit as $k \to 0$, T_H, T_M and T_L all go to zero. This is consistent with Brett and Weymark (2008), who show in a model with perfectly mobile agents (that is, $k = 0$ in our model), that there does not exist any equilibrium in which the highest type pays positive taxes, or the lowest type receives positive subsidies under competitive taxation.

Case 3: One DIC binds and the other does not. Here we only consider the case when DIC-H is slack but DIC-M binds (the analysis for the other case is similar). In this case, we have $\lambda_H = 0$ and $\lambda_M > 0$. Based on the first order conditions, one can show that

$$u'(C_H^I) = \frac{1}{\theta_H}; \ u'(C_M^I) = \frac{1}{\theta_M}; \ u'(C_L^I) > \frac{1}{\theta_L}.$$

That is, there is no consumption distortion for types H and M, but the consumption of type L is distorted downward.[12] The expressions for λ_M, λ_R and $u'(C_L^I)$ are the same as those in (5). In this case, we have $T_H > T_M > T_L$.

To summarize, we have the following lemma:

LEMMA 4 *In the symmetric equilibrium under independent taxation, $T_H > T_M \geq T_L$.*

So as in the unified taxation case, in the equilibrium of the competitive taxation regime, the rich pay taxes, and the poor receive subsidies. The middle class, however, may pay taxes or receive subsidies. We now turn to the comparisons of the two taxation systems.

---

[11] So C_j's are the same as in the autarky case, though Q_j's are different.

[12] In the opposite case that DIC-M binds and DIC-H is slack, we can show that both C_H^I and C_L^I have no distortion but C_M^I is distorted downward.

2.3 Comparison

Our first comparison result shows that competition increases consumption for both the middle class and the poor (while the consumption stays the same or undistorted for the rich).

PROPOSITION 1 $C_L^I > C_L^U$ and $C_M^I > C_M^U$: competition increases consumption for both types M and L.

Proof. See Appendix. ∎

The proof is rather tedious, and is therefore relegated to the appendix. Although the consumption for the rich stays the same (undistorted), the induced productivity or income Q is different under the taxation regimes.

PROPOSITION 2 $T_H^I < T_H^U$ and $v_H^I > v_H^U$. That is, type H pays lower taxes and is better off under independent taxation.

Proof. We first consider the case that both DIC's bind under independent taxation. By RC, DIC-H, and DIC-M, we have

$$
\begin{aligned}
Q_H &= \mu_H C_H + \mu_M C_M + (1 - \mu_H)\theta_H[u(C_H) - u(C_M)] \\
&\quad + \mu_L[C_L + \theta_M(u(C_M) - u(C_L))]
\end{aligned}
\tag{6}
$$

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 Proposition 1. Then from (6), we have

$$
\Delta Q_H = \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H)\theta_H - \mu_L \theta_M]\Delta u(C_M) - \mu_L \theta_M \Delta u(C_L)
\tag{7}
$$

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

$$
\begin{aligned}
\Delta Q_H &\leq \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H)\theta_H - \mu_L \theta_M]u'(C_M^I)\Delta C_M - \mu_L \theta_M u'(C_L^I)\Delta C_L \\
&< \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H)\theta_H - \mu_L \theta_M]\frac{1}{\theta_M}\Delta C_M - \mu_L \theta_M \frac{1}{\theta_L}\Delta C_L \\
&= (1 - \mu_H)\frac{\theta_M - \theta_H}{\theta_M}\Delta C_M + \mu_L \left(\frac{\theta_L - \theta_M}{\theta_L}\right)\Delta C_L < 0 = \Delta C_H
\end{aligned}
\tag{8}
$$

Thus $T_H^I < T_H^U$. Given that $C_H^I = C_H^U$, it follows that $v_H^I > v_H^U$.

Next, we consider the case that neither DIC binds under independent taxation. The equation (6) still holds under unified taxation. By the nonbinding DIC's, under independent taxation, the LHS

is strictly less than RHS in (6). As a result, the LHS is strictly less than RHS in (7). The rest of the proof is the same as in the previous case, except that in (8) the first inequality is replaced by a strict inequality, and the second inequality is replaced by an equality.

For the case that DIC-H is slack but DIC-M binds, the proof is exactly the same as in the case that neither DIC binds under independent taxation. ∎

PROPOSITION 3 $T_L^I > T_L^U$ and $v_L^I < v_L^U$. That is, type L receives less subsidies and is worse off under independent taxation.

Proof. Again, we first consider the case that both DIC's bind under independent taxation. Suppose $T_L^I \leq T_L^U$. Then $Q_L^I - Q_L^U \leq C_L^I - C_L^U$.

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

The first inequality is due to the fact that $Q_L^I - Q_L^U \leq C_L^I - C_L^U$. The second equality follows from the intermediate value theorem, where $C^* \in \left[C_L^U, C_L^I\right]$. The last inequality holds since $u'(C^*) \geq u'(C_L^I) > 1/\theta_L$. 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

$$
\begin{aligned}
v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\
&= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\
&\geq u(C_L^I) - u(C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} \\
&= u'(C^*)(C_L^I - C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} > 0,
\end{aligned}
\tag{9}
$$

where the last inequality follows from the fact that $u'(C^*) \geq u'(C_L^I) > 1/\theta_L > 1/\theta_M$. Thus $v_M^I > v_M^U$. Since we have already established that $C_H^I = C_H^U$ and $Q_H^I < Q_H^U$, we have $v_H^I > v_H^U$. The tax schedules under independent taxation $\left\{\left(Q_j^I, C_j^I\right)\right\}_{j \in \{H,M,L\}}$ 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 $\left\{\left(Q_j^U, C_j^U\right)\right\}_{j \in \{H,M,L\}}$ are the optimal solution for unified taxation.

17

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 $v_L^I \geq v_L^U$. This implies that

$$
\begin{aligned}
v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\
&= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\
&> u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_L} \\
&= v_L^I - v_L^U \geq 0.
\end{aligned}
\tag{10}
$$

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 $\left\{ \left(Q_j^I, C_j^I \right) \right\}_{j \in \{H,M,L\}}$ is feasible under unified taxation but the optimal solution is $\left\{ \left(Q_j^U, C_j^U \right) \right\}_{j \in \{H,M,L\}}$. Therefore, we must have $v_L^I < v_L^U$.

Next, we consider the case that neither DIC binds under independent taxation. The proof is very similar to that for the case with binding DIC's. We first show $T_L^I > T_L^U$. Suppose in negation $T_L^I \leq T_L^U$. Then following the same steps above, we can obtain the expressions for $v_L^I - v_L^U$ and $v_M^I - v_M^U$ (now the second equality in (9) should be replaced by a strict inequality, due to the strict inequality of DIC-M). Again the same contradiction can be reached. To show $v_L^I < v_L^U$, we follow similar steps as before. The only change in the proof is that the second equality in (10) should be replaced by a strict inequality.

Finally, consider the case that DIC-H is slack but DIC-M binds under independent taxation. Given that DIC-M binds, the proof is exactly the same as in the case when both DIC's are binding. ∎

Since the tax schedules under independent taxation $\left\{ \left(Q_j^I, C_j^I \right) \right\}_{j \in \{H,M,L\}}$ are also feasible under unified taxation, we have the following corollary:

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

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 (given that the rich always prefer the independent taxation system and the poor always prefer the unified taxation system). It is impossible to obtain general analytical results on the preferences of

the middle type as a function of relative productivities (distribution of θ's) and income distribution (distribution of μ's). However, the computations we now turn to, provide interesting results.[13]

2.4 Constitutional Choice

We now augment our previous model with a constitutional stage, where the taxation regime is decided by simple majority rule. That is, a taxation system (unified or independent) is chosen as long as more than 50% of citizens are in favor of that taxation system. The time line is now as follows. In the first period, all citizens in the Federation vote on the taxation regimes. In the second period, tax schedules for both States are proposed either by the Federation or by the two States independently and simultaneously (depending on the outcome from the constitutional stage). In the third period, citizens first decide on which State to reside and then decide on their labor supply/effort levels (or equivalently, decide on which contract (Q, C) to accept from the residence State). We assume that $\mu_i < 1/2$ for $i = H, L$. So the constitutional choice will be determined by the preference of the middle class.

Fix $\theta_H = 2$ and $\theta_L = 1$ for all the numerical computations in this section. Our computations first show, for any percentage of each type, that

OBSERVATION 1 *There exists a cutoff $\theta_M^* \in (\theta_L, \theta_H)$ such that type M prefers the independent taxation system if and only if her type is higher than θ_M^*.*

Our computations also show that given $\theta_M \in (1, 2)$ and $\mu_H = \mu_L = (1 - \mu_M)/2$:[14]

OBSERVATION 2 *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.

These results have an important implication in terms of welfare. Assume $\theta_H = 2, \theta_L = 1$ and $\theta_M = 1.51$. We can compute μ_M^* by keeping $\mu_H = \mu_L$. We can then compare the 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 sufficiently small (by Corollary 1 above). This means that

---

[13]Detailed computations and Matlab code used in this project are available upon request.

[14]When μ_M increases, we let μ_H and μ_L go down by the same compensating amount.

COROLLARY 2 *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 making at the constitutional stage.*

Another interesting observation comes from the following exercise: fix θ_M and μ_M (or μ_L); then our computations show that

OBSERVATION 3 *There exists μ_L^* such that type M prefers independent taxation if and only if $\mu_L > \mu_L^*$.*

This is very intuitive: as the percentage of the poor goes up, the fear for having to support the poor increases and the middle type becomes more likely to prefer the independent tax regime.

Our computations also reveal some less intuitive relationships between initial conditions and constitutional preferences by the middle type:

OBSERVATION 4 *Both θ_M^* and μ_M^* are decreasing in k.*

This suggests that when k decreases, for a given θ_M or μ_M, the middle type is more likely to prefer the unified taxation system. The schedules $\theta_M^*(k)$ and $\mu_M^*(k)$ are shown in Figure 1 below, 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 $\mu_H = \mu_L$, and $\theta_M = 1.3$.[15]
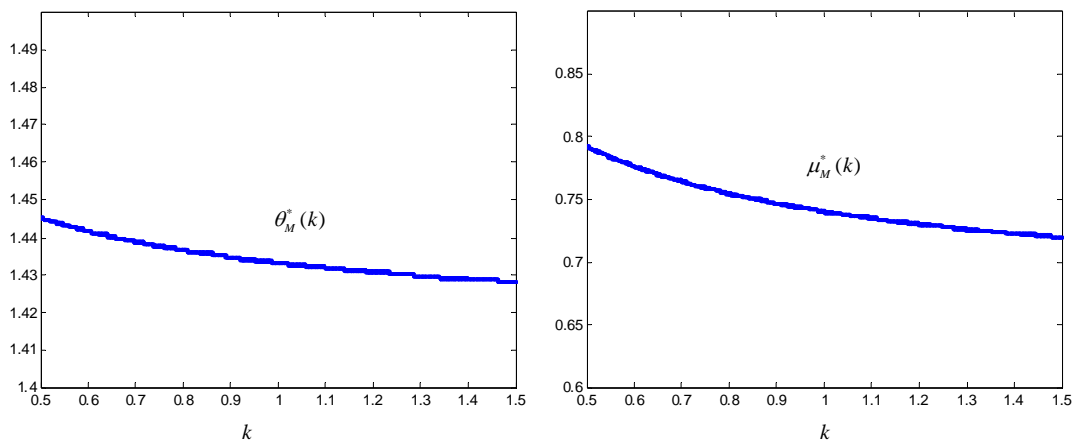

Figure 1: Schedules of $\theta_M^*(k)$ and $\mu_M^*(k)$

---

[15]Our computations show that the higher the selected value for θ_M, the lower the schedule of $\mu_M^*(k)$. This is consistent with our Observations 1 and 2.

An intuition for observation 4 is as follows: under both taxation regimes the middle class "benefits" from the existence of richer citizens who pay more taxes and "suffers" from the existence of poorer citizens who need to receive subsidies; under unified taxation these two effects do not depend on k, while under independent taxation when k goes down the "benefits" mentioned above go down, since the rich secures higher rents as the competition between two States becomes more intense. Given that there is no such competition effect for the poor, the relative attractiveness of the two regimes to the middle class must therefore change in the direction of a more likely preference for the unified system. The intuition for the monotonicity of $\mu_M^*(k)$ is similar: when k goes down, the previously indifferent type between the two systems should prefer the unified regime, and indifference can be restored if the middle class is larger, to compensate in terms of per capita share of the transfers to the poor.

In a picture with μ_M on horizontal axis and θ_M^* on vertical axis, our computations show that:

OBSERVATION 5 θ_M^* *decreases as* μ_M *increases (while the other two types decrease symmetrically at the same time).*

Figure 2 below is plotted with $k = 1$. 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.[16]

---

[16]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^*.
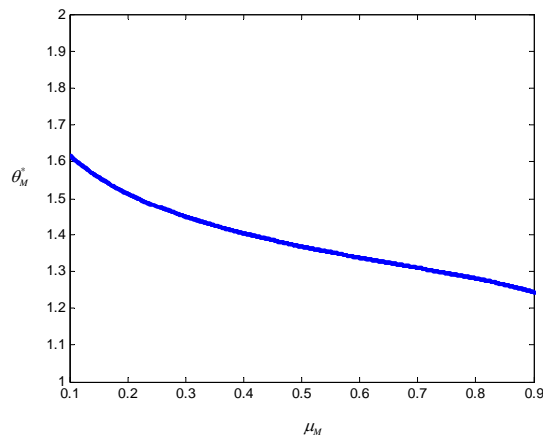
Figure 2: Schedule of $\theta_M^*(\mu_M)$

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 when studying the case of a continuum of ability types.

In summary, 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 mobility sharply increased in the 90's and 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.[17]

---

[17]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; but a majority in each State, if required, would be more difficult than before to materialize, since the median voters of the two richer countries would be against supporting also the more poor people of the new country added to the Union. All these issues are for future research and applications of the ideas in this paper.

3 The Continuous-type Model

In this section we extend our analysis to the continuous type case, which can be regarded as the limiting case of many finite types. 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 its 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 is determined by (1).

Again we will consider unified and independent taxation rules. Under either the unified or independent taxation rule, incentive compatibility has to hold for each type of citizen conditional on her State of residence. Define

$$V(\theta, \widehat{\theta}) \;= u(C(\widehat{\theta})) - \frac{Q(\widehat{\theta})}{\theta}$$

to be the utility for a citizen with (vertical) type θ who accepts contract $\{C(\widehat{\theta}), Q(\widehat{\theta})\}$. Incentive compatibility requires that

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

Let $v(\theta)$ denote the equilibrium rent provision to type-θ citizen: $v(\theta) = V(\theta, \theta)$. By the standard Constraint Simplification Theorem, the IC conditions are equivalent to the following two conditions:

$$v'(\theta) \;\; = \;\; \frac{Q(\theta)}{\theta^2} = \frac{1}{\theta}[u(C(\theta)) - v(\theta)] \tag{11}$$

$$Q'(\theta) \;\; \geq \;\; 0 \tag{12}$$

Constraint (12) is the monotonicity requirement as in the three-type model.

By (11), given $v(\theta)$, $Q(\theta)$ is uniquely determined and so is $C(\theta)$. For convenience, we will work with the rent provision contract $v(\theta)$.[18] It can be easily verified that $Q' = \theta u'(C)C'$. Thus, as in the three-type model, $Q'(\theta) \geq 0$ if and only if $C'(\theta) \geq 0$.

Given $v(\theta)$ provided by the State in question and the other State's rent provision $v_{-i}(\theta)$, the type-θ "market share" for the State in question is given by

$$x^*(\theta) = 1 + \frac{v(\theta) - v_{-i}(\theta)}{k}. \tag{13}$$

---

[18] This approach follows the lead of Armstrong and Vickers (2001), who model firms as supplying utility directly to consumers.

For ease of analysis, from now on we will work with the utility function $u(C) = 2\sqrt{C}$.[19]

3.1 Unified Taxation

Under unified taxation, the objective of the Federal authority is to maximize the weighted average utility of all the citizens in both States, where the weight function $w(\theta) = f(\theta)$, the density of the vertical types.[20] We focus on the symmetric solution in which the same menu of contracts is applied to both States and the resulting "market shares" are symmetric (no citizen moves). We can thus drop the State index to write $\{C_i(\theta), Q_i(\theta)\} = \{C(\theta), Q(\theta)\}$, $i = 1, 2$. Mathematically, this can be formulated as an optimal control problem:

$$\max \int_{\underline{\theta}}^{\overline{\theta}} v(\theta) f(\theta) d\theta$$

$$\text{s.t.} \quad v'(\theta) = \frac{1}{\theta} \left[2\sqrt{C(\theta)} - v(\theta) \right]$$

$$Q'(\theta) \geq 0$$

$$\int_{\underline{\theta}}^{\overline{\theta}} [Q(\theta) - C(\theta)] f(\theta) d\theta = 0$$

The last constraint is the resource or budget constraint (RC).

To solve this optimal control problem, as is standard in the literature, we first ignore the monotonicity constraint on $Q(\theta)$ to consider the relaxed program (and this approach will be justified if the solution of $Q(\theta)$ is indeed monotone). To deal with the resource constraint, we define the new state variable $J(\theta)$ as follows

$$J(\theta) = \int_{\underline{\theta}}^{\theta} [Q(\theta) - C(\theta)] f(\theta) d\theta, \text{ hence}$$

$$J'(\theta) = [Q(\theta) - C(\theta)] f(\theta).$$

Now (RC) is equivalent to $J(\overline{\theta}) = 0$ and $J(\underline{\theta}) = 0$. The Hamiltonian of the problem is:

$$H = vf + \lambda \frac{1}{\theta} \left[2\sqrt{C} - v \right] + \mu[\theta(2\sqrt{C} - v) - C]$$

---

[19] Our main results should not be altered as long as we work with concave utility functions.

[20] Theorem 6.3 in Coughlan (1992), establishing that a tax schedule maximizes the weighted average utility if and only if it is the equilibrium tax schedule in a standard two party electoral competition, has been extended to function spaces by Chen (2000).

Define $z = \sqrt{C}$, then the Hamiltonian can be rewritten as

$$H = vf + \lambda \frac{1}{\theta} [2z - v] + \mu[\theta(2z - v) - z^2]$$

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

$$\frac{\partial H}{\partial z} = 2\frac{\lambda}{\theta} + \mu[2\theta - 2z]f = 0 \tag{14}$$

$$\lambda' = -\frac{\partial H}{\partial v} = -f + \frac{\lambda}{\theta} + \mu\theta f \tag{15}$$

$$\mu' = -\frac{\partial H}{\partial J} = 0 \tag{16}$$

From (16), μ is a constant. From (14) and (15) we can get rid of λ to yield

$$z' + \frac{f'}{f} z = 2 - \frac{1}{\mu\theta} + \frac{f'}{f}\theta \tag{17}$$

We can further getting rid of μ by turning (17) into a second-order differential equation:

$$z'' = -\frac{1}{\theta}\left[z' - 2 + (z + \theta z' - 2\theta)\frac{f'}{f} + \theta(z - \theta)\left(\frac{f'}{f}\right)'\right] \tag{18}$$

$$z(\underline{\theta}) = \underline{\theta}, \; z(\overline{\theta}) = \overline{\theta}$$

where the boundary conditions above are directly implied from the transversality conditions $\lambda(\underline{\theta}) = \lambda(\overline{\theta}) = 0$ and (14). The above second-order (linear) differential equation system has a closed-form solution, which is given by

$$z(\theta) = \frac{f(\underline{\theta})}{f(\theta)}\left[\int_{\underline{\theta}}^{\theta} \frac{f(s)}{f(\underline{\theta})}\left(2 - \frac{1}{\mu s} + s\frac{f'(s)}{f(s)}\right) ds + \underline{\theta}\right], \tag{19}$$

$$where \; \mu = \int_{\underline{\theta}}^{\overline{\theta}} \frac{dF(s)}{s}.$$

3.2 Independent Taxation

Under the independent taxation regime, each State i chooses its taxation schedule simultaneously and independently. Given $v_{-i}(\theta)$, the rent provision provided by the other State, State i will choose a rent provision $v(\theta)$ to maximize the weighted average utility of the citizens residing in its own State.

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 $v(\theta)$, by (13) the type-θ "market share" for State 1 is given by $\eta(\theta) = 1 + \frac{1}{k}[v(\theta) - v^*(\theta)]$.

Now State 1's maximization problem can be formulated as the following optimal control problem:

$$\max \int_{\underline{\theta}}^{\overline{\theta}} v(\theta) f(\theta) d\theta$$

$$\text{s.t. } v'(\theta) = \frac{1}{\theta}\left[2\sqrt{C(\theta)} - v(\theta)\right]$$

$$Q'(\theta) \geq 0$$

$$J'(\theta) = [\theta(2\sqrt{C(\theta)} - v(\theta)) - C(\theta)]\eta(\theta)f(\theta)$$

$$J(\underline{\theta}) = 0, \ J(\overline{\theta}) = 0$$

where $J(\theta) = \int_{\underline{\theta}}^{\theta}[\theta(2\sqrt{C} - v) - C]\eta(\theta)f(\theta)d\theta$ is the state variable associated with the budget constraint. Note that the market share $\eta(\theta)$ does not directly enter the State's objective function. However, the States compete for high-type citizens as the market shares affect the resource constraints and hence the ability to redistribute.

We again drop the monotonicity constraint $Q'(\theta) \geq 0$ and define the Hamiltonian (with $z = \sqrt{C}$):

$$H = vf + \frac{\lambda}{\theta}(2z - v) + \mu\eta[\theta(2z - v) - z^2]f.$$

The optimality conditions for a symmetric equilibrium are given by

$$\frac{\partial H}{\partial z} = 2\frac{\lambda}{\theta} + \mu[2\theta - 2z]f = 0$$

$$\lambda'(\theta) = -\frac{\partial H}{\partial v} = -f - 2\frac{\lambda}{\theta} - \frac{\mu}{k}\left[\theta(2z - v) - z^2\right]f + \mu\theta f$$

$$\mu'(\theta) = -\frac{\partial H}{\partial J} = 0 \Rightarrow \mu \ is \ a \ constant$$

After getting rid of λ, we have:

$$z' = 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{\theta(2z - v) - z^2}{k\theta}$$

$$v' = \frac{1}{\theta}(2z - v)$$

$$J' = \theta(2z - v) - z^2$$

Letting $w = 2z - v$, the above system becomes

$$w' = 2z' - v' = 2z' - \frac{w}{\theta} \tag{20}$$

$$J' = \theta w - z^2 \tag{21}$$

$$z' = 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{\theta w - z^2}{k\theta} = 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{J'}{k\theta} \tag{22}$$

From (21), we have

$$w = \frac{1}{\theta}(J' + z^2), \tag{23}$$

$$w' = \frac{1}{\theta^2}\left[(J'' + 2zz')\theta - (J' + z^2)\right] \tag{24}$$

Substituting (23) and (24) into (20), we have

$$J'' = 2(\theta - z)z' \tag{25}$$

From (22), we have

$$J'' = 2k - k(\theta z'' + z') - k(z + \theta z' - 2\theta)\frac{f'}{f} - k\theta(\theta - z)\left(\frac{f'}{f}\right)' \tag{26}$$

Equating (25) and (26), and simplifying, we have

$$z'' = -\frac{1}{\theta}\left[z' - 2 + (z + \theta z' - 2\theta)\frac{f'}{f} + \theta(z - \theta)\left(\frac{f'}{f}\right)' + \frac{2}{k}(\theta - z)z'\right] \tag{27}$$

$$z(\underline{\theta}) = \underline{\theta}, \ z(\overline{\theta}) = \overline{\theta}$$

where the boundary conditions above, as in the unified taxation case, follow from the transversality conditions $\lambda(\underline{\theta}) = \lambda(\overline{\theta}) = 0$. Note that this is again a second-order differential equation system with two boundary values. It is nonlinear, however, in this case. The complication is that a closed-form solution is no longer available. The analysis can easily become intractable if we work with general distributions. For this reason in the next subsection we will focus on the uniform distribution case, where θ is distributed uniformly over $[\underline{\theta}, \overline{\theta}]$.

3.3 The Uniform Distribution Case

Under unified taxation, assuming that θ is uniformly distributed (i.e., $f' = 0$), (18) reduces to

$$z'' = -\frac{1}{\theta}\left[z' - 2\right] \tag{28}$$

$$z(\underline{\theta}) = \underline{\theta}, \; z(\overline{\theta}) = \overline{\theta}$$

Substituting $f(\theta) = 1/\left(\overline{\theta} - \underline{\theta}\right)$ into (19), we obtain the solution in the uniform distribution case:

$$z(\theta) = 2\theta - (\overline{\theta} - \underline{\theta})\frac{\log\theta - \log\underline{\theta}}{\log\overline{\theta} - \log\underline{\theta}} - \underline{\theta} \tag{29}$$

It can be easily verified that $z'(\theta) > 0$ if $\overline{\theta}/\underline{\theta} - 1 \le 2\log\left(\overline{\theta}/\underline{\theta}\right)$, or equivalently,

$$\overline{\theta}/\underline{\theta} \le \gamma^* \approx 3.55 \tag{30}$$

Note that $z'(\theta) > 0$ implies that $Q'(\theta) > 0$. Given our focus on perfect sorting equilibria and to justify our approach to solve the relaxed program by ignoring the monotonicity constraint, we maintain the sorting condition (30) throughout this section.[21] Intuitively, the higher the $\overline{\theta}/\underline{\theta}$, the more costly is sorting along the vertical dimension. When $\overline{\theta}/\underline{\theta}$ is large enough, pooling at the lower end is optimal.

It can be easily verified that $\theta - z > 0$ for all but $\theta = \underline{\theta}, \overline{\theta}$. Since $T'(\theta) = 2(\theta - z)z'$, $T'(\theta) > 0$ for $\theta \in \left(\underline{\theta}, \overline{\theta}\right)$ under unified regime. That is, the tax is increasing in the type. Given (RC), this also implies that the low types receive subsidies and the high types pay taxes.

Under independent taxation, given that θ is uniformly distributed, (27) becomes:

$$z'' = -\frac{1}{\theta}\left[z' - 2 + \frac{2}{k}(\theta - z)z'\right] \tag{31}$$

$$z(\underline{\theta}) = \underline{\theta}, \; z(\overline{\theta}) = \overline{\theta}$$

Despite the lack of closed-form solutions, we are able to explore some analytical properties of the equilibrium based on this ODE system. Our first result is that under independent taxation, consumption is downward distorted for all but the top and bottom:

LEMMA 5 $\theta - z_I > 0$ for $\theta \in (\underline{\theta}, \overline{\theta})$.

---

[21] This is a similar condition to the one that Rochet and Stole (2002) impose to guarantee separating equilibrium in a nonlinear pricing setting with random participation. When this assumption fails, pooling occurs at the lower end.

Proof. Define $y(\theta) = \theta - z_I(\theta)$. Then $y(\underline{\theta}) = y(\overline{\theta}) = 0$, $y'(\theta) = 1 - z'_I(\theta)$, and $y''(\theta) = -\frac{1}{\theta}[1 + y' - \frac{2}{k}(1 - y')]$. It is equivalent to show that y never drops strictly below the zero line $(y = 0)$.

First, we show that the curve is initially shooting above, i.e., $y'(\underline{\theta}) > 0$. Suppose not, then there are two cases:

Case 1: $y'(\underline{\theta}) < 0$. Since $y(\underline{\theta}) = 0$, in this case we have $y(\underline{\theta}^+) < 0$. That is, the y curve is initially shooting below. Given the endpoint condition $y(\overline{\theta}) = 0$, at some point the curve has to shoot back to the zero line. So there is $\hat{\theta} \in (\underline{\theta}, \overline{\theta})$, such that $y'(\hat{\theta}) = 0$ and $y(\theta) < 0$ for all $\theta \in (\underline{\theta}, \hat{\theta}]$. In that case,

$$y''(\hat{\theta}) = -\frac{1}{\hat{\theta}}[1 - \frac{2}{k}y(\hat{\theta})] < 0.$$

This implies that $y(\hat{\theta}^+) < y(\hat{\theta}) < 0$, i.e., the curve keeps shooting below right after $\hat{\theta}$. However, given the endpoint condition, the curve has to come back at some later point. But our preceding argument suggests that the curve can never come back to the zero line, contradicting the endpoint condition.

Case 2: $y'(\underline{\theta}) = 0$. In this case,

$$y''(\underline{\theta}) = -\frac{1}{\underline{\theta}} < 0.$$

Thus $y(\underline{\theta}^+) < 0$. Now connecting our argument from here with the argument in the first case above, we establish contradiction again.

Thus we show that the curve is initially shooting above $(y'(\underline{\theta}) > 0)$. Given the endpoint condition, the curve will eventually drop back to the zero line. If it drops back to zero exactly at $\theta = \overline{\theta}$, we are done; otherwise, there is $\hat{\theta} \in (\underline{\theta}, \overline{\theta})$, such that $y'(\hat{\theta}) = 0$ and $y(\hat{\theta}) < 0$. Now following the same argument above, y can never get back to zero, contradiction. This establishes that $y(\theta) > 0$ except $\theta = \underline{\theta}, \overline{\theta}$. ∎

The result of efficiency at the top is standard in the screening literature. Efficiency at the bottom, which is implied from the transversality condition, however, is different from what we have seen from our base model with three types.[22]

The next lemma establishes that the equilibrium under independent taxation exhibits perfect sorting.

LEMMA 6 *Suppose condition (30) holds, then $z'_I(\theta) > 0$ and hence $T'_I(\theta) > 0$ for any $\theta \in [\underline{\theta}, \overline{\theta}]$.*

---

[22]A reconciliation is provided in the nonlinear pricing literature by Rochet and Stole (2002), who demonstrate that in a finite type model, the quality distortion for the lowest type disappears as the number of types goes to infinity.

Proof. See Appendix. ∎

The proof of Lemma 6 suggests that whenever the optimal solution under unified taxation exhibits perfect sorting, the equilibrium under independent taxation must exhibits perfect sorting. On the other hand, it is possible that pooling occurs under unified regime but the equilibrium under independent taxation exhibits perfect sorting.[23] The implication is that sorting occurs more easily under a competition regime. The intuition is similar to that provided in Yang and Ye (2008): higher types receive higher rents under competition, which relaxes the IC constraint, making it easier to sort the agents.

The next proposition displays interesting comparative statics with respect to the role of mobility:

PROPOSITION 4 *Let $k_2 < k_1$. Under independent taxation, (i) $\theta > z_2 > z_1$ for all $\theta \in (\underline{\theta}, \overline{\theta})$; (ii) $T_1(\underline{\theta}) > T_2(\underline{\theta})$ and $T_2(\overline{\theta}) < T_1(\overline{\theta})$; (iii) the tax schedule for (relatively) rich people is flatter under k_2.*

Proof. See Appendix. ∎

By continuity, we also have $T_1(\theta) > T_2(\theta)$ for types sufficiently close to $\underline{\theta}$, and $T_2(\theta) < T_1(\theta)$ for types sufficiently close to $\overline{\theta}$. As k goes down, the competition between two States becomes more intense. Proposition 4 suggests that as mobility (or competition) increases, the consumption distortion is reduced, the rich (types sufficiently close to the top) pay less taxes, and the poor (types sufficiently close to the bottom) receive less subsidies. While these results are obtained computationally in our three type model, they are obtained analytically in this continuous type model. Thus the result that *increased mobility leads to lower progressivity* is a fairly robust prediction. As in the three type model, as $k \to 0$, $T(\theta) = 0$. The solution under unified taxation, on the other hand, is independent of k, which can be regarded as the limiting case when $k \to +\infty$ (this can be seen from comparing (18) and (27)).

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 model.[24] The difference arises for the following reasons. In Simula and Trannoy, higher types have lower moving cost than lower cost types. This means that competition for top types is stronger than the competition for middle types, thus a

---

[23]Consider the following example. θ is uniformly distributed on $[1, 4]$, $k = 0.5$. Under unified taxation, the monotonicity constraint is violated and pooling occurs in the neighborhood of the low end. However, the equilibrium under independent taxation exhibits perfect sorting.

[24]Under independent taxation, $T' = 2(\theta - z)z'$ is always positive as $(\theta - z) \geq 0$ and $z' > 0$.

negative marginal tax rate might occur at the top. In our model, all (vertical) types have the same moving cost given the same horizontal type. We have thus demonstrated that the "curse" of middle types may not arise in a model with outside options endogenously determined.

We next turn to comparing the two taxation systems. This will be done by comparing the ODE systems (28) and (31). Using subscripts U and I to denote the unified and independent taxation regimes, respectively, we can state the following comparison results:

PROPOSITION 5 *(i) There is a $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $z_I'(\widehat{\theta}) = z_U'(\widehat{\theta})$, $z_I'(\theta) > z_U'(\theta)$ for $\theta \in [\underline{\theta}, \widehat{\theta})$ and $z_I'(\theta) < z_U'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta}]$; (ii) $z_I(\theta) > z_U(\theta)$ for any $\theta \in (\underline{\theta}, \overline{\theta})$; (iii) $T_I'(\theta) < T_U'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta})$.*

Proof. Part (i) is established in the proof of Lemma 6.

Part (ii) follows from (i) given the boundary conditions $z_I(\underline{\theta}) - z_U(\underline{\theta}) = z_I(\overline{\theta}) - z_U(\overline{\theta}) = 0$. For $\theta \in (\widehat{\theta}, \overline{\theta}]$, that $z_U < z_I$ and $z_U' > z_I'$ implies that $T_I'(\theta) < T_U'(\theta)$, as $T' = 2(\theta - z)z'$ under both taxation regimes. ∎

Therefore, under competition all types $\theta \in (\underline{\theta}, \overline{\theta})$ receive strictly higher consumption. Moreover, the tax schedule is flatter for the rich (those with sufficiently high types).

PROPOSITION 6 *(i) There is a $\widetilde{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $v_I(\widetilde{\theta}) = v_U(\widetilde{\theta})$, $v_I(\theta) < v_U(\theta)$ for $\theta \in [\underline{\theta}, \widetilde{\theta})$ and $v_I(\theta) > v_U(\theta)$ for $\theta \in (\widetilde{\theta}, \overline{\theta}]$; (ii) $T_I(\underline{\theta}) > T_U(\underline{\theta})$ and $T_I(\overline{\theta}) > T_U(\overline{\theta})$.*

Proof. From the first order conditions of the IC constraints, we have

$$v_I' - v_U' = \frac{1}{\theta}[2(z_I - z_U) - (v_I - v_U)]. \tag{32}$$

Over $(\underline{\theta}, \overline{\theta})$, given $z_I > z_U$, from (32) we have $v_I' > v_U'$ whenever $v_I = v_U$. This implies that over $(\underline{\theta}, \overline{\theta})$, v_I and v_U cross at most once, and at the intersection v_I must cross v_U from below.

Next we rule out the case that v_I and v_U never cross in the interior domain. Suppose $v_I(\underline{\theta}) \geq v_U(\underline{\theta})$. Then $v_I(\theta) \geq v_U(\theta)$ for all θ and $v_I(\theta) > v_U(\theta)$ for any $\theta > \underline{\theta}$. This contradicts the fact that $v_U(\theta)$ is the optimal solution under the unified regime, while $v_I(\theta)$ is one of the feasible schedules under the unified regime. Therefore, $v_I(\underline{\theta}) < v_U(\underline{\theta})$. Given that $z_I(\underline{\theta}) = z_U(\underline{\theta})$, it must be the case that $T_I(\underline{\theta}) > T_U(\underline{\theta})$.

Next we rule out the case that $v_I(\overline{\theta}) \leq v_U(\overline{\theta})$. Suppose this is the case. Then $v_I(\theta) < v_U(\theta)$ for all $\theta < \overline{\theta}$. At $\underline{\theta}$, $v_I(\underline{\theta}) < v_U(\underline{\theta})$, which implies that $T_I(\underline{\theta}) > T_U(\underline{\theta})$. At $\overline{\theta}$, $v_I(\overline{\theta}) \leq v_U(\overline{\theta})$, which implies $T_I(\overline{\theta}) \geq T_U(\overline{\theta})$. For any interior $\theta \in (\underline{\theta}, \overline{\theta})$,

$$v_I(\theta) - v_U(\theta) = \left[\left(2z_I(\theta) - \frac{z_I^2(\theta)}{\theta}\right) - \left(2z_U(\theta) - \frac{z_U^2(\theta)}{\theta}\right)\right] + \frac{T_U(\theta) - T_I(\theta)}{\theta}.$$

31

The first term in the bracket is positive since $\theta > z_I(\theta) > z_U(\theta)$. If $v_I(\theta) < v_U(\theta)$, we must have $T_U(\theta) < T_I(\theta)$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. Therefore, $\int_{\underline{\theta}}^{\overline{\theta}} T_I(\theta)d\theta > \int_{\underline{\theta}}^{\overline{\theta}} T_U(\theta)d\theta$, violating the resource constraint $\int_{\underline{\theta}}^{\overline{\theta}} T_I(\theta)d\theta = \int_{\underline{\theta}}^{\overline{\theta}} T_U(\theta)d\theta = 0$.

Thus, v_I crosses v_U (from below) exactly once at some interior $\theta \in (\underline{\theta}, \overline{\theta})$. This proves part (i). Part (ii) follows from part (i) and the boundary conditions. $\blacksquare$

So the rich (high-type citizens) are better off while the poor (low-type citizens) are worse off moving from unified to competitive taxation. The highest type (and the types sufficiently close to the highest type) pay less tax and the lowest type (and the types sufficiently close to the lowest type) get less subsidy under independent taxation.

To illustrate, we consider the example with $\underline{\theta} = 1$ and $\overline{\theta} = 2$. We can plot the tax schedules under both taxation regimes for any given value of k. The case with $k = 0.5$ is given in Figure 3 below. It is evident that for this case the tax schedule under independent regime is everywhere flatter, which strengthens our analytical result given in Proposition 6. Generally speaking, higher types are taxed less and lower types get less subsidy under the independent system.
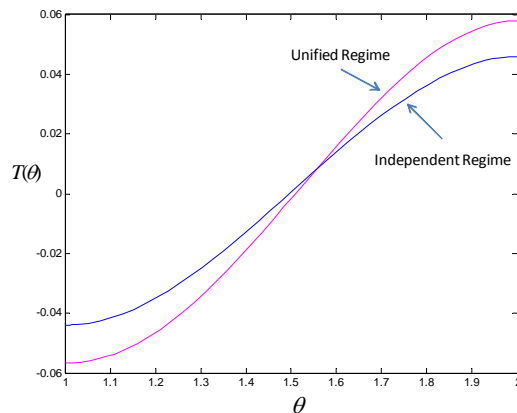

Figure 3: Tax Schedule Comparison with Uniform Distribution

With these results at hand, we are now ready to examine the determinants of constitutional choice with a continuum of types.

3.4 Constitutional Choice

Here the constitutional choice is determined by the median voter's preference. As in the three-type model, the preference of the median type can only be obtained using numerical computations. We thus go back to our model with general distributions for vertical types to characterize constitutional choice as a function of the mobility parameter, the distribution of relative classes (the types), and the distribution of income.

With any given distribution F (density function f), our computations can be done based on (18) and (27). Since the Pareto distribution is commonly adopted to proxy real world income inequality in the taxation literature, we consider the following truncated Pareto distribution family:

$$f(\theta) = \frac{\alpha\theta^{-\alpha-1}}{1-4^{-\alpha}} \text{ and } 1 - F(\theta) = \frac{\theta^{-\alpha} - 4^{-\alpha}}{1-4^{-\alpha}}, \theta \in [1,4].^{25} \tag{33}$$

Note that the uniform distribution is a special case of the Pareto distribution family (with $\alpha = -1$). As α increases, the density becomes more tilted toward lower types (more poor people). The tax schedules under two taxation systems are compared in Figure 4 below (plotted for the case $\alpha = 1$ and $k = 0.5$), which exhibits the same pattern as in the case of uniform distribution.
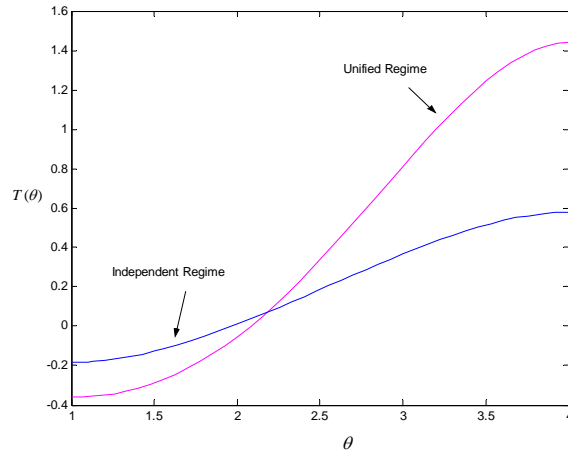

Figure 4: Tax Schedule Comparison with Pareto Distribution

Recall that with uniform distribution we established that the utility schedule v_I crosses v_U once from below. Our computation shows that this pattern of single crossing holds for truncated Pareto distributions as well. Let θ^* be the indifference type at which v_I crosses v_U. Then all the types below θ^* prefer the unified regime and all the types above θ^* prefer the independent regime. The following

---

[25] With the support of θ being $[1,4]$, the highest type's pre-tax income is 16 times that of the lowest type.

table shows how the indifference type θ^* shifts as k changes (for the truncated Pareto distribution, the computations are done based on the case $\alpha = -0.15$).

Table 1: How θ^* shifts as k changes

	$k = 1$	$k = 0.5$	$k = 0.3$	$k = 0.2$	$k = 0.1$	$k = 0.03$
Uniform $[1,3]$	1.8422	1.8529	1.8577	1.8635	1.8711	1.8815
Pareto $[1,4]$, $\alpha = -0.15$	2.0471	2.0626	2.0728	2.0798	2.0889	2.0965

The above table indicates that θ^* is monotonically decreasing in k. This is consistent with Observation 4 in the three type model. *Therefore, as the moving cost decreases, the measure of citizens who prefer the unified regime increases. As a result, the unified regime is more likely to be chosen at the constitutional stage for a smaller moving cost, other things equal.* The intuition for this result is analogous to that provided in the three type model. As k decreases, the previously indifferent type (the median type) "benefits" less from the presence of the rich (all the types above her), hence will switch her preferences toward the unified regime, whose solution does not depend on k.

For the range of mobility parameter k reported in the table, the unified regime is always chosen in the uniform distribution case (since the median type $\theta_m = 2$). However, for the truncated Pareto distribution case, the median type is $\theta_m = 2.0732$. Hence the independent regime will be chosen for cases $k = 0.3, 0.5$, and 1, and unified regime will be chosen for cases $k = 0.01, 0.1$, and 0.2.

We are also interested in how changes in the (type) income distribution affect the constitutional choice. Fix $k = 0.5$, and consider the truncated Pareto distributions given in (33). The following table reports how the indifference type θ^* and the median type θ_m change as α varies:

Table 2: How θ^* and θ_m shift as α changes

α	-0.5	-0.3	-0.2	-0.15	-0.1	0.5	1	1.5
θ^*	2.136	2.0933	2.0731	2.0626	2.0519	1.9437	1.8431	1.7645
θ_m	2.25	2.1484	2.0981	2.0732	2.0486	1.7778	1.60	1.4675

For all the cases we examined, the solutions exhibit perfect sorting. Two observations are worth noting. First, as α increases (more poor around), the *indifferent type* monotonically decreases. Again this is consistent with what we found from the three type model. This is intuitive: having more poor implies more taxes from the higher types in the unified regime, while in the independent regime the solution is closer to autarky. Therefore, the indifference type will decrease, as in Observation

3. However, if α is sufficiently large ($\alpha > -0.15$), the *median type* prefers the unified regime. Thus having more poor people in this continuous type case makes the choice of the unified system more likely, which seems to be inconsistent with our finding in the three type model. This happens in this Pareto distribution case simply because the *indifference type* decreases slower than the *median type*: as the size of the poor increases, the median type becomes even poorer. This observation highlights a difference between our three-type model and the continuous type model, that is, *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.*

Finally, we study how the degree of inequality affects constitutional choice by examining a distribution family with mean preserving spread. Again, we fix $k = 0.5$. Consider the following distribution family:

$$f_a(\theta) = \frac{1}{20 - \frac{2}{3}a}[10 - a(2 - \theta)^2], \ \theta \in [1, 3]$$

with $a \in [0, 10)$. The case $a = 0$ corresponds to the uniform distribution. As a increases, the distribution becomes more concentrated around the mean or median (which is 2 in this case), so inequality decreases. The computation results are reported in the following table. (θ^* is once again the cutoff type who is indifferent between the two tax regimes):

Table 3: How θ^* shifts as inequality parameter changes

	$a = 0$	$a = 3$	$a = 5$	$a = 7$	$a = 9$
θ^*	1.8813	1.8615	1.8561	1.8672	1.8728

The table shows that the relationship between inequality and the indifference type is not monotonic in this particular continuous type distribution case.

4 Concluding remarks

This paper has extended the analysis of optimal income taxation to the case in which strategic authorities compete for heterogeneous citizens, and where the heterogeneity is in productivity as well as mobility characteristics. Every agent's productivity and ability to move 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.

The independent taxation system yields lower progressivity than the unified case. Under competition the rich are better off and the poor are worse off, and whether the middle type is better off or worse off depends on mobility and on the distribution of income. In particular, in our base model with three types, we have shown that the middle type is more likely to choose the unified system when the mobility level is high (k is smaller), or when the proportion of the poor is not too large. Our analysis of the continuous type model confirms most of the main findings from our three type model, and provides some additional insights for this competitive nonlinear taxation framework.

An important extension of this model will be the consideration of asymmetric initial conditions. 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. 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 a static model cannot suffice to analyze the important relationship between inequality, redistribution, and institutions.

Appendix

Proof of Lemma 1: This can be shown in the following 4 steps:

1. The monotonicity of Q, and hence C.

   Adding DIC_{HM} to UIC_{MH}, we have $(Q_H - Q_M)(\frac{1}{\theta_M} - \frac{1}{\theta_H}) \geq 0$. This implies that $Q_H \geq Q_M$. By similar arguments, we can show that $Q_H \geq Q_M \geq Q_L$. By DIC_{HM}, $Q_H \geq Q_M$ implies that $C_H \geq C_M$. By similar arguments, $Q_H \geq Q_M \geq Q_L$ implies that $C_H \geq C_M \geq C_L$.

2. DIC_{HL} and UIC_{LH} are inactive.

   Adding DIC_{HM} to DIC_{ML}, we have

   $$u(C_H) - u(C_L) \geq \frac{Q_H - Q_M}{\theta_H} + \frac{Q_M - Q_L}{\theta_M} \geq \frac{Q_H - Q_L}{\theta_H}.$$

   The last inequality follows from $Q_M \geq Q_L$. The above inequality implies DIC_{HL}. By a similar argument, one can show that UIC_{LH} is inactive. Now we have 4 IC constraints left, plus the monotonicity of Q and C.

3. Under unified taxation, DIC_{HM} and DIC_{ML} bind, and all the other constraints are inactive.

We first show that DIC_{HM} must bind. Suppose not. Then we can increase Q_H by ε, and decrease Q_M and Q_L by the same amount $\frac{\mu_H}{1-\mu_H}\varepsilon$, where ε is strictly positive and sufficiently small. Note that this change does not affect the resource constraint. With sufficiently small ε, DIC_{HM} still holds. The other three constraints hold as well. But this change leads to the following change in the objective function:

$$\mu_H \varepsilon \left[-\frac{1}{\theta_H} + \frac{1}{1-\mu_H} \left(\frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L} \right) \right] > 0.$$

Therefore, DIC_{HM} must bind. By similar arguments, if DIC_{ML} does not bind, then we can construct the following change: increase Q_H and Q_M by the same amount ε, and decrease Q_L by $\frac{1-\mu_L}{\mu_L}\varepsilon$. The RC and 4 IC constraints are still satisfied, but it leads to an increase in the objective function. Therefore, DIC_{ML} must bind.

4. Under unified taxation, the two UIC's are inactive.

   A binding DIC_{HM} and the monotonicity of Q jointly imply

   $$u(C_H) - u(C_M) = \frac{Q_H - Q_M}{\theta_H} < \frac{Q_H - Q_M}{\theta_M},$$

   which in turn implies UIC_{MH}. By a similar argument, one can show that UIC_{LM} is inactive.

   Therefore, under unified taxation IC holds if and only if the monotonicity constraint holds and the two DIC's (3) bind. ∎

Proof of Lemma 3: We illustrate this point by considering the following case. Suppose in equilibrium UIC binds for both M and L types (the proofs for the cases that only one UIC binds are similar). In this case, one can show that C_L is not distorted, but both C_M and C_H are distorted upwards ($C_L < C_M \leq C_H$). The binding UIC's imply that

$$
\begin{aligned}
T_H - T_M &= \theta_M \left\{ [u(C_H) - \frac{C_H}{\theta_M}] - [u(C_M) - \frac{C_M}{\theta_M}] \right\} \leq 0; \\
T_M - T_L &= \theta_L \left\{ [u(C_M) - \frac{C_M}{\theta_L}] - [u(C_L) - \frac{C_L}{\theta_L}] \right\} < 0.
\end{aligned}
$$

The terms in the brackets are negative since $C_H \geq C_M > C_L = C_L^*$. Given that $T_H \leq T_M < T_L$, by the resource constraint we have $T_H < 0$ and $T_L > 0$. Now we construct a profitable deviation for one State. Suppose State 1 decreases T_M by ε, decreases T_L by ε and increases T_H by $\frac{\mu_M}{\mu_H}\varepsilon$ ($\varepsilon > 0$ but small). Note that under the new tax schedule, UIC-L still binds but UIC-M is slack. The change of budget for State 1 is:

$$\mu_M \left(\frac{\varepsilon}{k} \frac{T_M}{\theta_M} - \varepsilon \right) + \mu_L \left(\frac{\varepsilon}{k} \frac{T_L}{\theta_L} - \varepsilon \right) + \mu_H \left(-\frac{\varepsilon}{k} \frac{T_H}{\theta_H} \frac{\mu_M + \mu_L}{\mu_H} + \frac{\mu_M + \mu_L}{\mu_H} \varepsilon \right)$$

$$= \frac{\varepsilon}{k} \left[\mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_L} + \mu_H \frac{T_H}{\theta_H} - \frac{T_H}{\theta_H} \right] > \frac{\varepsilon}{k} \left[\mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_L} + \mu_H \frac{T_H}{\theta_H} \right]$$

$$> \frac{\varepsilon}{k} \left[\mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_M} + \mu_H \frac{T_H}{\theta_M} \right] = 0.$$

The inequalities are based on $T_H < 0$ and $T_L > 0$. Therefore, the new tax schedule is feasible for State 1. Now we compute the change in the value of the objective function:

$$-\mu_H \frac{\mu_M + \mu_L}{\theta_H \mu_H} \varepsilon + \mu_M \frac{\varepsilon}{\theta_M} + \mu_L \frac{\varepsilon}{\theta_L}$$

$$= \varepsilon \left[\mu_M \left(\frac{1}{\theta_M} - \frac{1}{\theta_H} \right) + \mu_L \left(\frac{1}{\theta_L} - \frac{1}{\theta_H} \right) \right] > 0.$$

Therefore, it constitutes a profitable deviation for State 1. ∎

Proof of Proposition 1: We start with the case that both DIC's bind under independent taxation. First we show that $u'(C_L^I) < u'(C_L^U)$.

$$u'(C_L^I) = \frac{\mu_L \lambda_R}{\mu_L - \lambda_M + \mu_L \lambda_R T_L / k}$$

$$= \frac{1}{\frac{1}{\lambda_R} \frac{\theta_L - \theta_M}{\theta_L} + \theta_M + T_L \frac{1}{k} \frac{\theta_L - \theta_M}{\theta_L}}$$

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

$$u'(C_L^U) - u'(C_L^I) = \frac{1}{\frac{\theta_L - \theta_M}{\lambda_R^U \theta_L} + \theta_M} - \frac{1}{\frac{1}{\lambda_R} \frac{\theta_L - \theta_M}{\theta_L} + \theta_M + T_L \frac{1}{k} \frac{\theta_L - \theta_M}{\theta_L}}$$

$$\sim \left[\frac{1}{\lambda_R} \frac{\theta_L - \theta_M}{\theta_L} + \theta_M + T_L \frac{1}{k} \frac{\theta_L - \theta_M}{\theta_L} \right] - \left[\frac{\theta_L - \theta_M}{\lambda_R^U \theta_L} + \theta_M \right]$$

$$= \frac{\theta_M - \theta_L}{\theta_L} \left[\frac{1}{\lambda_R^U} - \frac{1}{\lambda_R} - \frac{T_L}{k} \right]$$

$$\sim \frac{1}{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}} - \frac{1 - \left(T_H \frac{\mu_H}{k\theta_H} + T_M \frac{\mu_M}{k\theta_M} + T_L \frac{\mu_L}{k\theta_L} \right)}{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}} - \frac{T_L}{k}$$

$$\sim \left(T_H \frac{\mu_H}{k\theta_H} + T_M \frac{\mu_M}{k\theta_M} + T_L \frac{\mu_L}{k\theta_L} \right) - \frac{T_L}{k} \left(\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L} \right)$$

$$\sim (T_H - T_L) \frac{\mu_H}{\theta_H} + (T_M - T_L) \frac{\mu_M}{\theta_M}$$

Given that $T_H > T_M \geq T_L$, we have $u'(C_L^U) - u'(C_L^I) > 0$. Next we show that $u'(C_M^I) < u'(C_M^U)$.

$$u'(C_M^U) = \frac{\mu_M \lambda_R^U}{\mu_M - \lambda_H + \lambda_M} = \frac{1}{\frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R^U} - \frac{\mu_H}{\mu_M}\theta_H - \frac{\mu_L}{\mu_M}\theta_M}$$

$$u'(C_M^I) = \frac{\mu_M \lambda_R}{\mu_M - \lambda_H + \lambda_M + \mu_M \lambda_R \frac{T_M}{k}}$$

$$= \frac{1}{\frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R} - \frac{\mu_H}{\mu_M}\theta_H - \frac{\mu_L}{\mu_M}\theta_M + \frac{\mu_H}{\mu_M}\frac{T_H}{k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{k} + \frac{T_M}{k}}$$

$$u'(C_M^U) - u'(C_M^I) \sim \frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R} - \frac{\mu_H}{\mu_M}\theta_H - \frac{\mu_L}{\mu_M}\theta_M + \frac{\mu_H}{\mu_M}\frac{T_H}{k} + \frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{k} + \frac{T_M}{k}$$

$$-\frac{1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}}{\lambda_R^U} + \frac{\mu_H}{\mu_M}\theta_H + \frac{\mu_L}{\mu_M}\theta_M$$

$$\sim \left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)\lambda_R^U - \left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)\lambda_R$$

$$+\left(\frac{\mu_H}{\mu_M}\frac{T_H}{k}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{k}+\frac{T_M}{k}\right)\lambda_R^U \cdot \lambda_R$$

$$\sim \left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)\times\left[1-\left(T_H\frac{\mu_H}{k\theta_H}+T_M\frac{\mu_M}{k\theta_M}+T_L\frac{\mu_L}{k\theta_L}\right)\right]$$

$$-\left(1+\frac{\mu_H}{\mu_M}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\right)+\left(\frac{\mu_H}{\mu_M}\frac{T_H}{k}+\frac{\mu_L}{\mu_M}\frac{\theta_M}{\theta_L}\frac{T_L}{k}+\frac{T_M}{k}\right)\left(\frac{\mu_H}{\theta_H}+\frac{\mu_M}{\theta_M}+\frac{\mu_L}{\theta_L}\right)$$

$$= \mu_H\left(\frac{1}{\theta_M}-\frac{1}{\theta_H}\right)\left(\frac{T_H}{k}-\frac{T_M}{k}\right)+\frac{\mu_H}{\mu_M}\frac{\mu_L}{\theta_L}\left(1-\frac{\theta_M}{\theta_H}\right)\left(\frac{T_H}{k}-\frac{T_L}{k}\right) > 0$$

Now consider the case that neither DIC binds under independent taxation. Given that $C_M^I = C_M^*$ and $C_L^I = C_L^*$, we clearly have $u'(C_L^U) > u'(C_L^I)$ and $u'(C_M^U) > u'(C_M^I)$.

Finally, consider the case that DIC-H is slack but DIC-M binds under independent taxation. Given that $C_M^I = C_M^*$, we clearly have $u'(C_M^U) > u'(C_M^I)$. The proof for $u'(C_L^U) > u'(C_L^I)$ is exactly the same as that in the first case above, as the expressions for λ_M, λ_R and $u'(C_L^I)$ are exactly the same under both cases. ∎

Proof of Lemma 6: First, whenever $z_I' = 0$, $z_I'' = \frac{2}{\theta} > 0$. By the single-crossing lemma, z_I' has the single crossing property. That is, z_I' crosses zero line from below at most once.[26]

What remains to be shown is that $z_I'(\underline{\theta}) > 0$. Now compare two differential equation systems (28) and (31). Whenever $z_I' = z_U'$ (> 0), we have $z_I'' < z_U''$ (since $\theta - z_I > 0$ by Lemma 5). By the single-crossing lemma, the curve $z_I'(\theta) - z_U'(\theta)$ crosses zero line from above at most once. Given

---

[26]Therefore, if there is pooling, it must happen at the low end.

the boundary conditions $z_I(\underline{\theta}) - z_U(\underline{\theta}) = z_I(\overline{\theta}) - z_U(\overline{\theta}) = 0$, we conclude that $z_I'(\theta) - z_U'(\theta)$ has to cross zero line exactly once. That is, there is a $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $z_U'(\theta) < z_I'(\theta)$ for $\theta \in [\underline{\theta}, \widehat{\theta})$, and $z_U'(\theta) > z_I'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta}]$. Given that $z_U'(\underline{\theta}) > 0$, we have $z_I'(\underline{\theta}) > z_U'(\underline{\theta}) > 0$. This completes the proof for $z_I' > 0$.

Given $z_I' > 0$ and $(\theta - z_I) > 0$, we have $T_I'(\theta) = 2(\theta - z_I)z_I' > 0$ for $\theta \in (\underline{\theta}, \overline{\theta})$. ∎

Proof of Proposition 4: (i) The two differential equations under independent taxation are as follows:

$$
\begin{aligned}
z_1'' &= \frac{1}{\theta}[2 - z_1' - \frac{2}{k_1}(\theta - z_1)z_1'], \\
z_2'' &= \frac{1}{\theta}[2 - z_2' - \frac{2}{k_2}(\theta - z_2)z_2'].
\end{aligned}
\tag{34}
$$

Let $y = z_2 - z_1$. We have $y(\underline{\theta}) = y(\overline{\theta}) = 0$. We need to show that $y(\theta) > 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. The proof idea resembles that of Lemma 5.

First we show that $y'(\underline{\theta}) > 0$. Suppose in negation, $y'(\underline{\theta}) \leq 0$.

Case 1: $y'(\underline{\theta}) < 0$. Given that $y(\overline{\theta}) = 0$, there exists $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $y'(\widehat{\theta}) = 0$ and $y(\theta) < 0$ for all $\theta \in (\underline{\theta}, \widehat{\theta}]$. But then it is easily verified that $y''(\widehat{\theta}) < 0$. This implies that y will always remain strictly below zero after initially shooting below, a contradiction.

Case 2: $y'(\underline{\theta}) = 0$. It is easily verified that in this case all higher derivatives at $\underline{\theta}$ are zero: $y^{(n)}(\underline{\theta}) = 0$ for all $n \geq 2$. This, combined with $y(\underline{\theta}) = 0$, implies that there exists $\widehat{\theta}$ sufficiently close to $\underline{\theta}$, such that $y(\widehat{\theta}) = y'(\widehat{\theta}) = y''(\widehat{\theta}) = 0$. However, with notation $z(\widehat{\theta}) = z_1(\widehat{\theta}) = z_2(\widehat{\theta})$ and $z'(\widehat{\theta}) = z_1'(\widehat{\theta}) = z_2'(\widehat{\theta})$, we can demonstrate that

$$
y''(\widehat{\theta}) = \frac{1}{\widehat{\theta}}\left[2(\widehat{\theta} - z(\widehat{\theta}))z'(\widehat{\theta})\left(\frac{1}{k_1} - \frac{1}{k_2}\right)\right].
$$

Since $z'(\widehat{\theta}) > 0$ and $\widehat{\theta} - z(\widehat{\theta}) > 0$, the above expression implies that $y''(\widehat{\theta}) < 0$, a contradiction.

So the y curve is initially shooting up. Given the endpoint condition, it will eventually come back to the zero line. If it comes back exactly at $\overline{\theta}$, we are done with the proof; otherwise it drops below zero before reaching the end point $\overline{\theta}$. But then there is $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $y'(\widehat{\theta}) = 0$ and $y(\theta) < 0$ for all $\theta \in (\underline{\theta}, \widehat{\theta}]$. Applying the same argument to rule out Case 1 above, we can establish the contradiction. So y has to stay above zero except two boundary points.

(ii) Similarly to the previous proof, that $\theta > z_2 > z_1$ implies that v_2 cross v_1 at most once from below. Again, the case that $v_1 > v_2$ for all θ can be ruled out. But so far the case $v_1 < v_2$ for all θ cannot be ruled out. Therefore, we can only show $T_2(\overline{\theta}) < T_1(\overline{\theta})$.

(iii) Note that we have $z_1(\theta) < z_2(\theta)$ for any interior θ. This implies that at the neighborhood of $\overline{\theta}$, $z_1' > z_2'$. As a result, in this neighborhood, $T_1' > T_2'$ as well. ∎

References

[1] Armstrong, Mark and John Vickers (2001): "Competitive Price Discrimination," *Rand Journal of Economics* 32(4): 579-605.

[2] Bhagwati, J., Hamada, K. (1982): "Tax Policy in the Presence of Emigration," *Journal of Public Economics* 18: 291–317.

[3] Brett, Craig and John Weymark (2008): "Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility," Working Paper.

[4] Chen, Yan (2000): "Electoral Systems, Legislative Process and Income Taxation," *Journal of Public Economic Theory* 2(1): 71-100.

[5] Coughlan, Peter (1986): "Elections and Income Redistribution," *Public Choice* 50 (1-2): 27-91.

[6] Coughlan, Peter (1992): *Probabilistic Voting Theory*, Cambridge, Cambridge University Press.

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

[8] Epstein, Larry and Michael Peters (1999): "A Revelation Principle for Competing Mechanisms," *Journal of Economic Theory* 88: 119-60.

[9] Hamilton, J., and P. Pestieau (2005): "Optimal Income Taxation and the Ability Distribution : Implications for Migration Equilibria," *International Tax and Public Finance* 12: 29–45.

[10] Hindriks, J. (1999): "The Consequences of Labour Mobility for Redistribution: Tax vs. Transfer Competition," *Journal of Public Economics* 74: 215–234.

[11] Huber, Bernd (1999): "Tax Competition and Tax Coordination in an Optimal Income Tax Model," *Journal of Public Economics* 71: 441–458.

[12] Krause, A. (2007): "Education and Taxation Policies in the Presence of Countervailing Incentives," *Economica*, forthcoming.

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

[14] Leite-Monteiro, M. (1997): "Redistributive Policy with Labour Mobility Accross Countries," *Journal of Public Economics* 65: 229–244.

[15] Lindbeck, Assar and Joerg Weibull (1987): "Balanced-budget Redistribution as the Outcome of Political Competition," *Public Choice*, 273-97.

[16] Martimort, David and Lars Stole (2002): "The Revelation and Delegation Principles in Common Agency Games," *Econometrica* 70: 1659–1673.

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

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

[19] Mirrlees, James (1982): "Migration and Optimal Income Taxes," *Journal of Public Economics* 18: 319–341.

[20] 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.

[21] Osmundsen, P. (1999): "Taxing Internationally Mobile Individuals–A Case of Countervailing Incentives," *International Tax and Public Finance* 6: 149–164.

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

[23] Page Jr., H. Frank and Paulo K. Monteiro (2003): "Three Principles of Competitive Nonlinear Pricing," *Journal of Mathematical Economics* 39: 63-109.

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

[25] Peters, Michael (2001): "Common Agency and the Revelation Principle," *Econometrica* 69(5): 1349-72.

[26] Piaser, Gwenael (2007): "Labor Mobility and Income Tax Competition," In: Gregoriou, G., Read, C. (Eds.), International Taxation Handbook: Policy, Practice, Standards and Regulation. CIMA Publishing, Oxford, pp. 73–94.

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

[28] 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.

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

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

[31] Tiebout, Charles (1956): "A Pure Theory of Local Government Expenditure," *Journal of Political Economy* 64: 416-24

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

[33] Wilson, John Douglas (1980): "The Effect of Potential Emigration on the Optimal Linear Income Tax," *Journal of Public Economics* 14, 339–353.

[34] Wilson, John Douglas (1992): "Optimal Income Taxation and International Personal Mobility," *American Economic Review*, Papers and Proceedings 82, 191–196.

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

[36] Wilson, John Douglas (2006): "Income taxation and skilled migration: The analytical issues", unpublished manuscript, Department of Economics, Michigan State University.

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