

# Competitive Nonlinear Taxation and Constitutional Choice\*

Massimo Morelli,<sup>†</sup> Huanxing Yang,<sup>‡</sup> and Lixin Ye<sup>§</sup>

November 2010

## Abstract

In an economy where agents have different productivities and mobility, we compare a unified nonlinear optimal taxation with the equilibrium taxation 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; the “rich” are always in favor of competing authorities, whereas the “poor” are always in favor of unified taxation; the preferences of the middle class depend on the initial conditions in terms of the distribution of abilities, the relative power of the various classes, and mobility costs.

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

## 1 Introduction

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

---

\*We thank Daron Acemoglu, Stefania Albanesi, Francesco Caselli, Steve Coate, Oriol Carbonell-Nicolau, Dennis Epple, Bernard Salanie, John Weymark, and participants in the political economy and theory workshops at Princeton University, London School of Economics, UC Santa Barbara, University of Hong Kong, Peking University, Ohio State University, and at various meetings for very helpful comments and suggestions on the early versions of our paper. The usual disclaimer applies.

<sup>†</sup>Columbia University and European University Institute. mm3331@columbia.edu.

<sup>‡</sup>The Ohio State University. huanxing@econ.ohio-state.edu.

<sup>§</sup>The Ohio State University. lixinye@econ.ohio-state.edu.

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 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 thinking 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.<sup>1</sup> On the other hand, we are used to thinking of the institutional 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 the view that this issues can be separated, 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.

In order to compare the *effects* and the *origins* of centralized versus decentralized taxation systems, we consider a framework 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 a weighted average of all the citizens’ utilities. Under the independent taxation system, each State’s objective is to choose a tax schedule to maximize the weighted 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.

In our main 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).<sup>2</sup> Under the independent authority regime, a taxation authority has to take into account not only the resource constraints and incentive

---

<sup>1</sup>Taxation is called progressive when the average tax rate (not necessarily the marginal rate) increases with income.

<sup>2</sup>For robustness check, we have also worked out a continuous type model. The full analysis is reported in Appendix B.

compatibility constraints of a standard optimal taxation designer, but also the additional individual rationality constraint derived from location preferences. In this independent taxation regime the tax for the high type is lower and the subsidy for the low type is lower accordingly. Moreover, we show that under the independent regime the total output and consumption are higher, but the total 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 also means that the subsidy to the low type decreases accordingly. Finally, we show that under independent taxation *bunching* is less likely to occur, implying that under independent taxation there might be more tax brackets. The main reason for this result is that the incentive compatibility constraints are relaxed under independent taxation due to competition for high types.

At the constitutional stage, the collective choice between independent and centralized taxation would be trivial if we assumed a Rawlsian “veil of ignorance”: behind a veil of ignorance, in fact, there would be unanimous support for a centralized taxation system. In reality we observe plenty of decentralization choices, and this may be an indication that in most realistic constitutional stages in history the collective decision is made by agents that “represent” various classes, “in front” of the veil of ignorance. There is variation of course about the procedures and voting rules with which such delegates decide on the institution that will govern future government decision making, but here we assume simple majority.<sup>3</sup> The representatives of the interests of low productivity types (the poor) should always be in favor of a unified taxation regime. On the other hand, the representatives of the high productivity types (the rich) should prefer the independent regime. Hence the constitutional choice between the two regimes can always be thought of as determined by the preferences of the middle class (excluding the trivial cases in which one of the two extreme types has the absolute majority at the constitutional stage). Even though a unified regime always yields higher welfare, we can show that a country with better initial conditions (higher average productivity) may end up with lower welfare because the majority decision can favor decentralization 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

---

<sup>3</sup>See Barbera and Jackson (2004) for an interesting perspective on the consistent selection of majority rules.

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.

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, is the constitutional choice problem that best fits our framework, and where our analysis offers some important interpretations. 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, Wilson (2006), Krause (2007), and Simula and Trannoy (2010) study how allowing agent migration affects the optimal non-linear income tax schedule of a State, taking the other States’ tax schedules as exogenous outside options.<sup>4</sup> Hamilton and Pestieau (2005) provide a general equilibrium analysis of tax competition

---

<sup>4</sup>In particular, Simula and Trannoy (2010) 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

among a large number of small countries. They consider two skilled types and that only one type can move.

To the best of our knowledge, Piaser (2007), Brett and Weymark (2008), and Gordon and Cullen (2010) 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. The analysis with three types involves problems that do not arise with two types, as will be clarified below.

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  $k \rightarrow 0$ .

Gordon and Cullen (2010) also model a Nash game in (nonlinear) tax schedules between states, but a key difference is that in their analysis, the states and federal government both engage in income redistribution.<sup>5</sup> Since the federal government is also active when states set their tax policies, there exist “vertical tax externalities” in the sense that higher taxes at the state level tend to reduce tax revenue collected at the federal level. As a result, it is possible for state governments to redistribute too much income in equilibrium. Given that our objective in this research is to compare unified and competitive taxation with varying mobility costs, we ignore this kind of vertical externalities by assuming that redistribution occurs at only one level of government.

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.<sup>6</sup> The most related paper to ours in the literature on capital 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.

<sup>5</sup> A similar approach is taken by Boadway et al. (1998), who consider linear income taxation only.

<sup>6</sup> 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). See also Huber (1999).

competition is perhaps the recent one by Hatfield and Padro i Miquel (2008), because they too study the preferences of different citizens for the different levels of decentralization of taxes. They model both the constitutional stage and the tax implementation stage as a median voter's choice, whereas in our view the constitutional choice is the only one that makes sense to relate to voters' preferences directly. The choice of a tax schedule in a given system is instead an outcome of political competition, which leads, under standard probabilistic voting assumptions,<sup>7</sup> to a policy outcome that is equivalent to the solution of an average utility maximization problem.

The connection between mobility and redistribution of income was 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.

In terms of modeling and technical issues, our paper is most closely related to Armstrong and Vickers (2001) and Rochet and Stole (2002), who study competitive nonlinear pricing with both vertically and horizontally differentiated agents.<sup>8</sup> Our analysis differs from theirs in two main features at the technical level: first, we need to take into account the additional resource constraint; second, given our focus on the preference of the middle class, we need to solve a three-type model for the unified and decentralized system, which calls for additional care in dealing with the incentive compatibility constraints.

The paper is organized as follows. We lay out the model with three ability types in Section 2. In Sections 3 we analyze the unified and independent taxation and then compare two taxation regimes. In Section 4 we discuss constitutional choice over two taxation regimes. Section 5 provides concluding remarks with some discussion of the restrictions of our analysis. Appendix A contains most of the proofs and Appendix B offers an extended analysis of the case with continuous ability types.

---

<sup>7</sup>See e.g. Lindbeck and Weibull (1987).

<sup>8</sup>Also see Yang and Ye (2008) for a similar framework allowing for partial market coverage along vertical dimension.

## 2 The Model

Citizens (or workers/consumers) are characterized by identical preferences and different abilities (i.e., marginal productivities). Given consumption (or after-tax income)  $C$  and labor supply  $l$ , the preferences can be represented by the following quasi-linear utility function:

$$U(C, l) = u(C) - l \tag{1}$$

where  $u(\cdot)$  is strictly increasing, strictly concave and twice continuously differentiable.<sup>9</sup> Let  $Q$  denote the total product or before-tax income, then  $C = Q - T(Q)$ , where  $T(\cdot)$  is the tax schedule set by the tax authority. A citizen's ability is denoted by  $\theta$ , which captures the (constant) marginal productivity.

We assume that the labor market is competitive, and the wages are bid up to the marginal productivities of workers, which implies that  $Q = \theta l$ . The utility function (1) can be rewritten as follows:

$$U(C, Q; \theta) = u(C) - Q/\theta \tag{2}$$

We consider three ability types: type  $H$  (the “rich”), type  $M$  (the middle type), and type  $L$  (the “poor”), with abilities  $\theta_H$ ,  $\theta_M$  and  $\theta_L$ , respectively ( $\theta_H > \theta_M > \theta_L$ ). In Appendix B, we extend our analysis to the case with a continuum of ability types.

We 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.<sup>10</sup> Each State  $i$ ,  $i = 1, 2$ , has a total measure (population) of 1 original citizens attached to it. The State that a citizen is initially attached to is called her home State. Citizens can move from their home state to the other state. The cost of moving is given by  $(1 - x)k$ , where  $x$  denotes a locational preference which is individual specific,  $x \in [0, 1]$ , and  $k$  is a common factor affecting the moving cost for all

---

<sup>9</sup>We assume that preferences are quasi-linear in labor. There is a tradition of using such preferences, see, for example, Lollivier and Rochet (1983), Rochet (1987), and Boadway et al. (2000). Some more recent work has tended to opt for preferences that are quasi-linear in consumption (e.g., Diamond, 1998, Saez, 2001, and Salanie, 2003). We have tried both utility specifications. For the discrete type model, the qualitative results are the same. But for the continuous type model, with quasi-linearity in consumption the differential equation system characterizing the equilibrium under independent taxation becomes too complicated, which makes it hard to compare with the solution under unified taxation. For tractability we thus follow the more traditional approach, assuming that the preferences are quasi-linear in labor.

<sup>10</sup>Our 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.

the citizens. More specifically,  $x$  measures the degree of flexibility of a citizen: the smaller is  $x$ , the larger is the moving cost, or the greater the attachment to the home State.<sup>11</sup> On the other hand, the smaller  $k$ , 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 locational 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.

We assume that  $k$  is a (strictly) positive constant that is commonly known, but neither the ability  $\theta$  nor the locational preference parameter  $x$  is observable to the tax authority. Thus a citizen is characterized by a two-dimensional private type  $(\theta, x)$ . Using the jargon in the industrial organization literature,  $\theta$  can be regarded as the “vertical” type, while  $x$  can be regarded as the “horizontal” type in a Hotelling-type model (so that a citizen with a smaller  $x$  can be regarded as being located closer to the base of her home State).

We denote the corresponding proportions of the three types by  $\mu_H$ ,  $\mu_M$  and  $\mu_L$ , respectively, and we assume that  $x$  is uniformly distributed on the interval  $[0, 1]$ .<sup>12</sup>

Each State  $i$  decides on a tax schedule  $T_i(Q)$ . Given  $(T_1(\cdot), T_2(\cdot))$ , workers choose their State of residence and then  $Q$ , to maximize  $u(Q - T(Q)) - Q/\theta$ . The single crossing property holds only along the vertical dimension, with the consequence that the tax authorities can only design tax schedules to sort agents along that dimension.

In the environment of competitive mechanism design, it is not without loss of generality to restrict attention to direct contracts (see e.g. Martimort and Stole, 1997 and Peck, 1997). To sidestep this problem, we restrict attention to deterministic tax schedules, allowing to consider direct contracts of the form  $\{T_1(\theta), T_2(\theta)\}_{\theta \in \{\theta_H, \theta_M, \theta_L\}}$ .<sup>13</sup> For brevity of exposition, from now on we will often refer to vertical types as simply the types, especially when there is no confusion in the context.

Instead of choosing tax schedules of the form  $T_i(\theta)$ , each State may equivalently be assumed to choose direct contracts of the form  $(Q_i(\theta), C_i(\theta))$ ,  $i = 1, 2$ .<sup>14</sup> Since everyone is required to participate

---

<sup>11</sup>The citizen with  $x = 0$  is the least mobile, while the citizen with  $x = 1$  is the most mobile.

<sup>12</sup>Assuming some other distributions would not alter our main results, as we focus on symmetric equilibria in which no citizens move. However, doing so would necessarily complicate our equilibrium analysis.

<sup>13</sup>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).

<sup>14</sup>The (direct) tax schedule can be recovered from  $T_i(\theta) = Q_i(\theta) - C_i(\theta)$ .

in one of the tax systems, the individual rationality constraint only concerns which State to live in. Let  $v_i \equiv u(C_i) - Q_i/\theta_i$  be the rent provision to type- $\theta_i$  citizen who accepts contract  $(Q_i, C_i)$  from her home State  $i$ ,  $i \in \{H, M, L\}$ . Suppose the other State's taxation rule leads to rent provisions  $v_i^*$  for this type- $\theta_i$  citizen. Then this citizen with a horizontal type  $x$  will stay with her home State if and only if

$$v_i \geq v_i^* - k(1 - x), \quad i \in \{H, M, L\} \quad (3)$$

It is well known that if the menu of contracts,  $(Q_i(\cdot), C_i(\cdot))$ ,  $i = 1, 2$ , satisfy the self-selection constraint (3), then it can be supported by an anonymous income tax schedule (see, for example, Guesnerie and Seade, 1982). For this reason we can think of each tax authority as choosing a tax schedule  $T_i(Q)$  or, equivalently, as directly choosing a menu of contracts of the form  $(Q_i(\theta), C_i(\theta))$ ,  $i = 1, 2$ , for its residents, subject to the self-selection constraints. In what follows, we assume that the tax authority directly chooses the menu of contracts  $(Q(\theta), C(\theta))$ ,  $\theta \in \{\theta_H, \theta_M, \theta_L\}$ .<sup>15</sup>

Formally, under the independent taxation regime, the time line is as follows. In period  $t = 1$ , each State chooses its menu of contract of the form  $\{C_i(\theta), Q_i(\theta)\}$ ,  $i = 1, 2$ , simultaneously and independently. In period  $t = 2$ , given  $(\{C_i(\theta), Q_i(\theta)\})$ ,  $i = 1, 2$ , citizens decide on the location and the labor supply (or equivalently, the contract  $(C, Q)$  to accept). In period  $t = 3$ , production (or pre-tax income) is realized and consumption is determined (or equivalently, taxes are collected) according to the menu of contracts pre-announced at  $t = 1$ .

The tax authorities are benevolent. They share the same social preferences over the utility space, represented by the welfare function  $W(U_H, U_M, U_L)$ , where  $U_i = u(C_i) - Q_i/\theta_i$ , the utility per capita of type  $\theta_i$  who accepts a contract  $(C_i, Q_i)$ ,  $i = H, M, L$ . The tax authority maximizes a weighted utilitarian social welfare function, with the weights being the proportions of the three ability types.<sup>16</sup> Thus each tax authority's objective is to maximize the following welfare function:

$$W(U_H, U_M, U_L) = \mu_H U_H + \mu_M U_M + \mu_L U_L. \quad (4)$$

Alternatively, moving costs could be included in the objective function, so that the welfare depends on the horizontal "market shares" (i.e., the horizontal measure of citizens for each ability type). We do not follow that approach due to the following considerations. First, should we include the

---

<sup>15</sup>See e.g. Weymark (1986a, b) and Brett and Weymark (2008) for related papers that follow the same approach.

<sup>16</sup>The weighted utilitarian social welfare can be regarded as a linear approximation of a general quasiconcave social welfare function at the initial utility levels (Weymark, 1987). We choose the weights to be the proportions of the three ability types because the treatment of each class should intuitively reflect its relative size.

horizontal market shares in the welfare function, then by simply attracting more high type workers, the total welfare will increase even if the utility per capita for each type remains unchanged, which is an undesirable feature. Second, the weights used in a social choice function are usually exogenously given. If we include the endogenously determined horizontal market shares in the weights, our analysis can easily become intractable.<sup>17</sup> Also note that although horizontal market shares do not enter the objective functions, competing for higher type workers along the horizontal dimension is still important, as redistribution is the only purpose of taxation in our model, and hence the tax authority always has an incentive to attract more “rich” to subsidize the “poor” to improve the (weighted) welfare.

As the citizens’ utilities and the resulting market shares for the two States are functions of the menu of contracts offered (or the tax schedules), we can focus on the analysis of stage 1 alone. This can be done by replacing stages 2 and 3 with the corresponding payoffs as functions of the offered menu of contracts. Our solution concept in this reduced one-shot game is Bertrand-Nash equilibrium, which is characterized by the pair of menus of contracts  $\{(C_1(\theta), Q_1(\theta)), (C_2(\theta), Q_2(\theta))\}_{\theta \in \{\theta_H, \theta_M, \theta_L\}}$ : given  $(C_{-i}(\theta), Q_{-i}(\theta))$ ,  $(C_i(\theta), Q_i(\theta))$  maximizes the welfare function (4) among the workers who choose to reside in its State subject to the usual incentive compatibility and resource constraints,  $i = 1, 2$ .<sup>18</sup>

---

<sup>17</sup>In that case we may even need to worry about the existence of a perfect sorting equilibrium. Suppose a perfect sorting equilibrium exists. Then a State may want to deviate by choosing  $T(\theta_L) = Q(\theta_L)$  trying to get rid of all the poor: by maximizing taxes on the poor like this, the State is basically forcing all the poor to move to the other State, and hence increasing the total (or average) utility (including the moving costs) of its residents. But this cannot be sustainable, as the other State can react the same way, leaving the existence of a perfect sorting equilibrium in question. However, when individuals are heterogeneous only with respect to their attachment to home states (the vertical ability types are all the same), the non-existence problem mentioned above does not arise and it is feasible to include moving costs in the objective functions, see, for example, Mansoorian and Myers (1993, 1997).

<sup>18</sup>Note that the solution concept of Nash equilibrium in this context can be controversial, as when State 1 changes its menu of contracts (tax schedule) State 2’s resource constraint may be violated due to the induced migration flows to or from State 2. So the tax schedule of State 2 may not be sustainable. This issue is similar to the well known feature of Rothschild-Stiglitz type Nash equilibrium, where the constraint of having nonnegative profit for one firm can be easily violated following an off-equilibrium deviation by the other firm (Rothschild and Stiglitz, 1976). Note that imposing the restriction that State 2’s resource constraint be satisfied after State 1 changes its menu of contracts will be inconsistent with the notion of Nash equilibrium, as such a restriction requires that State 1 correctly anticipate a resource-balancing response by State 2. Using Nash equilibrium as our solution concept, we implicitly require that the resource constraint be satisfied only on equilibrium path. The solution concept of Nash equilibrium in competitive taxation setting is also adopted in Piarser (2007), Brett and Weymark (2008), and Gordon and Cullen (2010).

This completes the description of the model with independent taxation. For the model 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, whose objective is to maximize the welfare function (4) for *all* the citizens living in the Federation.

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

$$u'(C^*) = 1/\theta. \quad (5)$$

The optimal consumption or before-tax 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  $\theta$ .

### 3 Unified Taxation versus Independent Taxation

We will start with an analysis of the unified taxation, followed by an analysis of the independent taxation. We will then compare these two taxation regimes and discuss the constitutional choice by augmenting our above model with a constitutional stage ( $t = 0$ ), where the taxation regime is decided by simple majority rule.

#### 3.1 Unified Taxation

Under unified taxation, we solve for the tax schedule that maximizes the weighted utility of the citizens in the Federation. 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.<sup>19</sup>

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

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

---

<sup>19</sup>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 the symmetric solution is optimal for the Federation.

and the incentive compatibility (IC) constraints, which basically require that no type has incentive to mimic any of the other types. With three types there will be 6 inequality conditions: for any  $i = L, M, H$  and  $j \neq i$

$$u(C_i) - \frac{Q_i}{\theta_i} \geq u(C_j) - \frac{Q_j}{\theta_j} \quad (6)$$

The above problem is essentially a standard optimal taxation problem with discrete types. In such problems, it is well known that the IC constraints can be simplified as a monotonicity constraint  $Q_H \geq Q_M \geq Q_L$  plus (two) binding local DIC's. The proof of the following lemma can be found in Weymark (1986a,b, 1987).

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

$$\begin{aligned} u(C_H) - \frac{Q_H}{\theta_H} &= u(C_M) - \frac{Q_M}{\theta_H}, \quad (\text{DIC-H}) \\ u(C_M) - \frac{Q_M}{\theta_M} &= u(C_L) - \frac{Q_L}{\theta_M} \quad (\text{DIC-M}) \end{aligned} \quad (7)$$

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  $\lambda_H$ ,  $\lambda_M$  and  $\lambda_R$  respectively. The first order conditions can be written as follows:

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

From the above equations, we can obtain

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

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  $u(\cdot)$ ). Thus 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,  $T_H^U > T_M^U \geq T_L^U$ .*

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

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

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

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

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

Given  $T_H > T_M \geq T_L$ ,<sup>20</sup> 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

---

<sup>20</sup>Similar results regarding the ordering of tax payments can be found in Brito et al. (1990) and Matthews and Moore (1997).

$T_M$  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.

So far we have ignored the monotonicity constraint. When the monotonicity constraint fails, bunching may occur.<sup>21</sup> Given that type  $M$  and type  $H$  never pool in the optimal solution, we only need to check whether  $C_L^U \leq C_M^U$ , or equivalently, whether  $u'(C_L^U) \geq u'(C_M^U)$ . From the previous analysis, we have

$$\begin{aligned} u'(C_L^U) &\geq u'(C_M^U) \Leftrightarrow \mu_L \lambda_H^U \leq (\mu_L + \mu_M) \lambda_M^U \Leftrightarrow \mu_H (\theta_H \lambda_R^U - 1) \leq (\mu_L + \mu_M) \left( \frac{\theta_M}{\theta_L} - \theta_M \lambda_R^U \right) \\ &\Leftrightarrow \mu_H \left( \frac{\theta_H}{\theta_L} - 1 \right) [(\mu_L + \mu_M) \frac{\theta_M}{\theta_H} - \mu_L] + \mu_M [(\mu_L + \mu_M) \left( \frac{\theta_M}{\theta_L} - 1 \right) - \mu_H \left( \frac{\theta_H}{\theta_M} - 1 \right)] \geq 0 \end{aligned} \quad (8)$$

Therefore, if (8) is satisfied, then there is no bunching in the optimal solution. Observing (8), we can see that bunching will not occur if the high and middle types' abilities are not too different ( $\frac{\theta_M}{\theta_H}$  is not too low) and the proportion of high type,  $\mu_H$ , is not too large. If (8) is violated, then bunching can occur. In the bunching case, denote  $C_{ML}^U$  and  $Q_{ML}^U$  to be the consumption and output, respectively, for both types  $M$  and  $L$ . The programming problem now becomes:

$$\begin{aligned} \max \quad & \mu_H \left[ u(C_H) - \frac{Q_H}{\theta_H} \right] + \mu_M \left[ u(C_{ML}) - \frac{Q_{ML}}{\theta_M} \right] + \mu_L \left[ u(C_{ML}) - \frac{Q_{ML}}{\theta_L} \right] \\ \text{s.t.} \quad & \mu_H (Q_H - C_H) + (\mu_M + \mu_L) (Q_{ML} - C_{ML}) = 0 \\ & u(C_H) - \frac{Q_H}{\theta_H} = u(C_{ML}) - \frac{Q_{ML}}{\theta_H} \end{aligned}$$

The solution to the above problem is characterized by

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

With bunching, the resource constraint can be written as  $\mu_H T_H + (\mu_M + \mu_L) T_{ML} = 0$ . By slightly modifying the proof of Lemma 2, we can show that  $T_H > 0 > T_{ML}$ .

Note that the output,  $Q_i$ , should be nonnegative. With the explicit requirement that outputs are nonnegative, in a continuous type model Seade (1977) shows that it might be optimal for the lowest types to produce zero outputs. Since incorporating the nonnegativity constraint of output would complicate the analysis without adding much additional insight, it is typically ignored in the optimal taxation literature (e.g., Lollivier and Rochet, 1983, and Guesnerie and Laffont, 1984). In our model, by the binding resource constraint and the two DIC's,  $Q_L$  can be expressed as:

$$Q_L = \mu_H \{ C_H - \theta_M [u(C_M) - u(C_L)] - \theta_H [u(C_H) - u(C_M)] \} + \mu_M \{ C_M - \theta_M [u(C_M) - u(C_L)] \} + \mu_L C_L$$

<sup>21</sup>Weymark (1986b) provides a detailed analysis of the possibility of bunching in a one-country model.

If the ability of the low type,  $\theta_l$ , is high enough relative to those of higher types, or if the proportion of the low type,  $\mu_l$ , is relatively large, then the nonnegativity constraint of output will not bind.

### 3.2 Independent Taxation

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

We follow the same notation defined earlier to let  $v_i \equiv u(C_i) - Q_i/\theta_i$  be the rent provision to type- $\theta_i$  citizen who accepts contract  $(Q_i, C_i)$  from her home State  $i$ ,  $i \in \{H, M, L\}$ , and  $v_i^*$  be the corresponding rent provision to this type- $\theta_i$  citizen from the other State. Then by the self selection condition (3), a citizen with vertical type  $\theta_i$  and horizontal type  $x$  will stay with her home State if and only if

$$x \leq \min \left\{ 1 + \frac{v_i - v_i^*}{k}, 1 \right\}.$$

When  $v_i \geq v_i^*$ , all the type- $\theta_i$  citizens in the State in question will stay with their home State, and all the types  $(\theta_i, x)$  where  $x \geq 1 - (v_i - v_i^*)/k$  in the other State will move to the State in question. Therefore for vertical type  $\theta_j$ , the total measure of horizontal types that will reside in the State in question will be  $1 + (v_i - v_i^*)/k$ .<sup>22</sup> For this reason,  $x_i$  defined below can be regarded as the “market share” of type  $\theta_i$ ,  $i \in \{H, M, L\}$ , for the State in question:

$$x_i = 1 + \frac{v_i - v_i^*}{k}. \quad (9)$$

The State's objective 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  $x_i$ 's are given by (9).

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 six IC's (6) can be reduced to four local IC's ( $DIC_{HM}$ ,  $DIC_{ML}$ ,  $UIC_{MH}$ , and  $UIC_{LM}$ ) plus the monotonicity constraint  $Q_H \geq Q_M \geq Q_L$ .

---

<sup>22</sup>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.<sup>23</sup> 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 A. ■

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

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

where  $x_i$ 's are given by (9).

Unlike in the unified taxation case, under independent taxation the DIC's may not bind simultaneously.<sup>24</sup> 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 consider multiple cases.

1. Both DIC's bind. Let  $\lambda_H$  and  $\lambda_M$  be the multipliers of DIC-H and DIC-M, respectively, and let  $\lambda_R$  be the multiplier of RC. We first derive the first order conditions, then impose symmetry.

---

<sup>23</sup>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.

<sup>24</sup>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.

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

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

From the above equations, we obtain

$$\begin{aligned}
u'(C_H) &= 1/\theta_H, \\
\lambda_H^I &= \mu_H \left[ -1 + \theta_H \lambda_R^I \left( 1 - \frac{T_H}{k\theta_H} \right) \right] \\
\lambda_M^I &= \mu_L \left[ \frac{\theta_M}{\theta_L} - \theta_M \lambda_R^I \left( 1 - \frac{T_L}{k\theta_L} \right) \right] \\
\lambda_R^I &= \frac{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}}{1 - [T_H \frac{\mu_H}{k\theta_H} + T_M \frac{\mu_M}{k\theta_M} + T_L \frac{\mu_L}{k\theta_L}]} \\
u'(C_M) &= \frac{\mu_M \lambda_R^I}{\mu_M - \lambda_H^I + \lambda_M^I + \mu_M \lambda_R^I T_M/k} \\
u'(C_L) &= \frac{\mu_L \lambda_R^I}{\mu_L - \lambda_M^I + \mu_L \lambda_R^I T_L/k}
\end{aligned} \tag{10}$$

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.

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

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

Clearly, consumption is no longer distorted:  $C_j^I = C_j^*$ ,  $j = L, M, H$ . Moreover,  $T_H > T_M > T_L$ .<sup>25</sup> Now the DIC's can be rewritten as:

$$\begin{aligned} u(C_H^*) - \frac{C_H^*}{\theta_H} - \frac{T_H}{\theta_H} &\geq u(C_M^*) - \frac{C_M^*}{\theta_H} - \frac{T_M}{\theta_H}, \\ u(C_M^*) - \frac{C_M^*}{\theta_M} - \frac{T_M}{\theta_M} &\geq u(C_L^*) - \frac{C_L^*}{\theta_M} - \frac{T_L}{\theta_H}. \end{aligned}$$

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 small. So, the DICs will not bind when  $k$  is sufficiently small. In the limit as  $k \rightarrow 0$ ,  $T_H$ ,  $T_M$  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.

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

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

That is, there is no consumption distortion for types  $H$  and  $M$ , but the consumption of type  $L$  is distorted downward.<sup>26</sup> The expressions for  $\lambda_M$ ,  $\lambda_R$  and  $u'(C_L^I)$  are the same as those in (10). In this case, we have  $T_H > T_M > T_L$ .

Note that bunching can only occur when both DICs bind. When at least one DIC does not bind, from previous analysis we have  $u'(C_L^I) > u'(C_M^I)$ , which implies that the monotonicity constraint  $C_M > C_L$  is satisfied. If bunching between type  $M$  and  $L$  does occur when both DICs bind, the programming problem becomes:

$$\begin{aligned} &\max \mu_H v_H + \mu_M v_M + \mu_L v_L \\ &\text{s.t. } u(C_H) - \frac{Q_H}{\theta_H} = u(C_{ML}) - \frac{Q_{ML}}{\theta_H} \\ &\quad \mu_H x_H (Q_H - C_H) + (\mu_M x_M + \mu_L x_L) (Q_{ML} - C_{ML}) = 0 \end{aligned}$$

---

<sup>25</sup>So  $C_j$ 's are the same as in the autarky case, though  $Q_j$ 's are different.

<sup>26</sup>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.

Following the proof of Lemma 2, we can show that  $T_H^I > 0 > T_{ML}^I$ . Moreover, the solution to the above problem becomes:

$$\begin{aligned}
u'(C_H) &= 1/\theta_H, \\
\lambda_H^I &= \mu_H \left[ -1 + \theta_H \lambda_R \left( 1 - \frac{T_H}{k\theta_H} \right) \right] \\
\lambda_R^I &= \frac{\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}}{1 - \left[ T_H \frac{\mu_H}{k\theta_H} + T_{ML} \frac{\mu_M}{k\theta_M} + T_{ML} \frac{\mu_L}{k\theta_L} \right]} \\
u'(C_{ML}^I) &= \frac{(\mu_M + \mu_L)\lambda_R^I}{\mu_L + \mu_M - \lambda_H^I + (\mu_M + \mu_L)\lambda_R^I T_{ML}/k}
\end{aligned}$$

To summarize, we have the following lemma regarding the monotonicity of taxes.

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

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.

### 3.3 Comparison

Our first comparison result shows that bunching is more likely to occur under unified taxation.

PROPOSITION 1 *If bunching does not occur under unified taxation, then bunching does not occur under independent taxation.*

**Proof.** Given that under independent taxation bunching only occurs when both DICs bind, we only need to show that if bunching does not occur under unified taxation, bunching will not occur under independent taxation when both DICs bind. Under unified taxation, the condition that ensures no bunching is given by (8):

$$\mu_L \lambda_H^U < (\mu_L + \mu_M) \lambda_M^U. \quad (11)$$

When both DIC's bind under independent taxation, the no-bunching condition  $u'(C_L^I) > u'(C_M^I)$  can be written as:

$$\mu_L \lambda_H^I < (\mu_L + \mu_M) \lambda_M^I + \lambda_R^I \mu_M \mu_L \left( \frac{T_M^I}{k} - \frac{T_L^I}{k} \right). \quad (12)$$

Given that  $T_M^I \geq T_L^I$ , to show that (11) implies (12), it is sufficient to show that  $\lambda_H^I \leq \lambda_H^U$  and  $\lambda_M^I \geq \lambda_M^U$ .

We first prove that  $\lambda_R^I < \lambda_R^U$ . By the previously derived expressions, it is sufficient to show that  $T_H \frac{\mu_H}{k\theta_H} + T_M \frac{\mu_M}{k\theta_M} + T_L \frac{\mu_L}{k\theta_L} < 0$ . But

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

where the first inequality is due to  $T_L < 0$ , the first equality is due to the binding resource constraint, and the last inequality is due to  $T_H > 0$ . Therefore,  $\lambda_R^I < \lambda_R^U$ . Now given that  $\lambda_R^I < \lambda_R^U$  and  $T_H > 0$ , we can immediately see that  $\lambda_H^I \leq \lambda_H^U$  by comparing the expressions. Similarly,  $\lambda_M^I \geq \lambda_M^U$  follows the fact that  $\lambda_R^I < \lambda_R^U$  and  $T_L < 0$ . ■

It is easy to construct examples such that bunching occurs under unified taxation but does not occur under independent taxation. For instance, we can pick the parameter values such that the no-bunching condition under unified taxation (8) is violated, but  $k$  is small enough such that the DICs do not bind. Proposition 1 implies that bunching is more likely to occur under unified taxation. An empirical implication of this result is that there might be more tax brackets under the independent taxation regime.

Our second 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 2**  $C_L^I > C_L^U$  and  $C_M^I > C_M^U$ : *competition increases consumption for both types  $M$  and  $L$ .*

**Proof.** See Appendix A. ■

The proof is rather tedious, and is therefore relegated to the appendix. Although the consumption for the rich stays the same (undistorted), the induced output or income  $Q$  is different under different taxation regimes. Competition always increases consumption for both types  $M$  and  $L$ , regardless of whether bunching occurs or not, and regardless of whether it occurs under unified taxation alone or under both taxation regimes. Next we turn to the welfare comparisons. The rich pay less taxes under competitive taxation, and should always prefer the independent taxation.

**PROPOSITION 3**  $T_H^I < T_H^U$  and  $v_H^I > v_H^U$ . *That is, type  $H$  pays lower taxes and is better off under independent taxation.*

**Proof.** See Appendix A. ■

On the contrary, the poor receives lower subsidy under competitive taxation, and should always prefer a unified taxation system.

PROPOSITION 4  $T_L^I > T_L^U$  and  $v_L^I < v_L^U$ . That is, type  $L$  receives less subsidies and is worse off under independent taxation.

The results derived in the above propositions are quite intuitive. Under independent taxation, each state tries to attract more high types to increase its tax revenue so that it can redistribute more to low types. This competition for high types leads to lower taxes for high types and lower subsidies to low types.<sup>27</sup> Thus, high types are better off and low types are worse off under independent taxation. Since higher types will get higher utilities due to the competition (for higher types) under independent taxation, the IC constraints are more relaxed under independent taxation. As a result, the downward consumption distortion for the lower types, which are created to satisfy the IC constraints, are reduced with independent taxation.<sup>28</sup> Moreover, more relaxed IC constraints means that the monotonicity constraint will be more likely to hold, which reduces the possibility of bunching.

Since the tax schedules under independent taxation  $\{(Q_i^I, C_i^I)\}_{i \in \{H, M, L\}}$  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  $\theta$ 's) and income distribution (distribution of  $\mu$ 's). However, the computations we now turn to, provide interesting results.<sup>29</sup>

---

<sup>27</sup>Another way to understand this is as follows. The tax base of each state depends both on individuals' labor supply choice and the number of individuals in the State. Under unified taxation, the tax base is elastic only along the labor supply margin, while under independent taxation it is elastic along both margins. Therefore, the tax base of each state is more elastic under independent taxation; and as a result, less tax can be collected and less redistribution can be made. We thank one referee for suggesting this interpretation.

<sup>28</sup>The result that independent taxation reduces the consumption of the low types but increases their before-tax income even more has something to do with the quasilinear-in-leisure preferences. Weymark (1987) provides a discussion of a similar phenomenon in a comparative statics analysis regarding welfare weights in a one-country model.

<sup>29</sup>Detailed computations and Matlab code used in this project are available upon request.

## 4 Constitutional Choice

We now augment our model with a constitutional stage (stage 0), 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. 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

RESULT 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  $\theta_M^*$ .*

Our computation reveals that the utility differences between two taxation regimes for types  $H$  and  $L$  are both monotonic in  $k$ . More specifically,  $v_H^I - v_H^U$  is positive and decreasing in  $k$ , and  $v_L^I - v_L^U$  is negative and increasing in  $k$ : as  $k$  increases, the difference between two regimes diminishes, and the utility difference between two regimes also become smaller. For the middle type, the schedule of  $v_M^I - v_M^U$  is in general non-monotonic in  $k$  (though it is monotonically increasing over  $(.5, 1.5]$ ). But the whole schedule is higher when  $\theta_M$  is higher, which also gives rise to Result 1 above. Our computations also show that given  $\theta_M \in (1, 2)$  and  $\mu_H = \mu_L = (1 - \mu_M)/2$ :<sup>30</sup>

RESULT 2 *There exists a cutoff  $\mu_M^*$  such that type  $M$  prefers the independent taxation regime if and only if  $\mu_M > \mu_M^*$ .*

Intuitively, as  $\theta_M$  or  $\mu_M$  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  $\mu_M^*$  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  $\theta$  is *higher* in the second case, welfare is higher in the former Federation, for  $\epsilon$  sufficiently small (by Corollary 1 above). This means that

COROLLARY 2 *A country with “better” initial conditions (higher productivity, or higher average  $\theta$  here) may end up with lower welfare because of a suboptimal constitutional choice due to majority decision making at the constitutional stage.*

---

<sup>30</sup>When  $\mu_M$  increases, we let  $\mu_H$  and  $\mu_L$  go down by the same compensating amount.

Another interesting observation comes from the following exercise: fix  $\theta_M$  and  $\mu_M$  (or  $\mu_L$ ); then our computations show that

RESULT 3 *There exists  $\mu_L^*$  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:

RESULT 4 *Both  $\theta_M^*$  and  $\mu_M^*$  are decreasing in  $k$ .*

This suggests that when  $k$  decreases, for a given  $\theta_M$  or  $\mu_M$ , 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  $\theta_M = 1.3$ .<sup>31</sup>

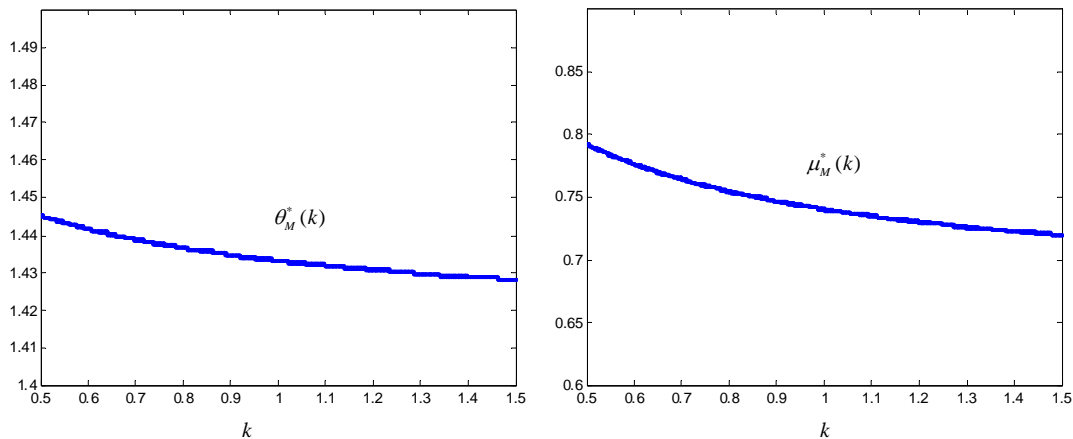


Figure 1: Schedules of  $\theta_M^*(k)$  and  $\mu_M^*(k)$

An intuition for Result 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

---

<sup>31</sup>Our computations show that the higher the selected value for  $\theta_M$ , the lower the schedule of  $\mu_M^*(k)$ . This is consistent with our Results 1 and 2.

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  $\mu_M$  on horizontal axis and  $\theta_M^*$  on vertical axis, our computations show that:

RESULT 5  $\theta_M^*$  decreases as  $\mu_M$  increases (while the other two types decrease symmetrically at the same time).

Figure 2 below is plotted with  $k = 1$ . Increasing  $\mu_M$  in this way reduces inequality but also reduces total productivity when  $\theta_M < 1.5$ . If  $\theta_M$  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.<sup>32</sup>

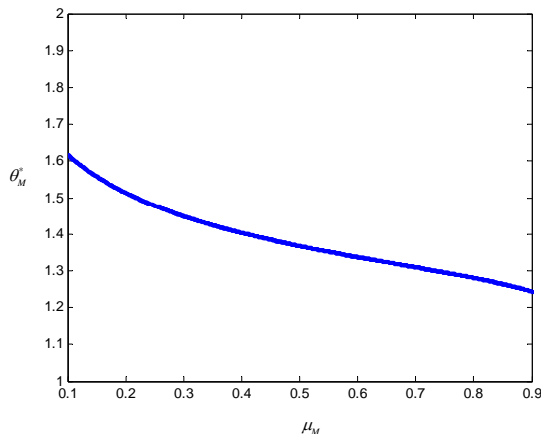


Figure 2: Schedule of  $\theta_M^*(\mu_M)$

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

---

<sup>32</sup>The pattern between  $\theta_M^*$  and  $\mu_M$  is a fortiori decreasing when the increase in  $\mu_M$  is balanced by a reduction in  $\mu_H$  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  $\theta_M^*$ .

confounding effects. We are able to say something clearer about the role of initial inequality when studying the case of a continuum of ability types (Appendix B).

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.<sup>33</sup>

## 5 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, 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 (reported in Appendix B) confirms most of the main findings from our three type model, and

---

<sup>33</sup>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 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.

provides some additional insights for this competitive nonlinear taxation framework.

An important direction for future research is the consideration of asymmetric initial conditions. Tracing the impact of different initial conditions on constitutional choice would also permit 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.

In our analysis, we assume that vertical types are independent of horizontal types. It would be reasonable to assume that higher (richer) types have lower moving costs, as in Simula and Trannoy (2010). However, we do not expect that allowing this correlation would change our results qualitatively. In that case, under independent taxation the competition for the high type would be more fierce. As a result, the high types would still be better off and the lower types would still be worse off.

Another restriction in our analysis is that we did not consider exogenous revenue requirements (or public good provision), thus redistribution is the only purpose of taxation, unlike in, for example, Boadway et al. (2000) and Stähler (2002). Nevertheless, we do not expect that incorporating an exogenous revenue requirement would change our main results qualitatively. The only change we expect is that under both taxation regimes the amount of redistribution will become smaller, as the aggregate tax must be positive in that case.

Finally, one feature of our current analysis is that the constitutional choice is made by the median voter, while the States are weighted utilitarian once the constitution has been chosen. If one considered the (fully normative) alternative in which at the constitutional stage institutions are chosen in a welfare-maximizing manner, then clearly the centralized taxation regime would always be chosen (regardless of type distributions or mobility costs). On the other hand, if one considered the opposite (fully positive) alternative in which the taxation policy is chosen based on direct voting, then there would be a technical question regarding how to offer alternatives of tax schedules to voters to choose from. This technical issue, together with the coexistence of centralized and decentralized taxation regimes in the real world, suggest that neither of the above alternative assumptions, albeit consistent, can be completely satisfactory. Even though the assumptions we have made for the two stages may appear somewhat inconsistent, this current research represents a first attempt to bridge constitutional choice and taxation design in a way that aims to shed light on when we should

expect to see one system or the other. In a sense we have provided a benchmark where citizens compare institutions under the most benevolent assumptions about their functioning. In future work, more realistic political economy models of the different regimes could replace our optimal taxation framework, and their equilibrium outcomes (and consequent constitutional choice incentives) will be usefully contrasted with the benchmark we established here.

## Appendix A

**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

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

The first inequality above follows from the upward distortion of both  $C_M$  and  $C_H$ , which implies that  $u(C) - \frac{C}{\theta_M}$  is a decreasing function of  $C$  in the region under consideration. The second inequality follows as  $C_H \geq C_M > C_L = C_L^*$ . 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  $\varepsilon$ , decreases  $T_L$  by  $\varepsilon$  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:

$$\begin{aligned} & \mu_M \left( \frac{\varepsilon T_M}{k \theta_M} - \varepsilon \right) + \mu_L \left( \frac{\varepsilon T_L}{k \theta_L} - \varepsilon \right) + \mu_H \left( -\frac{\varepsilon T_H}{k \theta_H} \frac{\mu_M + \mu_L}{\mu_H} + \frac{\mu_M + \mu_L}{\mu_H} \varepsilon \right) \\ &= \frac{\varepsilon}{k} \left[ \mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_L} + \mu_H \frac{T_H}{\theta_H} - \frac{T_H}{\theta_H} \right] > \frac{\varepsilon}{k} \left[ \mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_L} + \mu_H \frac{T_H}{\theta_H} \right] \\ &> \frac{\varepsilon}{k} \left[ \mu_M \frac{T_M}{\theta_M} + \mu_L \frac{T_L}{\theta_M} + \mu_H \frac{T_H}{\theta_M} \right] = 0. \end{aligned}$$

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:

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

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

**Proof of Proposition 2:** Given the possibility of bunching and the possibility that DIC's might not bind under independent taxation, we have several cases to consider.

1. No bunching under unified taxation. Since no bunching under unified taxation implies no bunching under independent taxation, we have the following subcases to consider.

(a) Both DIC's bind under independent taxation. First we show that  $u'(C_L^I) < u'(C_L^U)$ .

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

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

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

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

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

$$\begin{aligned}
u'(C_M^U) - u'(C_M^I) &\sim \frac{1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}}{\lambda_R} - \frac{\mu_H \theta_H}{\mu_M} - \frac{\mu_L \theta_M}{\mu_M} + \frac{\mu_H T_H}{\mu_M k} + \frac{\mu_L \theta_M T_L}{\mu_M \theta_L k} + \frac{T_M}{k} \\
&\quad - \frac{1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}}{\lambda_R^U} + \frac{\mu_H \theta_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M} \\
&\sim \left(1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}\right) \lambda_R^U - \left(1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}\right) \lambda_R \\
&\quad + \left(\frac{\mu_H T_H}{\mu_M k} + \frac{\mu_L \theta_M T_L}{\mu_M \theta_L k} + \frac{T_M}{k}\right) \lambda_R^U \cdot \lambda_R \\
&\sim \left(1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}\right) \times \left[1 - \left(T_H \frac{\mu_H}{k \theta_H} + T_M \frac{\mu_M}{k \theta_M} + T_L \frac{\mu_L}{k \theta_L}\right)\right] \\
&\quad - \left(1 + \frac{\mu_H}{\mu_M} + \frac{\mu_L \theta_M}{\mu_M \theta_L}\right) + \left(\frac{\mu_H T_H}{\mu_M k} + \frac{\mu_L \theta_M T_L}{\mu_M \theta_L k} + \frac{T_M}{k}\right) \left(\frac{\mu_H}{\theta_H} + \frac{\mu_M}{\theta_M} + \frac{\mu_L}{\theta_L}\right) \\
&= \mu_H \left(\frac{1}{\theta_M} - \frac{1}{\theta_H}\right) \left(\frac{T_H}{k} - \frac{T_M}{k}\right) + \frac{\mu_H \mu_L}{\mu_M \theta_L} \left(1 - \frac{\theta_M}{\theta_H}\right) \left(\frac{T_H}{k} - \frac{T_L}{k}\right) > 0
\end{aligned}$$

(b) 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  $u'(C_M^U) > u'(C_M^I)$ .

(c) 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 subcase 1(a), as the expressions for  $\lambda_M$ ,  $\lambda_R$  and  $u'(C_L^I)$  are exactly the same under both cases.

## 2. There is bunching under unified taxation.

(a) No bunching and both DIC's bind under independent taxation. It is sufficient to show that  $u'(C_{ML}^U) > u'(C_L^I)$ . Since in Subcase 1(a) we have shown that  $u'(C_L^U) > u'(C_L^I)$ , it is sufficient to show that  $u'(C_{ML}^U) \geq u'(C_L^U)$ . The condition  $u'(C_{ML}^U) \geq u'(C_L^U)$  can be written more explicitly as:

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

But given that the no bunching condition (8) is violated, the above condition is satisfied.

(b) Both DIC's bind and there is bunching under independent taxation. Our goal is to show  $u'(C_{ML}^U) > u'(C_{ML}^I)$ . Recall that

$$u'(C_{ML}^I) = \frac{\mu_M + \mu_L}{\frac{1}{\lambda_R^I} - \mu_H \theta_H + \mu_H \frac{T_H}{k} + (\mu_M + \mu_L) \frac{T_{ML}}{k}} = \frac{\mu_M + \mu_L}{\frac{1}{\lambda_R^I} - \mu_H \theta_H}$$

Where the second inequality uses the binding resource constraint:  $\mu_H T_H + (\mu_M + \mu_L) T_{ML} = 0$ . Now we have

$$\begin{aligned}
u'(C_{ML}^U) - u'(C_{ML}^I) &\sim \frac{1}{\frac{1}{\lambda_R^U} - \mu_H \theta_H} - \frac{1}{\frac{1}{\lambda_R^I} - \mu_H \theta_H} \sim \frac{1}{\lambda_R^I} - \frac{1}{\lambda_R^U} \\
&\sim - \left( T_H \frac{\mu_H}{\theta_H} + T_{ML} \frac{\mu_M}{\theta_M} + T_{ML} \frac{\mu_L}{\theta_L} \right) > - \left( T_H \frac{\mu_H}{\theta_H} + T_{ML} \frac{\mu_M}{\theta_L} + T_{ML} \frac{\mu_L}{\theta_L} \right) \\
&= \mu_H T_H \left( \frac{1}{\theta_L} - \frac{1}{\theta_H} \right) > 0
\end{aligned}$$

Note that we used the fact that  $T_{ML} < 0$  and the binding resource constraint along the way.

- (c) At least one DIC does not bind under independent taxation. Recall that there is no bunching under independent taxation. The proof follows closely those of previous subcases. ■

**Proof of Proposition 3:** Given the possibility of bunching and the possibility that DIC's may not bind under independent taxation, we have several cases to consider.

1. No bunching under unified taxation. Since no bunching under unified taxation implies no bunching under independent taxation, we have the following subcases to consider.

- (a) Both DIC's bind under independent taxation. By RC, DIC-H, and DIC-M, we have

$$\begin{aligned}
Q_H &= \mu_H C_H + \mu_M C_M + (1 - \mu_H) \theta_H [u(C_H) - u(C_M)] \\
&\quad + \mu_L [C_L + \theta_M (u(C_M) - u(C_L))]
\end{aligned} \tag{13}$$

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 2. Then from (13), we have

$$\Delta Q_H = \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H) \theta_H - \mu_L \theta_M] \Delta u(C_M) - \mu_L \theta_M \Delta u(C_L) \tag{14}$$

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

$$\begin{aligned}
\Delta Q_H &\leq \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H) \theta_H - \mu_L \theta_M] u'(C_M^I) \Delta C_M - \mu_L \theta_M u'(C_L^I) \Delta C_L \\
&< \mu_M \Delta C_M + \mu_L \Delta C_L - [(1 - \mu_H) \theta_H - \mu_L \theta_M] \frac{1}{\theta_M} \Delta C_M - \mu_L \theta_M \frac{1}{\theta_L} \Delta C_L \\
&= (1 - \mu_H) \frac{\theta_M - \theta_H}{\theta_M} \Delta C_M + \mu_L \left( \frac{\theta_L - \theta_M}{\theta_L} \right) \Delta C_L < 0 = \Delta C_H
\end{aligned} \tag{15}$$

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

- (b) Neither DIC binds under independent taxation. The equation (13) still holds under unified taxation. By the nonbinding DIC's, under independent taxation, the LHS is strictly less than RHS in (13). As a result, the LHS is strictly less than RHS in (14). The rest of the proof is the same as in the previous subcase, except that in (15) the first inequality is replaced by a strict inequality, and the second inequality is replaced by an equality.
- (c) DIC-H is slack but DIC-M binds under independent taxation. The proof is a combination of subcases 1(a) and 1b.

2. There is bunching under unified taxation.

- (a) Both DIC's bind and there is no bunching under independent taxation.

With  $C_M^U = C_L^U = C_{ML}^U$  and  $Q_M^U = Q_L^U = Q_{ML}^U$ , equation (13) still holds (though can be simplified further) under unified taxation. The rest of the proof follows exactly that of subcase 1(a).

- (b) Both DIC's bind and there is bunching under independent taxation. With  $C_M^U = C_L^U = C_{ML}^U$ ,  $Q_M^U = Q_L^U = Q_{ML}^U$ ,  $C_M^I = C_L^I = C_{ML}^I$ , and  $Q_M^I = Q_L^I = Q_{ML}^I$ , equation (13) still holds (though can be simplified further) under both taxation regimes. The rest of the proof follows exactly that of subcase 1(a).

- (c) At least one DIC does not bind under independent taxation. The proof follows exactly that of subcases 1(b) and 1(c), with slight modifications of notation. ■

#### Proof of Proposition 4:

1. No bunching under unified taxation.

- (a) Both DIC's bind under independent taxation. First, we show that  $T_L^I > T_L^U$ . Suppose  $T_L^I \leq T_L^U$ . Then  $Q_L^I - Q_L^U \leq C_L^I - C_L^U$ .

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

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  $\tilde{C} \in [C_L^U, C_L^I]$ . The last inequality

holds since  $u'(\tilde{C}) \geq u'(C_L^I) > 1/\theta_L$  and  $C_L^I > C_L^U$ . 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

$$\begin{aligned}
v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\
&= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\
&\geq u(C_L^I) - u(C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} \\
&= u'(\tilde{C})(C_L^I - C_L^U) - \frac{C_L^I - C_L^U}{\theta_M} > 0,
\end{aligned} \tag{16}$$

where the last inequality follows from the fact that  $u'(\tilde{C}) \geq u'(C_L^I) > 1/\theta_L > 1/\theta_M$ . 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. 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  $Q_L^I > Q_L^U$ .

Next we show that  $v_L^I < v_L^U$ . Suppose  $v_L^I \geq v_L^U$ . This implies that

$$\begin{aligned}
v_M^I - v_M^U &= u(C_M^I) - u(C_M^U) - \frac{Q_M^I - Q_M^U}{\theta_M} \\
&= u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_M} \\
&> u(C_L^I) - u(C_L^U) - \frac{Q_L^I - Q_L^U}{\theta_L} \\
&= v_L^I - v_L^U \geq 0.
\end{aligned} \tag{17}$$

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  $v_L^I < v_L^U$ .

- (b) Neither DIC binds under independent taxation. The proof is very similar to that for subcase 1(a). 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 (16) 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 (17) should be replaced by a strict inequality.

- (c) DIC-H is slack but DIC-M binds under independent taxation. Given that DIC-M binds, the proof is exactly the same as in subcase 1(a).

2. There is bunching under unified taxation.

- (a) Both DIC's bind and there is no bunching under independent taxation. The proof is similar to that in subcase 1(a). The proof that  $T_L^I \leq T_L^U$  implies  $v_L^I > v_L^U$  is exactly the same. The proof that  $T_L^I \leq T_L^U$  implies  $v_M^I > v_M^U$  still goes through. This is because with bunching DIC-M (and UIC-M) binds under unified taxation. By the same logic, the proof that  $v_L^I \geq v_L^U$  leads to  $v_M^I - v_M^U \geq 0$  holds as well.
- (b) Both DIC's bind and there is bunching under independent taxation. The proof is similar to (and simpler than) that in subcase 1(a). The proof that  $T_L^I \leq T_L^U$  implies  $v_L^I > v_L^U$  is exactly the same, while that  $T_L^I \leq T_L^U$  implies  $v_M^I > v_M^U$  and that  $v_L^I \geq v_L^U$  leads to  $v_M^I - v_M^U \geq 0$  follow immediately as types  $M$  and  $L$  are pooled together under both taxation regimes.
- (c) At least one DIC does not bind under independent taxation. The proof follows exactly that of subcases 1(b) and 1(c), with slight modifications of notation. ■

## Appendix B: The Continuous-type Model

For robustness check, in this section we extend our analysis to the continuous type model, which can be regarded as the limiting case of many finite types. As an overview, 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 (types sufficiently close to the lowest type) receive lower subsidy under competition; there exists a cutoff type  $\theta^*$  so that all types above  $\theta^*$  are better off, and all types below  $\theta^*$  are worse off with competition. Our computations confirm most of the findings from the three type model regarding the preferences of the median type, who is responsible for the constitutional choice.

Specifically, in the vertical dimension worker-consumers are distributed on  $[\underline{\theta}, \bar{\theta}]$  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  $\{C(\theta), Q(\theta)\}_{\theta \in [\underline{\theta}, \bar{\theta}]}$ . Define the tax function  $T(\theta) = Q(\theta) - C(\theta)$ . In the autarkic economy (no tax), a citizen's optimal consumption is determined by (5).

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

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

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

$$V(\theta, \theta) \geq V(\theta, \hat{\theta}) \quad \forall (\theta, \hat{\theta}) \in [\underline{\theta}, \bar{\theta}]^2.$$

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

$$v'(\theta) = \frac{Q(\theta)}{\theta^2} = \frac{1}{\theta} [u(C(\theta)) - v(\theta)] \quad (18)$$

$$Q'(\theta) \geq 0 \quad (19)$$

Constraint (19) is the monotonicity requirement as in the three-type model.

By (18), given  $v(\theta)$ ,  $Q(\theta)$  is uniquely determined and so is  $C(\theta)$ . For convenience, we will work with the rent provision contract  $v(\theta)$ .<sup>34</sup> 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- $\theta$  "market share" for the State in question is given by

$$x^*(\theta) = 1 + \frac{v(\theta) - v_{-i}(\theta)}{k}. \quad (20)$$

For ease of analysis, from now on we will work with the utility function  $u(C) = 2\sqrt{C}$ .<sup>35</sup>

---

<sup>34</sup>This approach follows the lead of Armstrong and Vickers (2001), who model firms as supplying utility directly to consumers.

<sup>35</sup>Our main results should not be altered as long as we work with concave utility functions.

## 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  $w(\theta) = f(\theta)$  (in the same spirit as in the three-type model). 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:

$$\begin{aligned} & \max \int_{\underline{\theta}}^{\bar{\theta}} v(\theta) f(\theta) d\theta \\ & \text{s.t. } v'(\theta) = \frac{1}{\theta} \left[ 2\sqrt{C(\theta)} - v(\theta) \right] \\ & Q'(\theta) \geq 0 \\ & \int_{\underline{\theta}}^{\bar{\theta}} [Q(\theta) - C(\theta)] f(\theta) d\theta = 0 \end{aligned}$$

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

$$\begin{aligned} J(\theta) &= \int_{\underline{\theta}}^{\theta} [Q(\theta) - C(\theta)] f(\theta) d\theta, \text{ hence} \\ J'(\theta) &= [Q(\theta) - C(\theta)] f(\theta). \end{aligned}$$

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

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

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  $\lambda$  and  $\mu$  are the two costate variables. The optimality conditions are as follows:

$$\frac{\partial H}{\partial z} = 2\frac{\lambda}{\theta} + \mu[2\theta - 2z]f = 0 \tag{21}$$

$$\lambda' = -\frac{\partial H}{\partial v} = -f + \frac{\lambda}{\theta} + \mu\theta f \tag{22}$$

$$\mu' = -\frac{\partial H}{\partial J} = 0 \tag{23}$$

From (23),  $\mu$  is a constant. From (21) and (22) we can get rid of  $\lambda$  to yield

$$z' + \frac{f'}{f}z = 2 - \frac{1}{\mu\theta} + \frac{f'}{f}\theta \quad (24)$$

We can further getting rid of  $\mu$  by turning (24) into a second-order differential equation:

$$\begin{aligned} z'' &= -\frac{1}{\theta} \left[ z' - 2 + (z + \theta z' - 2\theta) \frac{f'}{f} + \theta(z - \theta) \left( \frac{f'}{f} \right)' \right] \\ z(\underline{\theta}) &= \underline{\theta}, \quad z(\bar{\theta}) = \bar{\theta} \end{aligned} \quad (25)$$

where the boundary conditions above are directly implied from the transversality conditions  $\lambda(\underline{\theta}) = \lambda(\bar{\theta}) = 0$  and (21). The above second-order (linear) differential equation system has a closed-form solution, which is given by

$$\begin{aligned} z(\theta) &= \frac{f(\underline{\theta})}{f(\theta)} \left[ \int_{\underline{\theta}}^{\theta} \frac{f(s)}{f(\underline{\theta})} \left( 2 - \frac{1}{\mu s} + s \frac{f'(s)}{f(s)} \right) ds + \underline{\theta} \right], \\ \text{where } \mu &= \int_{\underline{\theta}}^{\bar{\theta}} \frac{dF(s)}{s}. \end{aligned} \quad (26)$$

## 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 (20) the type- $\theta$  "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:

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

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

$$\begin{aligned} \frac{\partial H}{\partial z} &= 2\frac{\lambda}{\theta} + \mu[2\theta - 2z]f = 0 \\ \lambda'(\theta) &= -\frac{\partial H}{\partial v} = -f - 2\frac{\lambda}{\theta} - \frac{\mu}{k} [\theta(2z - v) - z^2]f + \mu\theta f \\ \mu'(\theta) &= -\frac{\partial H}{\partial J} = 0 \Rightarrow \mu \text{ is a constant} \end{aligned}$$

After getting rid of  $\lambda$ , we have:

$$\begin{aligned} z' &= 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{\theta(2z - v) - z^2}{k\theta} \\ v' &= \frac{1}{\theta}(2z - v) \\ J' &= \theta(2z - v) - z^2 \end{aligned}$$

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

$$w' = 2z' - v' = 2z' - \frac{w}{\theta} \tag{27}$$

$$J' = \theta w - z^2 \tag{28}$$

$$z' = 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{\theta w - z^2}{k\theta} = 2 - \frac{1}{\mu\theta} - (z - \theta)\frac{f'}{f} - \frac{J'}{k\theta} \tag{29}$$

From (28), we have

$$w = \frac{1}{\theta}(J' + z^2), \tag{30}$$

$$w' = \frac{1}{\theta^2} [(J'' + 2zz')\theta - (J' + z^2)] \tag{31}$$

Substituting (30) and (31) into (27), we have

$$J'' = 2(\theta - z)z' \quad (32)$$

From (29), we have

$$J'' = 2k - k(\theta z'' + z') - k(z + \theta z' - 2\theta)\frac{f'}{f} - k\theta(\theta - z)\left(\frac{f'}{f}\right)' \quad (33)$$

Equating (32) and (33), and simplifying, we have

$$\begin{aligned} z'' &= -\frac{1}{\theta} \left[ z' - 2 + (z + \theta z' - 2\theta)\frac{f'}{f} + \theta(z - \theta)\left(\frac{f'}{f}\right)' + \frac{2}{k}(\theta - z)z' \right] \\ z(\underline{\theta}) &= \underline{\theta}, \quad z(\bar{\theta}) = \bar{\theta} \end{aligned} \quad (34)$$

where the boundary conditions above, as in the unified taxation case, follow from the transversality conditions  $\lambda(\underline{\theta}) = \lambda(\bar{\theta}) = 0$ . 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  $\theta$  is distributed uniformly over  $[\underline{\theta}, \bar{\theta}]$ .

### The Uniform Distribution Case

Under unified taxation, assuming that  $\theta$  is uniformly distributed (i.e.,  $f' = 0$ ), (25) reduces to

$$\begin{aligned} z'' &= -\frac{1}{\theta} [z' - 2] \\ z(\underline{\theta}) &= \underline{\theta}, \quad z(\bar{\theta}) = \bar{\theta} \end{aligned} \quad (35)$$

Substituting  $f(\theta) = 1/(\bar{\theta} - \underline{\theta})$  into (26), we obtain the solution in the uniform distribution case:

$$z(\theta) = 2\theta - (\bar{\theta} - \underline{\theta}) \frac{\log \theta - \log \underline{\theta}}{\log \bar{\theta} - \log \underline{\theta}} - \underline{\theta} \quad (36)$$

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

$$\bar{\theta}/\underline{\theta} \leq \gamma^* \approx 3.55 \quad (37)$$

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 (37) throughout this section.<sup>36</sup> Intuitively, the higher the  $\bar{\theta}/\underline{\theta}$ , the more costly is sorting along the vertical dimension. When  $\bar{\theta}/\underline{\theta}$  is large enough, pooling at the lower end is optimal.

It can be easily verified that  $\theta - z > 0$  for  $\theta \in (\underline{\theta}, \bar{\theta})$  and  $z = \theta$  for  $\theta = \underline{\theta}, \bar{\theta}$ . 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.<sup>37</sup>

Since  $T'(\theta) = 2(\theta - z)z'$ ,  $T'(\theta) > 0$  for  $\theta \in (\underline{\theta}, \bar{\theta})$  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  $\theta$  is uniformly distributed, (34) becomes:

$$\begin{aligned} z'' &= -\frac{1}{\theta} \left[ z' - 2 + \frac{2}{k}(\theta - z)z' \right] \\ z(\underline{\theta}) &= \underline{\theta}, \quad z(\bar{\theta}) = \bar{\theta} \end{aligned} \tag{38}$$

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  $\theta \in (\underline{\theta}, \bar{\theta})$ .

**Proof.** Define  $y(\theta) = \theta - z_I(\theta)$ . Then  $y(\underline{\theta}) = y(\bar{\theta}) = 0$ ,  $y'(\theta) = 1 - z'_I(\theta)$ , 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  $y(\bar{\theta}) = 0$ , at some point the curve has to shoot back to the zero line. So there is  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$ , such that  $y'(\hat{\theta}) = 0$  and  $y(\theta) < 0$  for all  $\theta \in (\underline{\theta}, \hat{\theta}]$ . In that case,

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

---

<sup>36</sup>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.

<sup>37</sup>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. In the literature of optimal taxation, Seade (1977) provides a good intuition for the “no-distortion-at-the-bottom” result.

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  $\theta = \bar{\theta}$ , we are done; otherwise, there is  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$ , 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  $\theta = \underline{\theta}, \bar{\theta}$ . ■

So as in the unified taxation case, consumption is also distorted downward for all but the top and the bottom types for any  $k > 0$ . Note that this is very different from a result obtained in the duopoly case in Rochet and Stole (2002), who show that when competition is sufficiently intense ( $k$  sufficiently small), quality distortions disappear completely.

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

LEMMA 6 *Suppose condition (37) holds, then  $z'_I(\theta) > 0$  and hence  $T'_I(\theta) > 0$  for any  $\theta \in [\underline{\theta}, \bar{\theta}]$ .*

**Proof.** 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.<sup>38</sup>

What remains to be shown is that  $z'_I(\underline{\theta}) > 0$ . Now compare two differential equation systems (35) and (38). 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 the boundary conditions  $z_I(\underline{\theta}) - z_U(\underline{\theta}) = z_I(\bar{\theta}) - z_U(\bar{\theta}) = 0$ , we conclude that  $z'_I(\theta) - z'_U(\theta)$  has to cross zero line exactly once. That is, there is a  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$  such that  $z'_U(\theta) < z'_I(\theta)$  for  $\theta \in [\underline{\theta}, \hat{\theta})$ , and  $z'_U(\theta) > z'_I(\theta)$  for  $\theta \in (\hat{\theta}, \bar{\theta}]$ . 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  $\theta \in (\underline{\theta}, \bar{\theta})$ . ■

---

<sup>38</sup>Therefore, if there is pooling, it must happen at the low end.

The proof of Lemma 6 suggests that whenever the optimal solution under unified taxation exhibits perfect sorting, the equilibrium under independent taxation must exhibit perfect sorting. On the other hand, it is possible that pooling occurs under unified regime but the equilibrium under independent taxation exhibits perfect sorting.<sup>39</sup> 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 5** *Let  $k_2 < k_1$ . Under independent taxation, (i)  $\theta > z_2 > z_1$  for all  $\theta \in (\underline{\theta}, \bar{\theta})$ ; (ii)  $T_1(\underline{\theta}) > T_2(\underline{\theta})$  and  $T_2(\bar{\theta}) < T_1(\bar{\theta})$ ; (iii) the tax schedule for (relatively) rich people is flatter under  $k_2$ .*

**Proof.** (i) The two differential equations under independent taxation are as follows:

$$\begin{aligned} z_1'' &= \frac{1}{\theta} \left[ 2 - z_1' - \frac{2}{k_1} (\theta - z_1) z_1' \right], \\ z_2'' &= \frac{1}{\theta} \left[ 2 - z_2' - \frac{2}{k_2} (\theta - z_2) z_2' \right]. \end{aligned} \quad (39)$$

Let  $y = z_2 - z_1$ . We have  $y(\underline{\theta}) = y(\bar{\theta}) = 0$ . We need to show that  $y(\theta) > 0$  for all  $\theta \in (\underline{\theta}, \bar{\theta})$ . 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  $y(\bar{\theta}) = 0$ , there exists  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$  such that  $y'(\hat{\theta}) = 0$  and  $y(\theta) < 0$  for all  $\theta \in (\underline{\theta}, \hat{\theta}]$ . But then it is easily verified that  $y''(\hat{\theta}) < 0$ . 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  $\hat{\theta}$  sufficiently close to  $\underline{\theta}$ , such that  $y(\hat{\theta}) = y'(\hat{\theta}) = y''(\hat{\theta}) = 0$ . However, with notation  $z(\hat{\theta}) = z_1(\hat{\theta}) = z_2(\hat{\theta})$  and  $z'(\hat{\theta}) = z_1'(\hat{\theta}) = z_2'(\hat{\theta})$ , we can demonstrate that

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

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

---

<sup>39</sup>Consider the following example.  $\theta$  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.

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  $\bar{\theta}$ , we are done with the proof; otherwise it drops below zero before reaching the end point  $\bar{\theta}$ . But then there is  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$  such that  $y'(\hat{\theta}) = 0$  and  $y(\theta) < 0$  for all  $\theta \in (\underline{\theta}, \hat{\theta}]$ . 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  $\theta$  can be ruled out. But so far the case  $v_1 < v_2$  for all  $\theta$  cannot be ruled out. Therefore, we can only show  $T_2(\bar{\theta}) < T_1(\bar{\theta})$ .

(iii) Note that we have  $z_1(\theta) < z_2(\theta)$  for any interior  $\theta$ . This implies that at the neighborhood of  $\bar{\theta}$ ,  $z'_1 > z'_2$ . As a result, in this neighborhood,  $T'_1 > T'_2$  as well. ■

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  $\bar{\theta}$ . As  $k$  goes down, the competition between two States becomes more intense. Proposition 5 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  $k \rightarrow 0$ ,  $T(\theta) = 0$ . The solution under unified taxation, on the other hand, is independent of  $k$ , which can be regarded as the limiting case when  $k \rightarrow +\infty$  (this can be seen from comparing (25) and (34)).

In Simula and Trannoy (2010), 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.<sup>40</sup> 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 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 (35) and (38). Using subscripts  $U$  and  $I$  to denote the unified and independent taxation regimes, respectively, we can state the following comparison results:

---

<sup>40</sup>Under independent taxation,  $T' = 2(\theta - z)z'$  is always positive as  $(\theta - z) \geq 0$  and  $z' > 0$ .

PROPOSITION 6 (i) There is a  $\hat{\theta} \in (\underline{\theta}, \bar{\theta})$  such that  $z'_I(\hat{\theta}) = z'_U(\hat{\theta})$ ,  $z'_I(\theta) > z'_U(\theta)$  for  $\theta \in [\underline{\theta}, \hat{\theta})$  and  $z'_I(\theta) < z'_U(\theta)$  for  $\theta \in (\hat{\theta}, \bar{\theta}]$ ; (ii)  $z_I(\theta) > z_U(\theta)$  for any  $\theta \in (\underline{\theta}, \bar{\theta})$ ; (iii)  $T'_I(\theta) < T'_U(\theta)$  for  $\theta \in (\hat{\theta}, \bar{\theta})$ .

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

Part (ii) follows from (i) given the boundary conditions  $z_I(\underline{\theta}) - z_U(\underline{\theta}) = z_I(\bar{\theta}) - z_U(\bar{\theta}) = 0$ . For  $\theta \in (\hat{\theta}, \bar{\theta}]$ , 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  $\theta \in (\underline{\theta}, \bar{\theta})$  receive strictly higher consumption. Moreover, the tax schedule is flatter for the rich (those with sufficiently high types).

PROPOSITION 7 (i) There is a  $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$  such that  $v_I(\tilde{\theta}) = v_U(\tilde{\theta})$ ,  $v_I(\theta) < v_U(\theta)$  for  $\theta \in [\underline{\theta}, \tilde{\theta})$  and  $v_I(\theta) > v_U(\theta)$  for  $\theta \in (\tilde{\theta}, \bar{\theta}]$ ; (ii)  $T_I(\underline{\theta}) > T_U(\underline{\theta})$  and  $T_I(\bar{\theta}) > T_U(\bar{\theta})$ .

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

$$v'_I - v'_U = \frac{1}{\theta} [2(z_I - z_U) - (v_I - v_U)]. \quad (40)$$

Over  $(\underline{\theta}, \bar{\theta})$ , given  $z_I > z_U$ , from (40) we have  $v'_I > v'_U$  whenever  $v_I = v_U$ . This implies that over  $(\underline{\theta}, \bar{\theta})$ ,  $v_I$  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  $\theta$  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  $v_I(\bar{\theta}) \leq v_U(\bar{\theta})$ . Suppose this is the case. Then  $v_I(\theta) < v_U(\theta)$  for all  $\theta < \bar{\theta}$ . At  $\underline{\theta}$ ,  $v_I(\underline{\theta}) < v_U(\underline{\theta})$ , which implies that  $T_I(\underline{\theta}) > T_U(\underline{\theta})$ . At  $\bar{\theta}$ ,  $v_I(\bar{\theta}) \leq v_U(\bar{\theta})$ , which implies  $T_I(\bar{\theta}) \geq T_U(\bar{\theta})$ . For any interior  $\theta \in (\underline{\theta}, \bar{\theta})$ ,

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

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  $\theta \in (\underline{\theta}, \bar{\theta})$ . Therefore,  $\int_{\underline{\theta}}^{\bar{\theta}} T_I(\theta) d\theta > \int_{\underline{\theta}}^{\bar{\theta}} T_U(\theta) d\theta$ , violating the resource constraint  $\int_{\underline{\theta}}^{\bar{\theta}} T_I(\theta) d\theta = \int_{\underline{\theta}}^{\bar{\theta}} T_U(\theta) d\theta = 0$ .

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

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  $\bar{\theta} = 2$ . 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 7. Generally speaking, higher types are taxed less and lower types get less subsidy under the independent system.

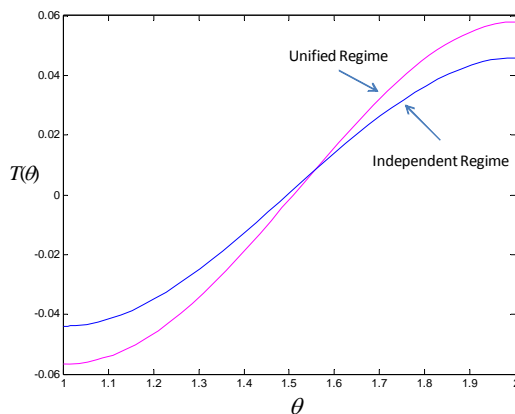


Figure 3: Tax Schedule Comparison with Uniform Distribution

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

## Constitutional Choice

With continuous types 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 (25) and (34). 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:

$$f(\theta) = \frac{\alpha\theta^{-\alpha-1}}{1-4^{-\alpha}} \text{ and } 1-F(\theta) = \frac{\theta^{-\alpha}-4^{-\alpha}}{1-4^{-\alpha}}, \theta \in [1, 4].^{41} \quad (41)$$

Note that the uniform distribution is a special case of the Pareto distribution family (with  $\alpha = -1$ ). As  $\alpha$  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.

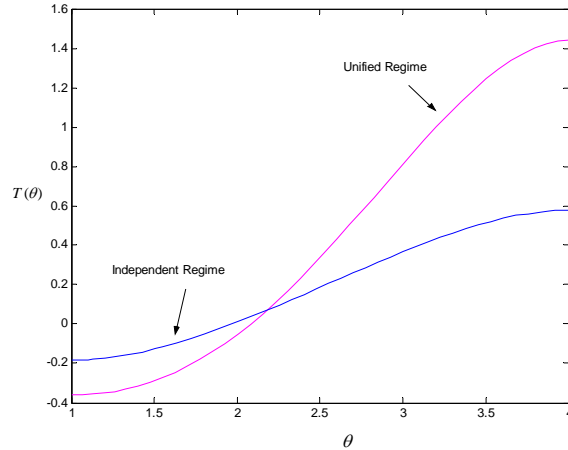


Figure 4: Tax Schedule Comparison with Pareto Distribution

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  $\theta^*$  be the indifference type at which  $v_I$  crosses  $v_U$ . Then all the types below  $\theta^*$  prefer the unified regime and all the types above  $\theta^*$  prefer the independent regime. The following table shows how the indifference type  $\theta^*$  shifts as  $k$  changes (for the truncated Pareto distribution, the computations are done based on the case  $\alpha = -0.15$ ).

Table 1: How  $\theta^*$  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  $\theta^*$  is monotonically decreasing in  $k$ . This is consistent with Result 4 in the three type model. *Therefore, as the moving cost decreases, the measure of citizens who*

<sup>41</sup>With the support of  $\theta$  being  $[1, 4]$ , the highest type's pre-tax income is 16 times that of the lowest type.

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 (41). The following table reports how the indifference type  $\theta^*$  and the median type  $\theta_m$  change as  $\alpha$  varies:

Table 2: How  $\theta^*$  and  $\theta_m$  shift as  $\alpha$  changes

$\alpha$	-0.5	-0.3	-0.2	-0.15	-0.1	0.5	1	1.5
$\theta^*$	2.136	2.0933	2.0731	2.0626	2.0519	1.9437	1.8431	1.7645
$\theta_m$	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  $\alpha$  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 Result 3. However, if  $\alpha$  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. ( $\theta^*$  is once again the cutoff type who is indifferent between the two tax regimes):

Table 3: How  $\theta^*$  shifts as inequality parameter changes

	$a = 0$	$a = 3$	$a = 5$	$a = 7$	$a = 9$
$\theta^*$	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.

## 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] Barbera, Salvador and Matthew O. Jackson (2004), “Choosing how to Choose: Self-Stable Majority Rules and Constitutions”, *Quarterly Journal of Economics*, **119-3**, 1011-48.
- [4] Boadway, Robin; Cuff, Katherine; and Marchand, Maurice (2000): “Optimal Income Taxation with Quasi-linear Preferences Revisited,” *Journal of Public Economic Theory* 2(4): 435-60.
- [5] Boadway, Robin; Marchand, Maurice; and Marianne Vigneault (1998): “The Consequences of Overlapping Tax Bases for Redistribution and Public Spending in a Federation,” *Journal of Public Economics* 68: 453-478.
- [6] Brett, Craig and John Weymark (2008): “Strategic Nonlinear Income Tax Competition with Perfect Labor Mobility,” Working Paper.
- [7] Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz: “Pareto Efficient Tax Structures,” *Oxford Economic Papers* 42, 61-77.
- [8] Diamond, Peter (1998): “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates,” *American Economic Review* 88 (1): 83-95.
- [9] Epple, Dennis, and Thomas Romer (1991): “Mobility and Redistribution,” *The Journal of Political Economy* 99 (4): 828-58.
- [10] Epstein, Larry and Michael Peters (1999): “A Revelation Principle for Competing Mechanisms,” *Journal of Economic Theory* 88: 119-60.
- [11] Guesnerie, R. and J. Laffont (1984): “A Complete Solution to a Class of Principal-Agent Problems with An Application to the Control of a Self-Managed Firm,” *Journal of Public Economics* 25, 329-69.
- [12] Guesnerie, R., Seade, J. (1982): “Nonlinear pricing in a finite economy,” *Journal of Public Economics* 17, 157–179.

- [13] 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.
- [14] Hatfield, J.W. and G. Padro i Miquel (2008): “A Political Economy Theory of Partial Decentralization,” Stanford University and BRED working paper.
- [15] Hindriks, J. (1999): “The Consequences of Labour Mobility for Redistribution: Tax vs. Transfer Competition,” *Journal of Public Economics* 74: 215–234.
- [16] Huber, Bernd (1999): “Tax Competition and Tax Coordination in an Optimal Income Tax Model,” *Journal of Public Economics* 71: 441–458.
- [17] Krause, A. (2007): “Education and Taxation Policies in the Presence of Countervailing Incentives,” *Economica*, forthcoming.
- [18] Ladd, Helen F. and Fred C. Doolittle (1982): “Which Level of Government Should Assist the Poor?” *National Tax Journal* 35 (3): 323-36.
- [19] Leite-Monteiro, M. (1997): “Redistributive Policy with Labour Mobility Across Countries,” *Journal of Public Economics* 65: 229–244.
- [20] Lindbeck, Assar and Joerg Weibull (1987): “Balanced-budget Redistribution as the Outcome of Political Competition,” *Public Choice*, 273-97.
- [21] Lollivier, Stefan; Rochet, Jean-Charles (1983): “Bunching and Second-Order Conditions: A Note on Optimal Tax Theory,” *Journal of Economic Theory* 31(2): 392-400.
- [22] Mansoorian, A. and G. Myers (1993): “Attachment to Home and Efficient Purchases of Population in a Fiscal Externality Economy,” *Journal of Public Economics* 52: 117-132.
- [23] Mansoorian, A. and G. Myers (1997): “On the Consequences of Government Objectives for Economies with Mobile Populations,” *Journal of Public Economics* 63: 265-281.
- [24] Martimort, David and Lars Stole (2002): “The Revelation and Delegation Principles in Common Agency Games,” *Econometrica* 70: 1659–1673.
- [25] Martimort, David and Lars Stole (1997): “Communication Spaces, Equilibria Sets and the Revelation Principle Under Common Agency,” unpublished working paper.

- [26] Matthews, S. and J. Moore: “Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multi-Dimensional Screening,” *Econometrica* 55, 441-467.
- [27] Mirrlees, James (1971): “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies* 38(2): 175-208.
- [28] Mirrlees, James (1982): “Migration and Optimal Income Taxes,” *Journal of Public Economics* 18: 319-341.
- [29] 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.
- [30] Osmundsen, P. (1999): “Taxing Internationally Mobile Individuals—A Case of Countervailing Incentives,” *International Tax and Public Finance* 6: 149-164.
- [31] OECD (2006): Revenue Statistics 1965-2005, Centre for Tax Policy and Administration.
- [32] Page Jr., H. Frank and Paulo K. Monteiro (2003): “Three Principles of Competitive Nonlinear Pricing,” *Journal of Mathematical Economics* 39: 63-109.
- [33] Peck, James (1997): “A Note on Competing Mechanisms and the Revelation Principle,” unpublished working paper.
- [34] Peters, Michael (2001): “Common Agency and the Revelation Principle,” *Econometrica* 69(5): 1349-72.
- [35] 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.
- [36] Piketty, Thomas and Emmanuel Saez (2006): “The Evolution of Top Incomes: A Historical and International Perspective,” *American Economic Review*, May 96 (2): 200.
- [37] 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.
- [38] Rochet, Jean-Charles (1987): “A Necessary and Sufficient Condition for Rationalizability in a Quasi-linear Context,” *Journal of Mathematical Economics* 16 (2): 191-200.

- [39] Rochet, Jean-Charles and Lars Stole (2002): “Nonlinear Pricing with Random Participation,” *Review of Economic Studies* 69(1): 277-311.
- [40] Rothschild, M. and J. Stiglitz (1976): “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information,” *Quarterly Journal of Economics* 90: 629-650.
- [41] Saez, Emmanuel (2001): “Using Elasticities to Derive Optimal Income Tax Rates,” *Review of Economic Studies* 68 (1): 205-29.
- [42] Salanie, Bernard (2003): *The Economics of Taxation*, MIT Press.
- [43] Seade, J. (1977): “On the Shape of Optimal Income Schedules,” *Journal of Public Economics* 7, 203-236.
- [44] Simula, L. and A. Trannoy (2010): “Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective,” *Journal of Public Economics* 94, 163-173.
- [45] Stähler, F. (2002): “Budget Cuts, Social Assistance and Voluntary Unemployment,” *Journal of Public Economic Theory* 4(4): 573-579.
- [46] Tiebout, Charles (1956): “A Pure Theory of Local Government Expenditure,” *Journal of Political Economy* 64: 416-24
- [47] Ticchi, Davide and Andrea Vindigni (2007): “Endogenous Constitutions,” Unpublished Working Paper, Princeton University.
- [48] Weymark, J. (1986a): “A Reduced-Form Optimal Nonlinear Income Taxes,” *Journal of Public Economics* 30, 199-217.
- [49] Weymark, J (1986b): “Bunching Properties of Optimal Nonlinear Income Taxes,” *Social Choice and Welfare* 3, 213-232.
- [50] Weymark, John (1987): “Comparative Static Properties of Optimal Nonlinear Income Taxes,” *Econometrica* 55(5): 1165-1185.
- [51] Wilson, John Douglas (1980): “The Effect of Potential Emigration on the Optimal Linear Income Tax,” *Journal of Public Economics* 14, 339–353.
- [52] Wilson, John Douglas (1992): “Optimal Income Taxation and International Personal Mobility,” *American Economic Review, Papers and Proceedings* 82, 191–196.

- [53] Wilson, John Douglas (1999): “Theories of Tax Competition,” *National Tax Journal* 52: 269-304.
- [54] Wilson, John Douglas (2006): “Income taxation and skilled migration: The analytical issues”, unpublished manuscript, Department of Economics, Michigan State University.
- [55] Yang, Huanxing and Lixin Ye (2008): “Nonlinear Pricing, Market Coverage, and Competition,” *Theoretical Economics* 3: 123-153.