Online Appendix: The Continuous-type Model of "Competitive Nonlinear Taxation and Constitutional Choice" by Massimo Morelli, Huanxing Yang, and Lixin Ye

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 θ^* so that all types above $\theta^*$are better off, and all types below θ^* 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 $[\underline{\theta}, \overline{\theta}]$ with density function $f(\theta)$, where $f(\theta)$ is continuous, strictly positive everywhere in its support. All the other assumptions are the same as those in the previous discrete type model.

As in the discrete type model, citizens can only be sorted in the vertical dimension. Thus, offering a tax schedule $T(Q)$ is equivalent to offering a menu of consumption and production pairs $\{C(\theta), Q(\theta)\}_{\theta\in[\underline{\theta},\overline{\theta}]}$. Define the tax function $T(\theta) = Q(\theta) - C(\theta)$. In the autarkic economy (no tax), a citizen's optimal consumption is determined by $u'(c^*) = 1/\theta$.

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

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

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

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

Let $v(\theta)$ denote the equilibrium rent provision to type-θ citizen: $v(\theta) = V(\theta, \theta)$. By the standard

Constraint Simplification Theorem, the IC conditions are equivalent to the following two conditions:

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

$$Q'(\theta) \geq 0 \tag{2}$$

Constraint (2) is the monotonicity requirement as in the three-type model.

By (1), given $v(\theta)$, $Q(\theta)$ is uniquely determined and so is $C(\theta)$. For convenience, we will work with the rent provision contract $v(\theta)$.[1] It can be easily verified that $Q' = \theta u'(C)C'$. Thus, as in the three-type model, $Q'(\theta) \geq 0$ if and only if $C'(\theta) \geq 0$.

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

$$x^*(\theta) = 1 + \frac{v(\theta) - v_{-i}(\theta)}{k}. \tag{3}$$

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

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:

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

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

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

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

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

---

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

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

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

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

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

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

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

$$\frac{\partial H}{\partial z} = 2\frac{\lambda}{\theta} + \mu[2\theta - 2z] f = 0 \tag{4}$$

$$\lambda' = -\frac{\partial H}{\partial v} = -f + \frac{\lambda}{\theta} + \mu\theta f \tag{5}$$

$$\mu' = -\frac{\partial H}{\partial J} = 0 \tag{6}$$

From (6), μ is a constant. From (4) and (5) we can get rid of λ to yield

$$z' + \frac{f'}{f} z = 2 - \frac{1}{\mu\theta} + \frac{f'}{f}\theta \tag{7}$$

We can further getting rid of μ by turning (7) into a second-order differential equation:

$$\begin{aligned} z'' &= -\frac{1}{\theta} \left[z' - 2 + (z + \theta z' - 2\theta)\frac{f'}{f} + \theta(z - \theta) \left(\frac{f'}{f} \right)' \right] \\ z(\underline{\theta}) &= \underline{\theta}, \; z(\overline{\theta}) = \overline{\theta} \end{aligned} \tag{8}$$

where the boundary conditions above are directly implied from the transversality conditions $\lambda(\underline{\theta}) = \lambda(\overline{\theta}) = 0$ and (4). The above second-order (linear) differential equation system has a closed-form solution, which is given by

$$z(\theta) = \frac{f(\underline{\theta})}{f(\theta)}\left[\int_{\underline{\theta}}^{\theta}\frac{f(s)}{f(\underline{\theta})}\left(2-\frac{1}{\mu s}+s\frac{f'(s)}{f(s)}\right)ds+\underline{\theta}\right], \tag{9}$$

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

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 (3) the type-θ "market share" for State 1 is given by $\eta(\theta) = 1 + \frac{1}{k}[v(\theta) - v^*(\theta)]$. Now State 1's maximization problem can be formulated as the following optimal control problem:

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

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

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

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

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

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

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

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

The optimality conditions for a symmetric equilibrium are given by

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

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

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

After getting rid of λ, 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

$$
\begin{aligned}
w' &= 2z' - v' = 2z' - \frac{w}{\theta} && (10) \\
J' &= \theta w - z^2 && (11) \\
z' &= 2 - \frac{1}{\mu\theta} - (z-\theta)\frac{f'}{f} - \frac{\theta w - z^2}{k\theta} = 2 - \frac{1}{\mu\theta} - (z-\theta)\frac{f'}{f} - \frac{J'}{k\theta} && (12)
\end{aligned}
$$

From (11), we have

$$
\begin{aligned}
w &= \frac{1}{\theta}(J' + z^2), && (13) \\
w' &= \frac{1}{\theta^2}\left[(J'' + 2zz')\theta - (J' + z^2)\right] && (14)
\end{aligned}
$$

Substituting (13) and (14) into (10), we have

$$
J'' = 2(\theta - z)z' \qquad (15)
$$

From (12), we have

$$
J'' = 2k - k(\theta z'' + z') - k(z + \theta z' - 2\theta)\frac{f'}{f} - k\theta(\theta - z)\left(\frac{f'}{f}\right)' \qquad (16)
$$

Equating (15) and (16), and simplifying, we have

$$
\begin{aligned}
z'' &= -\frac{1}{\theta}\left[z' - 2 + (z + \theta z' - 2\theta)\frac{f'}{f} + \theta(z-\theta)\left(\frac{f'}{f}\right)' + \frac{2}{k}(\theta - z)z'\right] && (17) \\
z(\underline{\theta}) &= \underline{\theta},\ z(\overline{\theta}) = \overline{\theta}
\end{aligned}
$$

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

The Uniform Distribution Case

Under unified taxation, assuming that θ is uniformly distributed (i.e., $f' = 0$), (8) reduces to

$$\begin{aligned} z'' &= -\frac{1}{\theta}\left[z' - 2\right] \\ z(\underline{\theta}) &= \underline{\theta},\ z(\overline{\theta}) = \overline{\theta} \end{aligned} \tag{18}$$

Substituting $f(\theta) = 1/\left(\overline{\theta} - \underline{\theta}\right)$ into (9), we obtain the solution in the uniform distribution case:

$$z(\theta) = 2\theta - (\overline{\theta} - \underline{\theta})\frac{\log\theta - \log\underline{\theta}}{\log\overline{\theta} - \log\underline{\theta}} - \underline{\theta} \tag{19}$$

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

$$\overline{\theta}/\underline{\theta} \leq \gamma^* \approx 3.55 \tag{20}$$

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 (20) throughout this section.[3] Intuitively, the higher the $\overline{\theta}/\underline{\theta}$, the more costly is sorting along the vertical dimension. When $\overline{\theta}/\underline{\theta}$ is large enough, pooling at the lower end is optimal.

It can be easily verified that $\theta - z > 0$ for $\theta \in \left(\underline{\theta}, \overline{\theta}\right)$ and $z = \theta$ for $\theta = \underline{\theta}, \overline{\theta}$. The result of efficiency at the top is standard in the screening literature. Efficiency at the bottom, which is implied from the transversality condition, however, is different from what we have seen from our base model with three types.[4]

[3] 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.

[4] 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.

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

Under independent taxation, given that θ is uniformly distributed, (17) becomes:

$$\begin{aligned} z'' &= -\frac{1}{\theta}\left[z' - 2 + \frac{2}{k}(\theta - z)z'\right] \\ z(\underline{\theta}) &= \underline{\theta},\ z(\overline{\theta}) = \overline{\theta} \end{aligned} \tag{21}$$

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

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

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

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

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

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

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

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

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

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

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 2 *Suppose condition (20) holds, then $z_I'(\theta) > 0$ and hence $T_I'(\theta) > 0$ for any $\theta \in [\underline{\theta}, \overline{\theta}]$.*

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.[5]

What remains to be shown is that $z_I'(\underline{\theta}) > 0$. Now compare two differential equation systems (18) and (21). Whenever $z_I' = z_U'$ (> 0), we have $z_I'' < z_U''$ (since $\theta - z_I > 0$ by Lemma 1). 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 $z_I(\underline{\theta}) - z_U(\underline{\theta}) = z_I(\overline{\theta}) - z_U(\overline{\theta}) = 0$, we conclude that $z_I'(\theta) - z_U'(\theta)$ has to cross zero line exactly once. That is, there is a $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $z_U'(\theta) < z_I'(\theta)$ for $\theta \in [\underline{\theta}, \widehat{\theta})$, and $z_U'(\theta) > z_I'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta}]$. Given that $z_U'(\underline{\theta}) > 0$, we have $z_I'(\underline{\theta}) > z_U'(\underline{\theta}) > 0$. This completes the proof for $z_I' > 0$. Given $z_I' > 0$ and $(\theta - z_I) > 0$, we have $T_I'(\theta) = 2(\theta - z_I)z_I' > 0$ for $\theta \in (\underline{\theta}, \overline{\theta})$. ■

The proof of Lemma 2 suggests that whenever the optimal solution under unified taxation exhibits perfect sorting, the equilibrium under independent taxation must exhibits perfect sorting. On the other hand, it is possible that pooling occurs under unified regime but the equilibrium under independent taxation exhibits perfect sorting.[6] 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:

---

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

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

Proposition 1 *Let $k_2 < k_1$. Under independent taxation, (i) $\theta > z_2 > z_1$ for all $\theta \in (\underline{\theta}, \overline{\theta})$; (ii) $T_1(\underline{\theta}) > T_2(\underline{\theta})$ and $T_2(\overline{\theta}) < T_1(\overline{\theta})$; (iii) the tax schedule for (relatively) rich people is flatter under k_2.*

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

$$\begin{aligned} z_1'' &= \frac{1}{\theta}[2 - z_1' - \frac{2}{k_1}(\theta - z_1)z_1'], \\ z_2'' &= \frac{1}{\theta}[2 - z_2' - \frac{2}{k_2}(\theta - z_2)z_2']. \end{aligned} \tag{22}$$

Let $y = z_2 - z_1$. We have $y(\underline{\theta}) = y(\overline{\theta}) = 0$. We need to show that $y(\theta) > 0$ for all $\theta \in (\underline{\theta}, \overline{\theta})$. The proof idea resembles that of Lemma 1.

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

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

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

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

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

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

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

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

By continuity, we also have $T_1(\theta) > T_2(\theta)$ for types sufficiently close to $\underline{\theta}$, and $T_2(\theta) < T_1(\theta)$ for types sufficiently close to $\overline{\theta}$. As k goes down, the competition between two States becomes more intense. Proposition 1 suggests that as mobility (or competition) increases, the consumption distortion is reduced, the rich (types sufficiently close to the top) pay less taxes, and the poor (types sufficiently close to the bottom) receive less subsidies. While these results are obtained computationally in our three type model, they are obtained analytically in this continuous type model. Thus the result that *increased mobility leads to lower progressivity* is a fairly robust prediction. As in the three type model, as $k \to 0$, $T(\theta) = 0$. The solution under unified taxation, on the other hand, is independent of k, which can be regarded as the limiting case when $k \to +\infty$ (this can be seen from comparing (8) and (17)).

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.[7] 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 (18) and (21). Using subscripts U and I to denote the unified and independent taxation regimes, respectively, we can state the following comparison results:

Proposition 2 *(i) There is a $\widehat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $z_I'(\widehat{\theta}) = z_U'(\widehat{\theta})$, $z_I'(\theta) > z_U'(\theta)$ for $\theta \in [\underline{\theta}, \widehat{\theta})$ and $z_I'(\theta) < z_U'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta}]$; (ii) $z_I(\theta) > z_U(\theta)$ for any $\theta \in (\underline{\theta}, \overline{\theta})$; (iii) $T_I'(\theta) < T_U'(\theta)$ for $\theta \in (\widehat{\theta}, \overline{\theta})$.*

Proof. Part (i) is established in the proof of Lemma 2.

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

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

Proposition 3 *(i) There is a $\widetilde{\theta} \in (\underline{\theta}, \overline{\theta})$ such that $v_I(\widetilde{\theta}) = v_U(\widetilde{\theta})$, $v_I(\theta) < v_U(\theta)$ for $\theta \in [\underline{\theta}, \widetilde{\theta})$ and $v_I(\theta) > v_U(\theta)$ for $\theta \in (\widetilde{\theta}, \overline{\theta}]$; (ii) $T_I(\underline{\theta}) > T_U(\underline{\theta})$ and $T_I(\overline{\theta}) > T_U(\overline{\theta})$.*

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

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

$$v_I' - v_U' = \frac{1}{\theta}[2(z_I - z_U) - (v_I - v_U)]. \tag{23}$$

Over $(\underline{\theta}, \overline{\theta})$, given $z_I > z_U$, from (23) we have $v_I' > v_U'$ whenever $v_I = v_U$. This implies that over $(\underline{\theta}, \overline{\theta})$, v_I and v_U cross at most once, and at the intersection v_I must cross v_U from below.

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

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

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

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

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

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

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

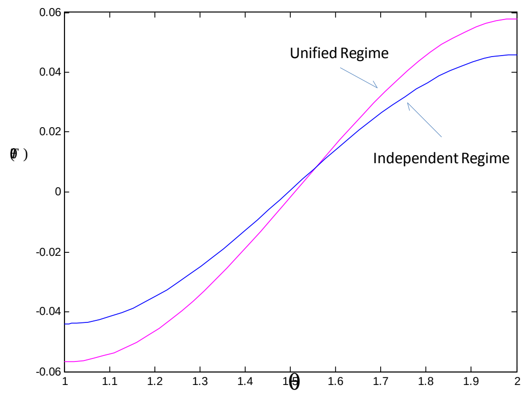

Figure 3: Tax Schedule Comparison with Uniform Distribution

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 (8) and (17). 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:

$$f(\theta) = \frac{\alpha\theta^{-\alpha-1}}{1-4^{-\alpha}} \text{ and } 1-F(\theta) = \frac{\theta^{-\alpha}-4^{-\alpha}}{1-4^{-\alpha}}, \theta \in [1,4].^{8} \tag{24}$$

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

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

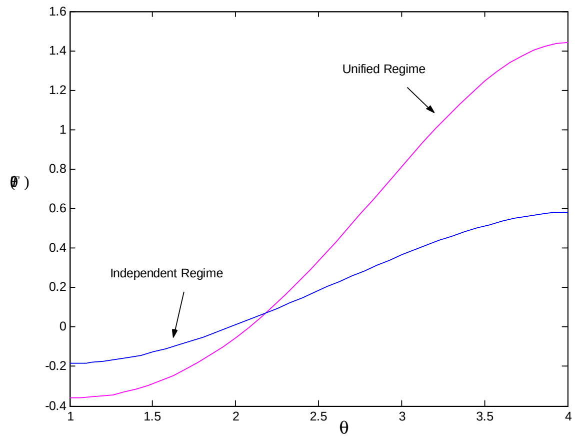

Figure 4: Tax Schedule Comparison with Pareto Distribution

Recall that with uniform distribution we established that the utility schedule v_I crosses v_U once from below. Our computation shows that this pattern of single crossing holds for truncated Pareto distributions as well. Let θ^* be the indifference type at which v_I crosses v_U. Then all the types below θ^* prefer the unified regime and all the types above θ^* prefer the independent regime. The following table shows how the indifference type θ^* shifts as k changes (for the truncated Pareto distribution, the computations are done based on the case $\alpha = -0.15$).

Table 1: How θ^* shifts as k changes

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

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

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

We are also interested in how changes in the (type) income distribution affect the constitutional choice. Fix $k = 0.5$, and consider the truncated Pareto distributions given in (24). The following table reports how the indifference type θ^* and the median type θ_m change as α varies:

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

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

For all the cases we examined, the solutions exhibit perfect sorting. Two observations are worth noting. First, as α increases (more poor around), the *indifferent type* monotonically decreases. Again this is consistent with what we found from the three type model. This is intuitive: having more poor implies more taxes from the higher types in the unified regime, while in the independent regime the solution is closer to autarky. Therefore, the indifference type will decrease, as in Result **??**. However, if α is sufficiently large ($\alpha > -0.15$), the *median type* prefers the unified regime. Thus having more poor people in this continuous type case makes the choice of the unified system more likely, which seems to be inconsistent with our finding in the three type model. This happens in this Pareto distribution case simply because the *indifference type* decreases slower than the *median type*: as the size of the poor increases, the median type becomes even poorer. This observation highlights a difference between our three-type model and the continuous type model, that is, *the median type is generically different from the type who is indifferent between the various constitutional choices, and they vary at different rates when the parameters change.*

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

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

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

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

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

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