Endogenous Income Distribution, Stratification and Fiscal Decentralization.*

David Bardey[†], Fernando Jaramillo[‡] and Fabien Moizeau[§]

February 2013.

Abstract

We develop a multijurisdiction model where individuals are heterogenous with respect to their productivity. The key feature of the framework is that before moving to a particular jurisdiction where the amount of local public good is determined by the median voter, individuals choose their level of effort which determines their income. Our findings suggest that the equilibrium is productivity-stratified, *i.e.* jurisdictions are inhabited by individuals with similar productivity. Further, the equilibrium level of effort is jurisdiction-dependent. It turns out that two individuals who are close in the productivity ladder earn dramatically different incomes if they do not reside in the same jurisdiction. Third, we study the planner's problem and characterize optimal allocations. Fourth, we study the design of the tax structure that implements optimal allocations in spite of asymmetric information with respect to productivity and effort. In particular, the optimal tax structure is such that externalities generated by free mobility of individuals are internalized. Finally, we analyze a computational model and focus on the effect of fiscal decentralization on income stratification, income inequality and welfare.

*We are very grateful to Nicolas Gravel for his valuable feedback. We also thank seminar participants at TEPP Federation (Paris), Universidad de Los Andes (Bogota), Universidad del Rosario (Bogota), Univeristé de Rennes 1, Univesité de Caen Basse-Normandie, as well as conference participants at the 2012 Journées Louis-André Gérard-Varet (Marseille).

[†]Universidad de Los Andes and Toulouse School of Economics (GREMAQ), email: david.bardey@gmail.com.

[‡]Universidad del Rosario, email: fernando.jaramillo@urosario.edu.co.

[§]CREM (Condorcet Center), Université de Rennes 1 and Institut Universitaire de France, email: fabien.moizeau@univ-rennes1.fr.

1 Introduction

Over the last decades, there has been a trend toward greater fiscal decentralization in most developed countries (see Arzaghi and Henderson, 2002, OECD, 2008, and the survey of Epple and Nechyba, 2004). It is well recognized that decentralized government activity in a context of increasing mobility of factor and population impacts the extent of segregation by income across local jurisdictions as well as income inequality. The theoretical analysis of fiscal decentralization dates back from the Tiebout model (1956) which examines horizontal government competition. The main result is that when perfectly mobile individuals have to choose among several communities, each one offering a particular taxation/public good package, the opportunity of migrating freely and without cost makes them choose the community they prefer. If individuals are heterogeneous only with income, then the "free mobility" equilibrium is income stratified. As income distribution is assumed exogenous, the Tiebout Model is silent about the effect of fiscal decentralization on income inequality. The literature on human capital accumulation with local externalities, pioneered by Bénabou (1996a,b) and Durlauf (1996), has analysed the consequences of decentralization of educational services on income dynamics. It is shown that the "free mobility" equilibrium leads to a polarized income distribution. From an empirical point of view, recent works that attempt to structurally estimate equilibrium models of local public goods find strong evidence for income stratification between local jurisdictions (see, for instance, Epple, Romer and Sieg, 2001, Bayer and Mc Millan, 2011). The picture is blurred regarding the impact of fiscal decentralization on income inequality. The reason lies on the role to be played by subnational governments in redistributive policies. For instance, using a panel of 34 developing countries on five five-year periods between 1976-2000, Sepulveda and Martinez-Vasquez (2011) find that fiscal federalism has a statistically significant effect on income inequalities but the sign and magnitude depend on the size of government sector in the economy. Sacchi and Salotti (2011), using a different sample of 23 OECD countries over the period 1971-2000, obtain a clear-cut result as they find that a higher degree of tax decentralization is associated with an increase of inequality within a country.

In this paper, we revisit the analysis of the consequences of fiscal decentralization on income inequality and income segregation. We develop a multijurisdiction model where individual income is not given *a priori* but depends on both a personal attribute and effort. Hence, our model stresses a key interplay between income inequality and income segregation. On the one hand, individual effort is assumed to be a determinant of individual income and thus influences the capacity to choose a particular bundle of local taxation and public good and to segregate. On the other hand, the return to individual effort turns out to depend on jurisdiction membership. This relationship sheds new light on the impact of decentralized provision of public goods on stratification and inequality.

Our approach on the key role played by individual effort is inspired by empirical results in the labor economics literature that emphasize the importance of local price variation in estimating either labor supply or educational effort. For instance, using 1990 U.S. census data on labor supply in the nation's largest 50 cities, Black, Kolesnikova and Taylor (2008) find that the correlation between labor supply and non-labor income differs across cities suggesting that labor decisions are location-dependent. Regarding returns to education, Black, Kolesnikova and Taylor (2009) for U.S. data that returns to college are overestimated for high-amenity locations when local prices are ignored. This result supports the evidence that returns to education differ across locations and that education decisions are location-dependent.

More specifically, we develop a model *à la* Tiebout where individuals are heterogenous with respect to their productivity. The key feature is that they choose an effort level -which can be interpreted as an educational effort or labor supply- that determines their income which in turn is crucial for their subsequent jurisdiction's choice. We assume that both productivity and effort are private information. Once in a jurisdiction, individuals vote on the level of local tax rate financing the local public good. Given this sequence of events, individuals while choosing their effort take into account their prospect of jurisdiction membership as their effort decision determines their income which is a key determinant of their local public good demand. This framework allows us to study the interaction between income distribution and social segregation. On the one hand, effort decisions determine the income distribution which shapes the social segregation. On the other hand, effort decisions depending on the prospect of jurisdiction membership, the social stratification in turn drives incentives to exert effort and may engender a particular pattern of income inequality.[1]

We first provide a characterization of the equilibrium. We show that the equilibrium is *productivity-stratified*, *i.e.* each jurisdiction is inhabited by individuals with productivity in a single interval. Jurisdictions are vertically differentiated and the degree of substitutability between private and local public goods is key to know how jurisdictions are ordered with respect to their local tax rate-local public good bundle. When private and local public goods are complements (substitutes), the "more productivite" a jurisdiction is, the higher (lower) is the local tax rate-local public good bundle. The equilibrium level of effort is also *jurisdiction-dependent* as it depends on the local tax rate determined by the jurisdiction median voter. Hence, the equilibrium level of effort depends on the equilibrium distribution of individuals among the jurisdictions. Further, individual income being monotonic with respect to productivity, the equilibrium is also *income-stratified*, *i.e.* individuals in the same jurisdiction have similar income. A direct consequence of the fact that effort is

[1] We ignore any peer-group effects that could be generated by the neighborhood. We also assume away housing markets. The main implications would remain in a more general setup.

jurisdiction-dependent is that two individuals who are close in the productivity ladder but do not belong to the same jurisdictions earn dramatically different incomes.

Second, we provide a normative analysis, mainly focused on the efficiency issue. We study the program of a social planner who is utilitarian and who decides levels of private consumption, public good consumption, income and jurisdiction's membership. We assume that the social planner is constrained by the the capacity of individuals to move freely across jurisdictions and must determine an allocation such that individuals have no incentives to move. At the optimum, we show that the provision of the local public good must satisfy the traditional Bowen-Lindhal-Samuelson condition. Efficient sorting is such that the migration of a boundary individual between two adjacent jurisdictions does not generate a net variation of resources in these jurisdictions that could make better off individuals residing in these jurisdictions. Hence, externalities generated by individual jurisdiction choice are internalized at the optimum.

Third, we show that, in spite of asymmetric information on both productivity and effort, there exists a tax structure designed at the central level that allows to decentralize the optimum, *i.e.* in the spirit of the first welfare theorem, one allocation of the contract curve being achieved. This tax structure combines a linear grant and a lump-sum tax that differs according to the jurisdiction membership. The grant exactly amounts the local tax paid in order to offset distorsions generated by the public good provision on the individuals' effort. As this grant is linear, within jurisdictions, nobody has incentives to mimic agents characterized by different productivity levels. This grant component is clearly anti-redistributive as the richest individuals receive a higher reimbursement than the poorest within each jurisdictions. Moreover, similarly to an optimal taxation scheme with tagging (see Cremer *et al.,* 2011), the social planner takes advantage of the fact that jurisdiction membership is observable and that each jurisdiction covers an interval of the productivity distribution to use a jurisdiction-dependent lump-sum tax. Indeed, it is precisely the membership which provides information on productivity levels and the linear grant that allows the central planner to achieve a first-best allocation despite asymmetric information. Further, the lump-sum tax is designed so that the local public good is efficiently provided. It is also shown that the lump-sum tax works as an instrument that allows the central planner to implement a redistribution from the richer to the poorer jurisdictions. The magnitude of the lump-sum tax is such that the migration of a boundary individual between two adjacent jurisdictions cannot make any individual better off. Precisely, the variation in the lump-sum tax paid by the migrating individual exactly equals the variation in the marginal costs of public good production in both adjacent jurisdictions. Hence, total resources do not vary implying that no individual can be made better off by this migration.

Finally, we proceed to numerical simulations of our model in order to study the impact of an increase of fiscal decentralization -assimilated to an increase in the number of jurisdictions- on total production, inequality and welfare. Results are threefold. First, it turns out that total production (defined by the sum of effort over all individuals) increases with decentralization. This result relies on the fact that an increase in the number of jurisdictions may lead to exert higher effort as it provides opportunities for individuals to secede from less productive agents and cluster with closer peers. Second, depicting different Lorenz curves for different values of the number of jurisdictions reveals that income inequality increases with the degree of fiscal decentralization. This result is consistent with the empirical evidence provided in Sacchi and Salotti (2012). We interpret this result saying that, due to the fact that income is jurisdiction-dependent, an increase in the number of jurisdictions may potentially exacerbate income heterogeneity. Third, the level of welfare decreases with the degree of fiscal decentralization. Although the degree of fiscal decentralization increases total production, individuals may not be better off as the potential gains from sorting into more homogenous jurisdictions may be outweighed by larger costs of effort. This result is in sharp contrast with the intuition that Tiebout competition provides efficiency gains compared to centralization.

Our paper belongs to the extensive literature on local public goods that examines conditions under which stratification arises (see, among others, Westhoff, 1977, Epple, Filimon and Romer, 1993, Fernandez and Rogerson, 1996, Hansen and Kessler, 2001, Gravel and Thoron, 2007). We differ from these works as we introduce in the standard Tiebout model effort decisions and thus endogenize income distribution. As we address the issue of the design of fiscal instruments implemented at the central level to decentralize efficient allocations, we are closely related to Biswas, Gravel and Oddou (2011). These authors examine the impact on income stratification arising at equilibrium when the central government uses equalization transfers in order to maximize some welfare function. We depart from their work as the design of fiscal instruments must also take into account their impact on incentives to exert effort. We are also closely related to Calabrese, Epple and Romano (2012) who study the welfare effects of fiscal decentralization. We draw the same conclusion that Tiebout competition may generate welfare losses compared to the central provision of public goods but our approach highlights different inefficiencies affecting incentives to exert effort.

Our model is also related to the strand of literature that emphasizes the role of local human capital externalities on inequality dynamics and productivity growth (see Bénabou, 1996a,b, Durlauf, 1996, Cooper, 1998, Kempf and Moizeau, 2009). These works focus on the interplay between income inequality and income segregation in dynamic frameworks where the current income inequality shapes the income segregation pattern which in turn drives the dynamics of inequality and the subsequent income distribution. We depart from these works as the interplay between inequality and segregation is static and relies on the choice of individual effort. Hence, our model allows us to discuss the impact of segregation on incentives to exert effort. We can mention that Bénabou (1996b) also studies effort choices as individuals allocate their time between work and education of their offspring. However, the equilibrium level of effort is not jurisdiction-dependent in his model.

Finally, our approach is also related to the literature on local labor markets (see the survey of Moretti, 2011) which integrates the spatial dimension in labor markets. We provide a unified analysis of labor supply decisions, migration choices among jurisdictions and voting decisions over local taxation. This general equilibrium approach could be viewed as a first step to better understand the effects of the decentralized provision of public goods on incentives to supply labor and on labor markets performance.

The plan of the paper is as follows. In the following section, we present the theoretical setup. In section 3, we study the equilibrium properties of the model. Section 4 provides the normative analysis of the model. Section 5 concludes.

2 The Set Up

Our model builds on Westhoff (1977)'s multijurisdiction model. The city is composed of J jurisdictions indexed by $j = 1, ..., J$. Each jurisdiction j provides a local public good q_j financed by a proportional tax rate τ_j on income. We consider individuals who are characterized by a productivity θ. This parameter is distributed according to the continuous cumulative density function $H(.)$ and corresponding density function $h(.)$ over the interval Θ=$[\underline{\theta}, \overline{\theta}] \in \mathbb{R}_{++}$. We normalize the mass of individuals to one. Exerting the effort l, an individual θ earns the gross income, $y(\theta) = \theta l$. l could be viewed either as human capital investment or as labor supply, θl being the income earned on the labor market. We assume that neither θ nor l is separately observable while income y is observable.

Let I denotes the total income in the economy. An individual θ who lives in jurisdiction j has the following utility function:

$$\mathcal{U}_j(\theta) \equiv u(c) + \beta u(q_j) - v\left(\frac{y}{\theta}\right),$$

with β a positive parameter and c the private good consumption. $u(.)$ is increasing and concave. We also assume that $\lim_{x \longrightarrow 0} u'(x) = +\infty$. The disutility of effort $v(.)$ is increasing and convex. Moreover, we set $v(0) = 0$. Let us stress that the separability assumption will prove to be crucial for effort choice.

We denote by n_j the size of jurisdiction j, $\zeta(n_j)$ the congestion cost in jurisdiction j with $\zeta'(.) > 0$ and by μ_j the average income in jurisdiction j, we have

$$c_j(y(\theta)) = (1-\tau_j)y, \tag{1}$$

$$\text{and } q_j = \frac{\tau_j n_j \mu_j}{\zeta(n_j)}. \tag{2}$$

Our framework is built on the following assumptions on preferences and congestion cost.

Assumption 1: If $(c, q, y) >> 0$ then $u(c) + \beta u(q) - v(y/\theta) > u(\overline{c}) + \beta u(0) - v(\overline{y}/\theta)$ over all $(\overline{c}, \overline{y}) \geq 0$.

This assumption will be necessary to obtain jurisdictions characterized by population of a size higher than a lower bound at equilibrium.

Assumption 2a: $-u''c/u' > 1$, $\forall c$.
Assumption 2b: $-u''c/u' < 1$, $\forall c$.

These alternative assumptions imply that the slope of the indifference curves of two individuals with different incomes cross at most once. Hence, denoting the slope by $S(q, \tau, y)$, which at point (q_j, τ_j) equals:

$$S(q_j, \tau_j, y) = \frac{\beta u'(q_j)}{y u'(c_j)}$$

we have,

$$\frac{\partial S(q_j, \tau_j, y)}{\partial y} = -\frac{\beta u'(q_j)[u'(c_j) + c_j u''(c_j)]}{[y u'(c_j)]^2} > (<) 0 \text{ under Assumption 2a (2b).}$$

As already stressed by Hansen and Kessler (2001), the slope of an indifference curve through any point of policy space (q, τ) increases (decreases) when the private and the public goods are complements (substitutes), *i.e.* when $-u''c/u' > 1$ $(-u''c/u' < 1)$ (see also the Gross Substitutability/Complementarity condition in Gravel and Thoron, 2007). These assumptions will thus allow us to rank individual preferences over bundles (q, τ) according to individual income. We will see that these assumptions imply that given jurisdiction membership, the agents' effort is a monotonous function of the local tax rate.

For instance, the following specific form

$$\mathcal{U}_j(\theta) = \frac{c^{1-\sigma}}{1-\sigma} + \beta \frac{q^{1-\sigma}}{1-\sigma} - \left(\frac{y}{\eta}\right)^{\eta}, \text{ with } \sigma > 0 \text{ and } \eta > 1$$

would satisfy all the above assumptions[2].

Assumption 3: $\partial \left(n_j/\zeta\left(n_j\right)\right)/\partial n_j > 0.$

Assumption 3 amounts to say that there are economies of scale. This cost function belongs to the class of cost functions considered by Westhoff (1977) that facilitates the equilibrium existence.

3 Equilibrium Analysis

We consider the following sequence of events:

Stage 1: Individuals choose their effort;

Stage 2: Individuals choose the jurisdiction where they decide to live;

Stage 3: Once jurisdictions are formed, individuals vote on local taxation.

Individuals have rational expectations and thus anticipate all continuation equilibrium values.

This sequence of events captures the key feature of our model. By considering that effort is decided first, it is assimilated as an investment which return is the jurisdiction membership allowed by the level of earned income. It turns out that the prospect to locate in a particular jurisdiction is driving the incentives an individual faces while choosing his effort. We thus depart from the literature on jurisdiction formation (see for instance Westhoff, 1977, Fernandez and Rogerson, 1996 and Hansen and Kessler, 2001) as we endogenize the distribution of income before individuals decide where to live.

Further, the crucial assumption we make while considering this sequential formulation is that individuals vote taking both residential choices and labor supplies as given. In this respect, we are thus following most models of the multijurisdiction literature by assuming "myopic voting" [3]. We thus do not consider tax competition between jurisdictions.

We solve this model backward and look at a Subgame Perfect Nash Equilibrium (SPNE hereafter) of the extensive form game defined as follows.

Definition 1 *An equilibrium with J jurisdictions and endogenous income distribution is such that: (i) each jurisdiction has a strictly positive mass of individuals, (ii) in each jurisdiction, the local*

[2] Assuming the following utility function

$$\mathcal{U}_j(\theta) \equiv u(c) + \beta f(q_j) - v\left(\frac{y}{\theta}\right)$$

would require to consider four combinations of relative risk aversions, $-cu''/u' \lesseqgtr 1$ and $-qf''/f' \lesseqgtr 1$. This would not change the key mechanisms of our framework.

[3] Epple and Romer (1991) is a notable exception that develops a model where individuals are fully aware of the migration effects on local public policies.

budget constraint is balanced and the local tax rate is chosen by the jurisdiction's median voter, (iii) no individual has an incentive to move to another jurisdiction and (iv) each individual chooses optimally her effort.

We call a free mobility equilibrium the SPNE of the subgame defined by stages 2 and 3. We will focus on non-symmetric equilibria, *i.e.* equilibria such that jurisdictions differ with respect to their bundles (τ_j, q_j). We do not pay attention to any symmetric equilibrium which always exists in this set-up (see, for instance, Fernandez and Rogerson, 1996, Calabrese, Epple and Romano, 2012).

Stage 3. We first begin by the voting decision on the local tax rate. The most preferred local tax rate for an individual with income y living in jurisdiction j satisfies the following first-order condition[4]:

$$yu'\left((1-\tau_j)y\right) = \frac{n_j \mu_j}{\zeta\left(n_j\right)} \beta u'(q_j). \tag{3}$$

We now turn to some comparative statics which will be useful to the analysis of the equilibrium.

Lemma 1 *Under Assumption 2a (2b) (i)* $\partial \tau_j / \partial y > (<) 0$, *(ii)* $\partial \tau_j / \partial \mu_j < (>) 0$, *(iii)* $\partial q_j / \partial \mu_j > 0$.

Proof. See Appendix. ■

This Lemma is standard in the literature (see for instance Fernandez and Rogerson, 1996, and Hansen and Kessler, 2001). The two first items are direct consequences of Assumptions 2a and 2b. First, according to item (i) under Assumption 2a (2b), other things being equal, richer individuals who consume a higher level of private good consumption are, due to complementarity (substituability) in favor of high (low) levels of public good and are more (less) willing to pay for the public good. Second, item (ii) of Lemma 1 also relies on Assumption 2a (2b) which implies that a marginal increase in the average income in jurisdiction j, other things being equal, increases the level of local public good and due to complementarity (substitutability) make individuals more (less) willing to consume the private good. Hence, the preferred level of local taxation is lowered (increased). Finally, a marginal increase in the average income of jurisdiction j generates two effects on the production of the public good. On the one hand, given the *per capita* tax rate τ_j and the size of the population, the richer the jurisdiction, the higher the proceeds of taxation and the higher is the level of public good. On the other hand, under the case of complementarity (substitutability) the richer the jurisdiction, the lower (higher) is the *per capita* tax rate individuals are willing to pay, *i.e.* $\partial \tau_j / \partial \mu_j < (>) 0$. Item (iii) points out that, whatever the private and public goods are

[4] Concavity of $u(.)$ ensures that (3) is necessary and sufficient to reach a maximum.

complements or substitutes, the overall effect of an increase in the average income of a jurisdiction is to raise the provision of public good.

Concavity of individual utility function allows us to apply the median voter theorem. Hence, there always exists a local tax rate chosen by a majority of voters which is the one most preferred by the median voter of jurisdiction j, denoted by y_j^m. The equilibrium tax rate for a jurisdiction j satisfies (3) with $y = y_j^m$. It is denoted $\tau^*(y_j^m)$, while $q_j^*(y_j^m)$ denotes the quality of the local public good.

Stage 2. We turn to the choice of individuals between communities that we shall refer to stage 2. The analysis borrows a lot from communities' formation frameworks (like Westhoff, 1977, Epple, Filimon and Romer, 1993, and Fernandez and Rogerson, 1996). Any individual y faces the following program:

$$\max_j u\left((1-\tau^*(y_j^m))y\right) + \beta u(q_j^*(y_j^m)).$$

Given Assumption 2a (2b), we can rank the slopes of the indifference curves with respect to income and define a threshold income $\widetilde{y}_j$ such that for two communities j and $j+1$ with $\tau_{j+1} > \tau_j$ and $q_{j+1} > q_j$ we have

$$u\left((1-\tau_{j+1})\widetilde{y}_j\right) + \beta u(q_{j+1}) = u\left((1-\tau_j)\widetilde{y}_j\right) + \beta u(q_j) \tag{4}$$

$$u\left((1-\tau_{j+1})y\right) + \beta u(q_{j+1}) \geq (\leq) u\left((1-\tau_j)y\right) + \beta u(q_j) \text{ for all } y \geq \widetilde{y}_j \tag{5}$$

$$u\left((1-\tau_{j+1})y\right) + \beta u(q_{j+1}) < (>) u\left((1-\tau_j)y\right) + \beta u(q_j) \text{ for all } y < \widetilde{y}_j \tag{6}$$

Under Assumption 2a (2b), all individuals richer (poorer), respectively poorer (richer), than $\widetilde{y}_j$ strictly prefer to live in the jurisdiction which levies the higher, respectively lower, level of tax rate and provides the higher, respectively lower, level of public good. In the following proposition, we characterize the SPNE of the subgame defined by stages 3 and 2.[5]

Proposition 1 *Under either Assumption 2a (2b), a free mobility equilibrium is characterized as follows:*

(i) There exists a vector $(\widetilde{y}_j)_{j=1,...,J-1}$ *such that* $u\left((1-\tau_{j+1})\widetilde{y}_j\right) + \beta u(q_{j+1}) = u\left((1-\tau_j)\widetilde{y}_j\right) + \beta u(q_j)$ *for* $j = 1, ..., J-1$.

(ii) Jurisdictions are vertically differentiated: under Assumption 2a (2b) $(\tau_J^*, q_J^*) >> (<<)(\tau_{J-1}^*, q_{J-1}^*) >> (<<)... >> (<<)(\tau_1^*, q_1^*)$.

[5] To keep notations simple, τ_j^*, respectively q_j^*, denotes $\tau^*(y_j^m)$, respectively $q^*(y_j^m)$.

(iii) Under Assumption 2a (2b), $\widetilde{y}_{j+1} > \widetilde{y}_j$ for all $j = 1, ...J-1$ and all individuals with $\widetilde{y}_{j+1} > y > \widetilde{y}_j$ live in community $j+1$.

(iv) The equilibrium set of jurisdictions represent a partition of the income support into $J-1$ intervals.

This proposition is well known in the literature (see, among many others, Epple and Romer, 1991, Fernandez and Rogerson, 1996, Hansen and Kessler, 2001, Gravel and Thoron, 2007, Calabrese *et al.*, 2012)[6]. The free mobility equilibrium is such that jurisdictions providing a higher level of public good are also taxing more heavily their inhabitants. Item (ii) is a necessary condition to have individuals $\widetilde{y}_j$ indifferent between two adjacent jurisdictions (item (i)). Finally, the free mobility equilibrium is income-stratified[7]. In other words, people with similar income levels choose to live in the same jurisdiction. When the private and public goods are complements (substitutes), richer individuals reside in jurisdictions with the highest (lowest) tax and public good package (see Hansen and Kessler, 2001).

At this stage, we do not provide any conditions guaranteeing uniqueness of the free mobility equilibrium. In particular, we have to find conditions such that the identity of an individual satisfying (4) for two communities j and $j+1$ is uniquely defined. We denote by $U_j^*(y) \equiv u\left((1-\tau_j^*)y\right) + \beta u(q_j^*)$ the best location for an individual θ earning income y and living in jurisdiction j.

Stage 1. Let us now turn to the choice of effort. This stage amounts to endogenize the income distribution. The program any individual θ faces is the following:

$$\max_y [\max_j U_j^*(y) - v\left(\frac{y}{\theta}\right)] \tag{7}$$

This program thus exhibits the fact that while deciding to exert an effort level an individual must take into account the consequences of his choice on the jurisdiction he will live in. Let $y^*(\theta) \equiv \arg\max_y [\max_j U_j^*(y) - v\left(\frac{y}{\theta}\right)]$.

In order to obtain the equilibrium, we define for each θ and any $j = 1, ..., J$, the income level $\widehat{y}_j(\theta)$ that solves

$$(1-\tau_j^*)u'\left((1-\tau_j^*)y\right) = \frac{1}{\theta} v'\left(\frac{y}{\theta}\right). \tag{8}$$

Let us now provide some information about $\widehat{y}_j(\theta)$. Some comparative statics lead to

[6] Except Hansen and Kessler (2001), most of the literature develops multijurisdiction models under Assumption 2a only.

[7] See Hansen and Kessler (2001) who discuss existence and non-existence of free mobility equilibria. Notice that they consider the case where the public good production is not characterized by any economies of scale.

Lemma 2 *Under Assumption 2a (2b), (i)* $\partial\widehat{y}_j(\theta)/\partial\theta > 0$*, (ii)* $\partial\widehat{y}_j(\theta)/\partial\tau_j^* > (<)0$*.*

Proof. See Appendix. ■

Lemma 2 gives useful insights about individual effort given the jurisdiction's choice. The more productive is the individual, other things being equal, the higher is his effort and consequently his income. Under Assumption 2a (2b), the higher is the tax rate levied in jurisdiction j, other things being equal, the higher (lower) is income of individual θ exhibiting the fact that the income effect exceeds (is outweighed by) the substitution effect. Item (ii) thus stresses the fact that jurisdiction membership impacts the choice of effort.

Corollary 1 *Under Assumption 2a (2b) we have* $\tau_{j+1}^* > (<)\tau_j^*$*, then* $\widehat{y}_{j+1}(\theta) > \widehat{y}_j(\theta)$ *for any* $\theta \in [\underline{\theta}, \overline{\theta}]$ *and any* $j = 1, ..., J-1$*.*

Corollary 1 highlights the fact that for any individual θ her income jumps according to jurisdiction membership. This corollary will be key for characterizing the equilibrium income distribution.

We define $V_j(\theta)$ the indirect utility level of individual θ living in jurisdiction j with income $\widehat{y}_j$

$$V_j(\theta) \equiv u\left((1-\tau_j^*)\widehat{y}_j(\theta)\right) + \beta u(q_j^*) - v\left(\frac{\widehat{y}_j(\theta)}{\theta}\right).$$

In order to solve (7), we need to rank preferences over tax rates and local public goods according to θ rather than y as in Proposition 1.

Lemma 3 *For any* $\tau_j, \tau_{j'}, q_j, q_{j'}$ *such that* $\tau_j < \tau_{j'}$ *and* $q_j < q_{j'}$*, under Assumption 2a (2b) if for an individual* θ' *we have* $V_j(\theta') \leq V_{j'}(\theta')$*, respectively* $V_j(\theta') \geq V_{j'}(\theta')$*, then for any* $\theta > (<)\theta'$*, respectively* $\theta < (>)\theta'$*, we have* $V_j(\theta) < V_{j'}(\theta)$*, respectively* $V_j(\theta) > V_{j'}(\theta)$*.*

Proof. See Appendix. ■

Lemma 3 amounts to say that preferences are intermediate (see Grandmont, 1978, or Demange, 1994). Hence, there is a "monotonicity" property in the ranking of jurisdictions according to type θ. When an individual has a particular ranking of two jurisdictions then either all less productive or more productive individuals agree upon this ranking. This Lemma is key to demonstrate the following proposition.

Proposition 2 *If there exists an equilibrium with* J *jurisdictions and endogenous income distribution then it is productivity-stratified, i.e. each jurisdiction is formed from a single productivity interval and the equilibrium set of jurisdictions is a partition of* $[\underline{\theta}, \overline{\theta}]$ *into* J *intervals.*

Proof. See Appendix. ■

Proposition 2 states that individuals with similar productivity parameter gather in the same jurisdictions. According to Proposition 2, at equilibrium any jurisdiction j is comprised by all individuals θ in the interval $[\widetilde{\theta}_{j-1}, \widetilde{\theta}_j]$ where $\widetilde{\theta}_{j-1}$ and $\widetilde{\theta}_j$ are the boundaries of jurisdiction j and are such that $1 = \sum_{j=1}^{J} \int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} dH(\theta)$ with $\widetilde{\theta}_0 = \underline{\theta}$ and $\widetilde{\theta}_J = \overline{\theta}$. We denote by

$$V_j^*(\theta) \equiv u\left((1-\tau_j^*)y^*(\theta)\right) + \beta u(q_j^*) - v\left(\frac{y^*(\theta)}{\theta}\right) = \max_{\widehat{y}_j(\theta)} V_j(\theta)$$

the indirect utility level obtained by an agent θ who lives in jurisdiction j at equilibrium. We have

Proposition 3 *If an equilibrium with J jurisdictions and endogenous income distribution exists, then it is characterized as follows:*

(i) There exists a vector $(\widetilde{\theta}_j)_{j=1,\dots,J-1}$ such that $V_j^*(\widetilde{\theta}_j) - V_{j+1}^*(\widetilde{\theta}_j) = 0$, for $j = 1, ..., J-1$;

(ii) Jurisdictions are vertically differentiated: under Assumption 2a (2b) $(\tau_J^*, q_J^*) >> (<<)(\tau_{J-1}^*, q_{J-1}^*) >> (<<) ... >> (<<)(\tau_1^*, q_1^*)$;

(iii) For any $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$, and any $j = 1, ..., J-1$, $\widetilde{\theta}_0 = \underline{\theta}$ and $\widetilde{\theta}_J = \overline{\theta}$, the equilibrium level of effort is $y^*(\theta) = \widehat{y}_{j+1}(\theta)$ and $\widetilde{y}_{j+1} > y^*(\theta) > \widetilde{y}_j$.

Proof. See Appendix. ■

Items (i) and (ii) are similar to the first two items of Proposition 1 considering now the productivity distribution. Given the continuity of $V_j^*(.)$ with respect to θ, item (i) states that there exist boundary individuals $\widetilde{\theta}_j$ who are indifferent between two adjacent jurisdictions j and $j+1$. According to Lemma 2, we know that more productive individuals earn higher incomes. Hence, item (ii) amounts to say that depending on whether private and public goods are complements, respectively substitutes, the more productive individuals live in jurisdictions with higher, respectively lower, bundles (q, τ). Finally, item (iii) characterizes the level of income and jurisdiction chosen at equilibrium by any individual θ. Precisely, for any individual $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$, his best effort leads to $y^*(\theta) = \widehat{y}_{j+1}(\theta)$ given by (8). Further, for any $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$, $\widehat{y}_{j+1}(\theta)$ belongs to the interval $]\widetilde{y}_j, \widetilde{y}_{j+1}[$ and thus their best jurisdiction choice at stage 2 is jurisdiction $j+1$. Hence, at equilibrium, a community j can be characterized by the interval $[\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$. In the Appendix, we show that under Assumption 2a, our model satisfies all the assumptions made by Westhoff (1977) to prove existence of an equilibrium.

To capture the intuition behind item (iii), let us consider a sketch of the proof as depicted in Figure 1. The bold curve is the upper envelope of utilities $U^*_{j+1}(y)$ and $U^*_j(y)$. It is shown in the Appendix 7.7 that this curve is characterized by a kink at $\widetilde{y}_j$. Its existence relies on Assumption 2a or 2b which implies the sorting condition. Graphically, $\widehat{y}_j(\theta)$ is the largest distance between $U^*_j(y)$ and $v(.)$ and $y^*(\theta)$ maximizes the distance between the upper envelope and the disutility $v(.)$. It is important to notice that due to the kink characterizing the upper envelope at $\widetilde{y}_j$, any individual θ must consider several income levels $\widehat{y}_j(\theta)$, each one satisfying (8) for a particular j. In the graph, we consider an individual $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$ whose highest level of utility is, according to Lemma 3, $V^*_{j+1}(\theta)$. Assume by contradiction that $y^*(\theta) = \widehat{y}_{j+1}(\theta) < \widetilde{y}_j$ for an individual θ, $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$. As $y^*(\theta) < \widetilde{y}_j$, the individual θ prefers to live in jurisdiction j rather than jurisdiction $j+1$, *i.e.* $V^*_{j+1}(\theta) < U^*_j(y^*(\theta)) - v(y^*(\theta)/\theta)$ with $U^*_j(y^*(\theta)) \equiv u((1-\tau^*_j)y^*(\theta)) + \beta u(q^*_j)$. By definition of $\widehat{y}_j(\theta)$, the individual θ in jurisdiction j would prefer earning the income $\widehat{y}_j(\theta)$ (which is assumed to be higher than $\widetilde{y}_{j-1}$) rather than $y^*(\theta)$. This can be clearly seen in the graph as $U^*_j(y^*(\theta)) - v(y^*(\theta)/\theta) < V_j(\theta)$. Hence, we have $V_j(\theta) > V^*_{j+1}(\theta)$, leading to a contradiction.

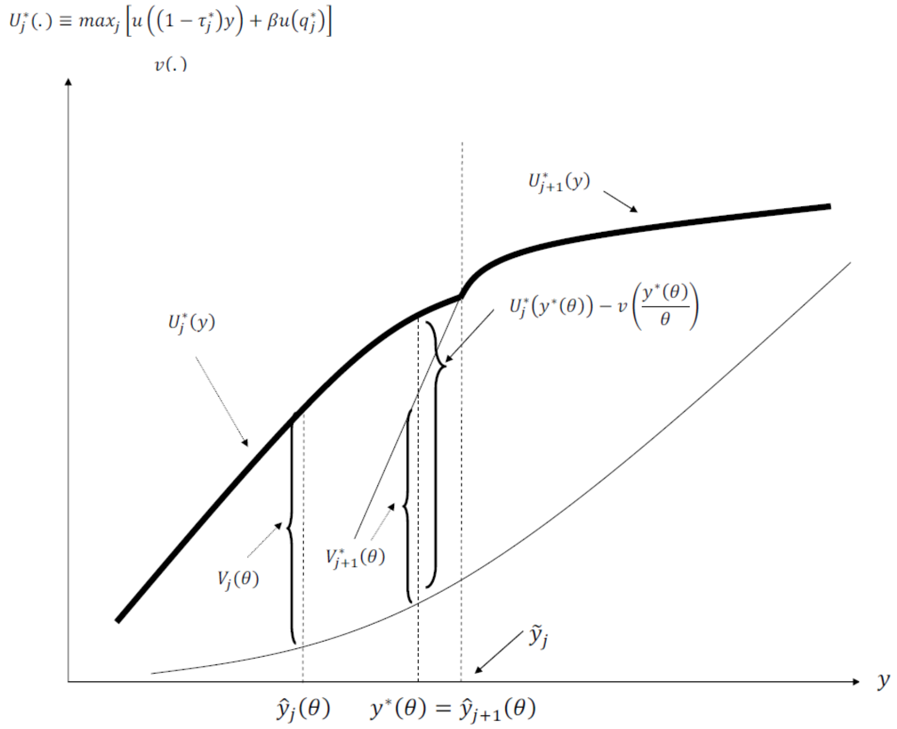

Figure 1: Item (iii) of Theorem 2.

The crucial feature of our framework is that income distribution that gives rise to the jurisdiction formation is endogenous. According to Proposition 3, the income distribution can be described by a set of income intervals, each one being a set of individuals residing in the same jurisdiction. Further, the income distribution exhibits discontinuities, each one corresponding to a change in jurisdiction membership. Such discontinuities highlight that effort and thus income is jurisdiction-dependent. It turns out that individuals with θ in the neighborhood of a jurisdiction boundary $\widetilde{\theta}_j$ may dramatically differ with respect to their effort, and consequently by their level of income, according to the jurisdiction they live in (see Lemma 2 and Corollary 2). Precisely, consider two individuals θ and θ' who are close in the productivity distribution but $\theta > \widetilde{\theta}_j$ and $\theta' < \widetilde{\theta}_j$. Individual θ resides in jurisidiction $j+1$ while θ' is in j. It turns out that their optimal level of effort, respectively given yielding $y^*(\theta) = \widehat{y}_{j+1}(\theta)$ and $y^*(\theta') = y_j^*(\theta)$, can differ substantially depending on values taken by τ_{j+1}^* and τ_j^*.

Figure 2 illustrates the stratified equilibrium in the $(\theta - y)$-space whatever Assumption 2a or 2b holds. The bold curve depicts the equilibrium income distribution. We see that for any θ such that $\widetilde{\theta}_{j-1} < \theta < \widetilde{\theta}_j$, we have $y^*(\theta) = \widehat{y}_j(\theta)$ and $\widetilde{y}_{j-1} < \widehat{y}_j(\theta) < \widetilde{y}_j$ implying that her best jurisdiction choice is j (see item (iii) of Proposition 3). For higher values of θ, *i.e.* $\theta > \widetilde{\theta}_j$, individuals exert an effort level that allows them to earn the income $\widehat{y}_{j+1}(\theta)$. We have $\widehat{y}_{j+1}(\theta) > \widetilde{y}_j$ and thus all individuals with $\theta > \widetilde{\theta}_j$ strictly prefer jurisdiction $j+1$ to jurisdiction j. As it is depicted in Figure 2, some levels of income will never arise. For instance, we see that no income in $\left[\widehat{y}_{j-1}\left(\widetilde{\theta}_{j-1}\right), \widehat{y}_j\left(\widetilde{\theta}_{j-1}\right)\right]$ and $\left[\widehat{y}_j\left(\widetilde{\theta}_j\right), \widehat{y}_{j+1}\left(\widetilde{\theta}_j\right)\right]$ will be observed at equilibrium. Moreover, item (iii) of Proposition 3 implies that $\widehat{y}_{j-1}\left(\widetilde{\theta}_{j-1}\right) < \widetilde{y}_{j-1} < \widehat{y}_j\left(\widetilde{\theta}_{j-1}\right)$ and $\widehat{y}_j\left(\widetilde{\theta}_j\right) < \widetilde{y}_{j-1} < \widehat{y}_{j+1}\left(\widetilde{\theta}_j\right)$. This result departs from most multijurisdiction models as boundary incomes satisfying (4) do not exist at equilibrium[8]. Jurisdictions are now characterized by boundary levels of productivity $(\widetilde{\theta}_j)$. Consequently, at Stage 2 of our game, a boundary individual $\widetilde{\theta}_j$ is no more indifferent between the two adjacent jurisdictions j and $j+1$. Our model of jurisdictions formation with endogenous income distribution may provide an explanation why, *once effort choices are made*, individuals may oppose changes in the jurisdictions' frontiers as they would imply large welfare effects[9].

[8]Empirical income distributions are not characterized by such discontinuities. In Section 5, we develop a stochastic version of this model and obtain a continuous equilibrium income distribution. Still, this stochastic version obtains that individual income is jurisdiction-dependent.

[9]See Epple and Romer (1989) for empirical findings on the scarcity of jurisdictional boundaries changes and De Bartolome and Ross (2007) for a theoretical explanation of boundary fixedness.

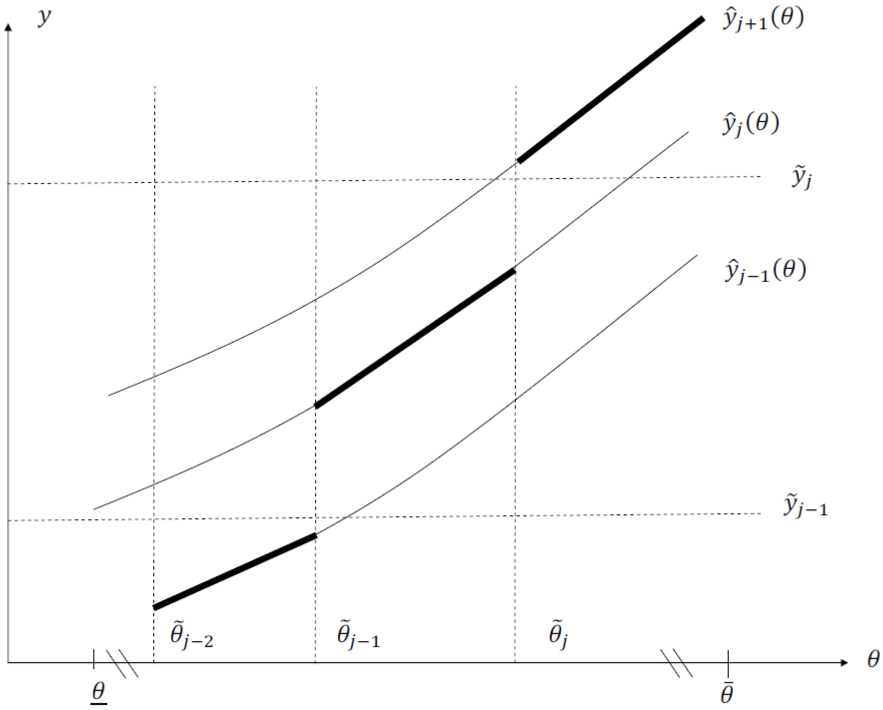

Figure 2: Equilibrium Income Distribution.

4 Optimal Taxation at the Central Level

This section addresses the issue of the optimal resource allocations in this economy. In a first step, we characterize the stratified-constrained optimum. We consider that the central government is constrained by the free mobility of individuals which will oblige him to build jurisdictions that are formed by a single interval of individuals. In a second step, our aim is to characterize the central tax structure that implements such efficient allocations.

4.1 Stratified-Constrained Optimum

We denote by $\omega(\theta)$ the weight given to individual θ in the social welfare function which can be written as follows:

$$\sum_{j=1}^{J} \int_{\tilde{\theta}_{j-1}}^{\tilde{\theta}_j} \omega(\theta)\mathcal{U}_j(\theta)h(\theta)d\theta.$$

The government has all latitude to choose private consumption levels, local public goods, effort levels and boundaries of jurisdictions. The government is constrained by resources scarcity and the freedom of individuals to move across jurisdictions. Hence, the government's program can be written as follows:

$$\max_{\left\{\left\{\{c_j(\theta),y_j(\theta)\}_{\widetilde{\theta}_{j-1}\leq\theta\leq\widetilde{\theta}_j},q_j\right\}_{j=1,...,J} \text{ and } \left\{\widetilde{\theta}_j\right\}_{j=1,...,J-1}\right\}} \sum_{j=1}^{J}\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\omega(\theta)\left[u(c_j(\theta))+\beta u(q_j)-v\left(\frac{y_j(\theta)}{\theta}\right)\right]h(\theta)d\theta$$

with respect to the following resource constraint

$$\sum_{j=1}^{J}\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} y_j(\theta)h(\theta)d\theta \geq \sum_{j=1}^{J}\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} c_j(\theta)h(\theta)d\theta + \sum_{j=1}^{J}\zeta(n_j)q_j \text{ with } \widetilde{\theta}_0=\underline{\theta} \text{ and } \widetilde{\theta}_J=\overline{\theta}, \tag{9}$$

and with respect to the "free mobility" constraints

$$u(c_j(\theta))+\beta u(q_j)-v\left(\frac{y_j(\theta)}{\theta}\right)\geq u(c_k(\theta))+\beta u(q_k)-v\left(\frac{y_k(\theta)}{\theta}\right) \tag{10}$$

$$\text{for any } j,k=1,...,J,\ k\neq j \text{ and any } \theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j].$$

According to (10), at the optimum, no individual in jurisdiction j must have any incentives to move into another jurisdiction. As our aim is to know whether an optimum can be supported as an equilibrium, we impose the planner's program to satisfy these free mobility conditions. All the computations of this optimization problem are relegated in the Appendix. The characterization of this stratified-constrained optimum allocation is provided in the following proposition.

Proposition 4 *The stratified constrained optimum allocation is characterized by the following equations:*

(i) Optimal consumption and effort:

$$\begin{aligned}
\omega(\theta)u'(c_j^{op}(\theta)) &= \lambda \\
\omega(\theta)\frac{1}{\theta}v'\left(\frac{y_j^{op}(\theta)}{\theta}\right) &= \lambda \\
\frac{v'\left(\frac{y_j^{op}(\theta)}{\theta}\right)}{u'(c_j^{op}(\theta))} &= \theta,
\end{aligned} \tag{11}$$

$$\forall\theta\in[\widetilde{\theta}_{j-1}^{op},\widetilde{\theta}_j^{op}],\ j=1,...,J \textit{ with } \widetilde{\theta}_0^{op}=\underline{\theta} \textit{ and } \widetilde{\theta}_J^{op}=\overline{\theta}$$

with λ the Lagrange multiplier associated with the resource constraint.

(ii) Optimal provision of local public good:

$$\beta u'(q_j^{op}) \int_{\widetilde{\theta}_{j-1}^{op}}^{\widetilde{\theta}_j^{op}} \frac{h(\theta)}{u'(c_j^{op}(\theta))} d\theta = \zeta(n_j^{op}), \ \forall j = 1, ..., J \ \textit{with} \ \widetilde{\theta}_0^{op} = \underline{\theta} \ \textit{and} \ \widetilde{\theta}_J^{op} = \overline{\theta}. \tag{12}$$

(iii) Optimal size:

$$\begin{aligned} & y_j^{op}(\widetilde{\theta}_j^{op}) - c_j^{op}(\widetilde{\theta}_j^{op}) - q_j^{op}\zeta'(n_j^{op}) \\ & = y_{j+1}^{op}(\widetilde{\theta}_j^{op}) - c_{j+1}^{op}(\widetilde{\theta}_j^{op}) - q_{j+1}^{op}\zeta'(n_{j+1}^{op}) \ \forall j = 1, ..., J \ \textit{with} \ \widetilde{\theta}_0^{op} = \underline{\theta} \ \textit{and} \ \widetilde{\theta}_J^{op} = \overline{\theta}. \end{aligned} \tag{13}$$

(iv) Boundary indifference:

$$u(c_j(\widetilde{\theta}_j^{op})) + \beta u(q_j) - v\left(\frac{y_j(\widetilde{\theta}_j^{op})}{\widetilde{\theta}_j^{op}}\right) = u(c_{j+1}(\widetilde{\theta}_j^{op})) + \beta u(q_{j+1}) - v\left(\frac{y_{j+1}(\widetilde{\theta}_j^{op})}{\widetilde{\theta}_j^{op}}\right) \ \textit{for} \ j = 1, .., J-1. \tag{14}$$

Equation (11) characterizes the optimal level of effort which is such that the marginal rate of substitution between consumption and cost of effort equals individual productivity. According to equation (12), each jurisdiction provides an efficient level of public good that satisfies the Bowen-Lindhal-Samuelson condition such that the sum of the marginal rates of substitution between private good consumption and public good consumtion equals the marginal cost of public good production, *i.e.* $\zeta(n_j^{op})$. Equation (13) characterizes the optimal size. At the optimum, individual $\widetilde{\theta}_j^{op}$ is such that his migration from j to $j+1$ generates a variation of the resources in jurisdiction j that exactly equals the variation of resources in jurisdiction $j+1$. Hence, individual $\widetilde{\theta}_j^{op}$'s migration does not generate any extra resources that could make any individual be better off. Equation (13) combined with free mobility constraints lead the externalities generated by free mobility of individuals to be fully internalized at the optimum. Finally, any optimal jurisdiction j can be characterized by the interval $[\widetilde{\theta}_j^{op}, \widetilde{\theta}_{j+1}^{op}]$ where any boundary individual $\widetilde{\theta}_j^{op}$ is indifferent between both adjacent communities j and $j+1$.

4.2 Properties of the Optimal Fiscal Scheme

We now tackle the issue whether it is possible to design a tax structure such that equilibrium allocations of resources coincide with optimal ones or if a second-best approach must be envisaged. We denote by $T(y, \tau_j, \mu_j, n_j, I)$ taxes used by the central government. The government can design taxes with respect to income y, any variables referring to jurisdiction membership, *i.e.* τ_j, μ_j and n_j, and total income I. However, θ and l being not observable, taxes cannot depend on individual productivity and effort. We denote by $T_i(.)$ the partial derivative of $T(.)$ with respect to its ith argument. Individual consumption thus can be rewritten as follows

$$c(y) = (1-\tau_j)y - T(y, \tau_j, \mu_j, n_j, I). \tag{15}$$

The local public good is still produced according to (2). The equilibrium with taxes denoted by $\left\{ \left\{ c_j^{**}(\theta), y_j^{**}(\theta), q_j^{**}, \tau_j^{**} \right\}_{\theta \in [\widetilde{\theta}_{j-1}^{**}, \widetilde{\theta}_j^{**}], j=1,...,J}, \left\{ \widetilde{\theta}_j^{**} \right\}_{j=1,...,J-1} \right\}$ is characterized by (15) and the following four equations

(i) For any $j = 1, ..., J$, any θ such that $\widetilde{\theta}_{j-1}^{**} \leq \theta \leq \widetilde{\theta}_j^{**}$ with $\widetilde{\theta}_0^{**} = \underline{\theta}$ and $\widetilde{\theta}_J^{**} = \overline{\theta}$, we have

$$\left(1 - \tau_j - T_1(y_j^{**}(\theta), \tau_j^{**}, \mu_j^{**}, n_j^{**}, I^{**})\right) u'(c_j^{**}(\theta)) = \frac{1}{\theta} v' \left(\frac{y_j^{**}(\theta)}{\theta} \right). \tag{16}$$

(ii) Tax rate chosen by the median voter of jurisdiction j,

$$\left[y_j^{**}(\theta^m) + T_2(y_j^{**}(\theta), \tau_j^{**}, \mu_j^{**}, n_j^{**}, I^{**}) \right] u'(c_j^{**}(\theta^m)) = \frac{n_j^{**} \mu_j^{**}}{\zeta(n_j^{**})} \beta u'(q_j^{**}) \tag{17}$$

with $y_j^{**}(\theta^m)$ the median income of jurisdiction j.[10]

(iii) For any $j = 1, ..., J-1$, we have $\widetilde{\theta}_j^{**}$ such that

$$u(c_{j+1}^{**}(\widetilde{\theta}_j^{**})) + \beta u(q_{j+1}^{**}) - v \left(\frac{y_{j+1}^{**}(\widetilde{\theta}_j^{**})}{\widetilde{\theta}_j^{**}} \right) - \left[u(c_j^{**}(\widetilde{\theta}_j^{**})) + \beta u(q_j^{**}) - v \left(\frac{y_j^{**}(\widetilde{\theta}_j^{**})}{\widetilde{\theta}_j^{**}} \right) \right] = 0. \tag{18}$$

(iv) Budget constraint,

$$\sum_{j=1}^{J} \int_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} T(y_j^{**}(\theta), \tau_j^m, \mu_j^{**}, n_j^{**}, I^{**}) = 0. \tag{19}$$

Further, we know that there exist some weights such that the fiscal scheme $T(y, \tau_j, \mu_j, n_j, I)$ independent of any θ and l is a solution of the central government's program (see also Calabrese, Epple and Romano, 2012). The following proposition provides a characterization of the tax structure that allows to support an optimum.

Proposition 5 *In a free mobility and efficient equilibrium, for any j, taxes are characterized as follows:*

[10] We can apply the median voter theorem if preferences are single-peaked with respect to τ, so that:

$$\left[-y - T_2(y(\theta), \tau_j, \mu_j, n_j, I)\right]^2 u''(c_j(\theta)) - T_{22}(y(\theta), \tau_j, \mu_j, n_j, I) u'(c_j(\theta)) + u''(q_j) \left[\frac{\int_{\theta_{j-1}}^{\theta_j} y(\theta) h(\theta) d\theta}{\zeta(n_j)} \right]^2 \leq 0.$$

(i) $T(y, \tau_j, \mu_j, n_j, I)$ *is specified as follows:* $T(y, \tau_j, \mu_j, n_j, I) = -\tau_j y + \varphi(\tau_j, \mu_j, n_j, I)$;

(ii) Bowen-Lindhal-Samuelson condition:

$$\varphi_1(\tau_j^m, \mu_j^{**}, n_j^{**}, I^{**}) = \mu_j^{**} \frac{\left(u'(c_j^{**}(\theta^m))\right)^{-1}}{\left(\frac{1}{n_j^{**}} \int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} \left(u'(c_j^{**}(\theta))\right)^{-1} h(\theta) d\theta\right)},$$

(iii) Optimal Size:

$$\begin{aligned} &\varphi(\tau_j^{**}, \mu_j^{**}, n_j^{**}, I^{**}) - \varphi(\tau_{j+1}^{**}, \mu_{j+1}^{**}, n_{j+1}^{**}, I^{**}) \\ &= q_j^{**}\zeta'(n_j^{**}) - q_{j+1}^{**}\zeta'(n_{j+1}^{**}) \ \forall j = 1, ..., J-1, \end{aligned}$$

(iv) National Budget Constraint:

$$\sum_{j=1}^{J} \int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} \left(-\tau_j^{**} y_j^{**}(\theta) + \varphi(\tau_j^{**}, \mu_j^{**}, n_j^{**}, I^{**})\right) h(\theta) d\theta = 0,$$

(v) Local Public Good Provision:

$$\tau_j^m \int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} y_j^{**}(\theta) h(\theta) d\theta = q_j^{**}\zeta(n_j^{**}) \ \textit{for any } j = 1, ..., J,$$

(vi) Optimal Effort:

$$u'(y_j^{**}(\theta) - \varphi(\tau_j^{**}, \mu_j^{**}, n_j^{**}, I^{**})) = \frac{1}{\theta} v' \left(\frac{y_j^{**}(\theta)}{\theta}\right).$$

Sketch of the Proof. The proof proceeds in designing $T(.)$ so that equations (11)-(13) and (9) characterizing the optimum coincide with the system of equations (16)-(19) of the equilibrium with taxes. Let us mention that in the Appendix we derive a necessary and sufficient condition which may be more restrictive than Assumptions 2a and 2b for the sorting condition to hold when this tax structure is implemented, *i.e.* $S_T(q_j, \tau_j, y)$ being the slope of individual preferences in the policy space when $T(.)$ is implemented $\frac{\partial S_T(q_j, \tau_j, y)}{\partial y} \gtreqless 0 \Leftrightarrow \frac{c_j}{y} \lesseqgtr -\frac{c_j u''(c_j)}{u'(c_j)}$. Besides, the intermediate preferences property is always satisfied. ■

In spite of the asymmetry of information on agents' productivity level, we find a central tax scheme that allows the central planner to implement a first best allocation, *i.e.* an allocation that belongs to the contract curve. The optimal central tax structure, $T(y, \tau_j, \mu_j, n_j, I)$, is designed as the sum of two components: a grant depending on individual income, *i.e.* $\tau_j y$, and a lump-sum transfer depending on both jurisdiction's characteristics (size, average income and local taxation) and total income, *i.e.* $\varphi(\tau_j, \mu_j, n_j, I)$. This lump-sum transfer applied on jurisdiction membership works in a similar way than optimal taxation schemes with tagging (see Cremer et al., 2011). It taxes or subsidizes jurisdiction membership which is observable and due to the productivity-stratification properties of the equilibrium it taxes or subsidies somehow the individual productivity. The key feature of our tagging is that jurisdiction membership of an individual θ does provide information on a productivity interval to which her θ belongs to. Moreover, within jurisdictions, the grant is proportional to the individuals' income such that incentives constraints are automatically satisfied.

With regard to item (i), the optimal national tax structure is designed so that it must exactly offset the distortion that the local taxation may generate on the effort choice (see item (vi)). The whole tax structure, *i.e.* national plus local taxes, amounts to lump-sum transfer given jurisdiction j's membership as $c(y) = y - \varphi(\tau_j, \mu_j, n_j, I)$. However, the equilibrium being productivity-stratified, the tax structure is non-linear as $\varphi(.)$ may move upward or downward with respect to jurisdiction's membership. Precisely, when the slope of an indifference curve through any point of policy space (q, τ) increases (decreases), we have for $q_j^{**} < (>) q_{j+1}^{**}$ and $\varphi(\tau_{j+1}^m, \mu_{j+1}^{**}, n_{j+1}^{**}, I^{**}) > (<) \varphi(\tau_j^m, \mu_j^{**}, n_j^{**}, I^{**})$ for any $j = 1, ..., J-1$. Hence, people who live in jurisdictions providing the higher level of local public goods pay higher lump sums.

According to item (ii), the tax structure is such that distortions generated by the median voter rule are corrected. The function $\varphi(.)$ is designed so that (17) and (12) coincide. It thus turns out that the median voter's marginal rate of substitution of public good for private good, *i.e.* $\left[\beta u'(q_j^{**}) \int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} y_j^{**}(\theta) h(\theta) d\theta \right] / \left[\varphi_1(\tau_j^m, \mu_j^{**}, n_j^{**}, I^{**}) u'(c_j^{**}(\theta^m))\right]$, equals the the total willingness to pay for the public good over the whole jurisdiction's population, *i.e.* $\beta u'(q_j^{op}) \int\limits_{\widetilde{\theta}_{j-1}^{op}}^{\widetilde{\theta}_j^{op}} \frac{h(\theta)}{u'(c^{op}(\theta))} d\theta$.

According to item (iii), lump-sum transfers are designed so that the jurisdictions 'sizes are optimal. Considering two adjacent jurisdictions j and $j+1$, lump-sum transfers must be such that the resources constraint is not modified by the migration of the marginal individual from jurisdiction j to $j+1$. The variation of transfers applied on the migrating individual, $\varphi(\tau_{j+1}^m, \mu_{j+1}^{**}, n_{j+1}^{**}, I^{**}) - \varphi(\tau_j^m, \mu_j^{**}, n_j^{**}, I^{**})$, is then equal to the variation of marginal cost of public good production.

Finally, the optimal tax structure must be such that the resource constraint as well as the local budget constraint are binding (see items (iv) and (v)).

We can notice that the following specific form of national tax structure

$$T(y, \tau_j, \mu_j, n_j, I) = -\tau_j y + b_j q_j + d_j + z$$
$$\text{with}$$
$$b_j = \frac{\zeta(n_j^{**}) \left(u'(c_j^{**}(\theta^m))\right)^{-1}}{\left(\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} \left(u'(c_j^{**}(\theta))\right)^{-1} h(\theta) d\theta\right)},$$
$$d_j - d_{j+1} = q_j^{**}\left(\zeta'(n_j^{**}) - b_j\right) - q_{j+1}^{**}\left(\zeta'(n_{j+1}^{**}) - b_{j+1}\right),$$
$$z \in \mathbb{R}$$

can satisfy the entire set of items of Proposition 5. In particular, for $\zeta'(.) > 0$, it can easily be checked that $b_j = \zeta'(n_j^{**})$, $d_j = 0$ and $z = 0$ would satisfy Proposition 5. Let us stress that the following equation would be obtained

$$\zeta'(n_j^{**}) = \frac{\zeta(n_j^{**})}{n_j^{**}} \frac{\left(u'(c_j^{**}(\theta^m))\right)^{-1}}{\mathbb{E}_j\left[\left(u'(c_j^{**}(\theta))\right)^{-1}\right]}. \tag{20}$$

with $\mathbb{E}_j\left[\left(u'(c_j^{**}(\theta))\right)^{-1}\right]$ denoting the average of jurisdiction j inverse marginal utility of private consumption. Equation (20) could be paralleled to the optimal-club-size-with-congestion-condition emphasized in the standard model of club formation with homogenous individuals stating that the marginal congestion cost equals the average congestion cost, i.e. $\zeta'(n_j^{**}) = \zeta(n_j^{**})/n_j^{**}$, leading the average cost of congestion to be at its minimum (see, for instance, Rubinfeld, 1987). Equation (20) takes into account individual consumption heterogeneity. Further, when $\zeta'(n_j^{**})$ equals 1, it turns out that $\varphi(\tau_j^m, \mu_j^{**}, n_j^{**}, I^{**}) = q_j$ for any j.[11]

5 Stochastic Model and Numerical Exercise

5.1 Model Specification

Our aim is to analyze the effect of decentralization on income distribution, aggregate income and welfare. We consider a stochastic version of the model. The gross income is now defined as follows

$$y(\theta) = \theta l \varepsilon$$

[11] This result is similar to Calabrese *et al.* (2012) where they assume publicly provided private good, or equivalently marginal congestion effects equal to 1 (see their Proposition 3).

with l denoting the labor supply and ε an income shock that agents observe once individual effort, the jurisdiction and the vote on local taxation are made. We assume that

$$\log \theta \rightsquigarrow \mathcal{N}\left(\widehat{\theta}, \varrho_\theta^2\right) \tag{21}$$

and

$$\log \varepsilon \rightsquigarrow \mathcal{N}\left(\psi, \varrho_\varepsilon^2\right) \text{ with } \mathbb{E}\left[\varepsilon\right] = 1. \tag{22}$$

We normalize the mass of individuals to one. Denoting by $\mathbb{E}[.]$ the expectation operator, we consider the following utility specification

$$\mathcal{U}_j(\theta) = \mathbb{E}\left[\frac{\left(c_j\left(\theta\right)\right)^{1-\sigma}}{1-\sigma} + \beta\frac{q^{1-\sigma}}{1-\sigma}\right] - \frac{\Phi}{\eta}\left(l_j\left(\theta\right)\right)^{\eta}, \text{ with } \Phi, \sigma > 0 \text{ and } \eta > 1,$$

and the public good of the jurisdiction j is equal to

$$q_j = \frac{\tau_j n_j \mu_j}{\zeta\left(n_j\right)} \text{ with } \zeta\left(n_j\right) \equiv n_j^{\nu}, \ \nu > 0. \tag{23}$$

In the Appendix, we develop in detail this model. In order to better understand the effects of decentralization, we are able to compare the level of inequality that arises at equilibrium between the centralized case, *i.e.* $J = 1$ and the decentralized case, *i.e.* $J > 1$. Denoting by $y^*\left(\theta, \varepsilon\right)$ the equilibrium level of income for individual θ under the shock ε and $var\left[\ln y^*\left(\theta, \varepsilon\right)\right]_J$ the variance of the lognormal income distribution at equilibrium when the number of jurisdictions is J, we are able to show the following

Proposition 6 *Whatever $\sigma > 0$, income inequality is higher under decentralization, that is*

$$var\left[\ln y^*\left(\theta, \varepsilon\right)\right]_{J>1} - var\left[\ln y^*\left(\theta, \varepsilon\right)\right]_{J=1} > 0, \text{ for any } J > 1.$$

The possibility given to individuals to segregate into local jurisdictions leads to a more unequal income distribution. The reason lies on the fact that effort is jurisdiction-dependent. Hence, decentralization which leads to different taxation/public good packages increases the range of effort levels yielding greater income heterogeneity.

However, we are not able to obtain analytical results on the effects of decentralization on welfare and effort. We thus proceed to numerical simulations of the model in order to further study the impact of an increase in decentralization, assimilated to an increase of J, on income inequality, labor supply and welfare.

To this aim, we set $\sigma = 2.5$ as we know from Westhoff (1977) that when $\sigma > 1$ an income-stratified equilibrium exists for any income distribution of income. Given that the elasticity of effort with respect to local taxation equals $\left(1-\sigma\right)/\left(\eta+\sigma-1\right)$ (see Appendix), we will suppose that $\eta = 1.15$ so that the elasticity is lower than 1. We assume no congestion effect, *i.e.* $\nu = 0$. Further, we calibrate the income distribution so that when there is only one jurisdiction it is the one considered by Calabrese, Epple and Romano (2012):

> "The distribution of MA income is calibrated using data from the 1999 American Housing Survey (AHS). Median income reported by the AHS is \$36,942. Using data for the 14 income classes reported by the AHS, we estimate mean household income to be \$54,710. These values and our assumption that the income distribution is lognormal imply $\ln y \rightsquigarrow \mathcal{N}(10.52, 0.785)$," (Calabrese, Epple and Romano, 2012).

To this aim, we set the mean income, denoted by μ^a, and the median income, denoted by μ^m, of the income distribution with only one community such that

$$\mu^a = 5.4710;\ \mu^m = 3.6942.$$

Given the properties of the lognormal distribution, mean income and the variance of logarithm of income for any individual θ thus equal

$$\mathbb{E}\left[\ln y\left(\theta, \varepsilon\right)\right] = \ln \mu^a = 1.6995,$$

$$var\left(\ln y\right) = 2 * \ln \frac{\mu^a}{\mu^m} = 0.7854.$$

In order to set ϱ_ε^2, we consider that ε can be interpreted as transitory income and use the ratio of transitory income to total earnings in U.S. calculated by Moffitt and Gottschalk (2001). It turns out that $\varrho_\varepsilon = 0.6267$. Further, if θ follows a lognormal distribution then the initial support considered must be such that $\underline{\theta} = 0$ and $\overline{\theta} = +\infty$. In our model $\underline{\theta} > 0$ and $\overline{\theta} < \infty$. As consequence, the $\theta-$distribution cannot follow a lognormal distribution. In order to use the lognormal distribution as an approximation of the real distribution of θ we will suppose that $\underline{\theta}$ is very close to zero and $\overline{\theta}$ is very high. In particular, we suppose that $\underline{\theta}$ and $\overline{\theta}$ are such that the 99.9% of θ of a lognormal distribution are inside the interval $\left(\underline{\theta}, \overline{\theta}\right)$

$$\underline{\theta} = H^{-1}(0.0005)$$

$$\overline{\theta} = H^{-1}(0.9995),$$

where $H^{-1}\left(x\right)$ is the inverse of the cumulative function of the lognormal distribution. Finally, we suppose that the local tax with only one community denoted by τ^{cen} is equal to 0.3. Parameters Φ and β are calibrated to obtain an average effort equal to 1, when there is only one community. The Appendix details how parameters are computed.

To sum up, we have

Parameters	η	σ	ν	Φ	β
	1.15	2.5	0	0.5493	0.3371
ε distribution		ψ		ϱ_{ε}	
		-0.1963		0.6267	
θ distribution		$\widehat{\theta}$		ϱ_{θ}	
		1.8372		1.444	
y distribution		$\mathbb{E}\left[\ln y\left(\theta,\varepsilon\right)\right]$		$var\left(\ln y\left(\theta,\varepsilon\right)\right)$	
		1.6995		0.7854	

5.2 Results

As our aim is to study the impact of fiscal decentralization on income inequality, labor supply and welfare, we compute the equilibrium that may arise for different values of the number of jurisdictions J from 1 to 5. When $J = 1$, the equilibrium is called centralized and we assimilate an increase of J as a rise of the degree of decentralization.

In the Appendix, we provide, for each J, jurisdictions' characteristics (size, boundary individuals, average income, local tax rate, local public good), the level of some aggregate variables (population average income, total welfare, total welfare in equivalent consumption units) and finally some inequality indicators (standard deviation of the lognormal distribution of income over the whole population, share of total inequality generated by the difference in the average income between communities (used in Calabrese, Epple and Romano, 2012). and standard deviation of the lognormal distribution of utility levels).

First, our simulations results presented in Tables A1-A3 and Figures A1-A2 describe the equilibria arising for each value of J. Given a number J of jurisdictions, we see from Tables A1-A3 that an equilibrium is productivity-stratified and that jurisdictions are ranked acccording to the bundle local taxation and local public good. As we assume that $\sigma > 1$, we satisty Assumption 2a and thus richer jurisdictions provide higher local public goods. Figure A1 depicts for each value of J the equilibrium level of individual labor supply with respect to the productivity parameter. Within each community, the labor supply decreases with productivity. Figure A1 also exhibits the jurisdiction-dependency of individual labor supply. The labor supply curve is characterized by discontinuities each one corresponding to a change in the jurisdiction membership. Further, when J varies, we can also see the impact of fiscal decentralization on stratification arising at equilibrium. Let us focus for instance on the equilibria when $J = 1$ and $J = 2$. When there are two communities, individuals whose $\theta \in [20.0833, 726.9035]$ belong to jurisdiction 2 (see Table A1). They vote for a higher level of taxation than the one chosen in the centralized equilibrium (see Table A2). However, from Table A3, fiscal decentralization leads them to consume less public good than in the centralized case, $q_2 = 1.5651$ when $J = 2$ while $q_1 = 1.6390$ when $J = 1$. The main reason why public good consumption drops for these individuals with respect to centralization is that the size of this jurisdiction, and thus the fiscal basis, is low ($n_2 = 0.2101$ in Table A1). However, let us mention some individuals may increase their local public good consumption when the degree of fiscal decentralisation rises. For instance, if we take individuals with $\theta \in [10.6136, 20.0833]$ who live in jurisidiction $j = 1$, respectively $j = 2$, when $J = 2$, respectively $J = 3$ (see Table A1) we see that they increase their consumption from $q_1 = 1.1483$ when $J = 1$ to $q_2 = 1.1532$ when $J = 2$ (see Tables A1 and A3). We deduce from Figure A1, that more fiscal decentralization incites individuals to increase their labor supply. Indeed, we see that any individual increases his labor supply when J increases. We conjecture that this can be explained by the fact that more jurisdictions provide greater opportunity for individuals to cluster with similar peers. Individuals thus exert higher effort in order to escape from the company of others who are lower in the productivity ladder.

Second, let us focus on the impact of an increase in J on aggregate variables provided in Table A4. We stress that fiscal decentralization may generate efficiency costs *via* its impact on labor supply. We can see that a more decentralized economy is characterized by a higher average income. This is a direct consequence of the result we explained above that more fiscal decentralization lead to greated labor supply. However, fiscal decentralization is costly for individuals. We see in Table A4 that both welfare variables V^a, V^{eq} decrease with J suggesting that the potential benefits from living in a more homogenous jurisdiction generated by more fiscal decentralization are outweighed by increased costs of effort.

Third, an increase in degree of fiscal decentralization may lead to higher disparities in income. In Table A5, we see that standard deviations of income increase with J. Figure A3 also corroborate this result as whatever $J > 1$ the income distribution obtained with some J is always Lorenz-dominated by the income distribution under the $J - 1$ case. This is due to the fact that labor supply is more disparate when fiscal decentralization increases. On the contrary, we can see that the distribution of utility V^{eq} is less unequal the more decentralized is the economy.[12]

[12]We also run simulations with $\sigma < 1$ (Assumption 2b). Under Assumption 2b, we still obtain that fiscal decentralization makes individuals supply mor labor while it is harmful from a welfare point of view. We also consider the case of a low ratio of transitory income to total earnings that is $\varrho_\varepsilon^2 / var\left(\ln y\left(\theta, \varepsilon\right)\right) = 0.05$. These simulations results are available upon request.

6 Conclusion

In this paper, we develop a multijurisdiction model where individuals differ with respect to their productivity and choose their effort before moving into their place of residence. In this set up, we show that an equilibrium with J jurisdictions and endogenous income distribution is productivity-stratified and jurisdictions are vertically differentiated. The degree of substitutability between public and private goods is key to know whether the rich jurisdictions are producing higher levels or lower levels of local public goods. Noticeably, the equilibrium income distribution is characterized by discontinuities each one corresponding to a change in jurisdiction membership. These income distribution discontinuities come from the fact that individual effort depends on the local taxation applied in the jurisdiction.

Further, we characterize the optimum. In particular, efficient sorting is such that the migration of any boundary individual between two adjacent jurisdictions does not generate any variation in total resources that could make some individuals better-off.

Despite asymmetric information on productivity, we show that there exists a tax structure that allows us to implement efficient allocations. This tax structure is such that the tax implemented by the government is characterized by a grant and a lump sum. The grant exactly compensates the distortion generated by local taxation on effort. The lump sum depends on jurisdiction membership. Its magnitude is such that sorting is efficient. It is also such that it fully corrects distortions generated by the voting procedure.

Finally, we develop a stochastic version of the model. We show that income distribution is more unequal under decentralization than under centralization. The reason comes from the fact that decentralization increases the variability of effort and thus income levels. Thanks to numerical simulations, we can see that an increase in fiscal decentralization leads to higher aggregate effort but a lower total welfare. This contradicts the intuition that competition between jurisdictions provides efficiency gains compared to centralization.

Our model opens up the avenue to investigate the issue whether individual effort decisions exacerbate or attenuate segregation forces. A possible strategy would be to compare how, after some exogenous macroeconomic shock, boundary individuals change in the standard Tiebout model with exogenous income distribution and in our model. We would then be able to study how segregation depends on ability inequality, on the one hand, and on different effort choices, on the other hand. This is left for further research.

References

[1] Arzaghi, M. and J. V. Henderson, 2002, "Why are Countries Fiscally Decentralized?", mimeo.

[2] Bayer, P. et R. McMillan, 2012, "Tiebout Sorting and Neighborhood Stratification", *Journal of Public Economics*, vol. 96, 1129-1143.

[3] Bénabou, R., 1996a, "Equity and Efficiency in Human Capital Investment: the Local Connection", *Review of Economic Studies*, vol. 63, 237-264.

[4] Bénabou, R., 1996b, "Heterogeneity, Stratification and Growth: Macroeconomic Implications of Community Structure and School Finance", *American Economic Review*, vol. 86, 584-609.

[5] Biswas, R., Gravel, N. and R. Oddou, 2011, "The Segregative Properties of Endogenous Formation of Jursidictions with a Welfarist Central Government", mimeo.

[6] Black, D., Kolesnikova, N. and L. Taylor, 2008, "Local Price Variation and Labor Supply Behavior", *Federal Reserve Bank of St Louis Review*, vol. 91, 613-625.

[7] Black, D., Kolesnikova, N. and L. Taylor, 2009, "Earnings Functions when Wages and Prices Vary by Location", *Journal of Labor Economics*, vol. 27, 21-47.

[8] Calabrese, S., Epple, D. and R. Romano, 2012, "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition", *Review of Economic Studies*, vol. 79, 1081-1111.

[9] Cooper, S., 1998, "A Positive Theory of Income Redistribution", *Journal of Economic Growth*, vol. 3, 171-195.

[10] Cremer, H., Gahvari, F. and J.M. Lozachmeur, 2010, "Tagging and Income Taxation: Theory and an Application", *American Economic Journal: Economic Policy*, vol. 2, 31-50.

[11] de Bartolome, Ch. and S. Ross, 2007, "Community Income Distributions in a Metropolitan Area", *Journal of Urban Economics*, vol. 61, 496-518.

[12] Demange, G., 1994, "Intermediate Preferences and Stable Coalition Structures", *Journal of Mathematical Economics*, vol. 23, 45-58.

[13] Durlauf, S., 1996, "A Theory of Persistent Income Inequality", *Journal of Economic Growth*, vol. 1, 75-93.

[14] Epple, D., Filimon, R. and T. Romer, 1993, “Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes”, *Regional Science and Urban Economics*, vol. 23, 585-610.

[15] Epple, D. and T. Romer, 1989, “On the Flexibility of Municipal Boundaries”, *Journal of Urban Economics*, vol. 26, 307-319.

[16] Epple, D. and T. Romer, 1991, “Mobility and Redistribution”, *Journal of Political Economy*, vol. 99, 828-858.

[17] Epple, D., Romer, T. et H. Sieg, 2001, “Interjurisdictional Sorting and Majority Rule: An Empirical Analysis”, *Econometrica*, vol. 69, 1437-1466.

[18] Epple, D. and T. Nechyba, 2004, “Fiscal Decentralization”, in J. V. Henderson and J.-F. Thisse (eds), *Handbook of Regional and Urban Economics*, vol. 4, 2423-2480, North Holland, Amsterdam.

[19] Fernandez, R. and Rogerson R., 1996, “Income Distribution, Communities and the Quality of Public Education”, *Quarterly Journal of Economics*, vol. 111, 135-164.

[20] Gravel, N. and S. Thoron, 2007, “Does Endogenous Formation of Jurisdictions Lead to Wealth Stratification?”, *Journal of Economic Theory*, vol. 132, 569-583.

[21] Grandmont, J.M., 1978, “Intermediate Preferences and the Majority Rule”, *Econometrica*, vol. 46, 317-330.

[22] Hansen, N. and A. Kessler, 2001, “(Non-) Existence of Equilibria in Multicommunity Models”, *Journal of Urban Economics*, vol. 50, 418-435.

[23] Kempf, H. and F. Moizeau, 2009, “Inequality, Growth and the Dynamics of Social Segmentation”, *Journal of Public Economic Theory*, vol. 11, 529-564.

[24] Moretti, E., 2011, “Local Labor Markets”, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics,* vol. 4, 1237-1313, North Holland, Amsterdam.

[25] Oates, W.E., 1972, Fiscal Federalism, New York; Harcourt Brace Jovanovich.

[26] OECD, 2008, “Growing Unequal? Income Distribution and Poverty in OECD Countries,” OECD Publications, Paris.

[27] Reardon S and K. Bischoff, 2011, “Growth in the Residential Segregation of Families by Income, 1970-2009,” *American Journal of Sociology*, 116(4), 1092-1153.

[28] Rubinfeld, D., 1987, "The Economics of the Local Public Sector", in A. Auerbach and M. Feldstein (eds), *Handbook of Public Economics,* vol. 2, , North Holland, Amsterdam, 571-645.

[29] Sacchi A. and S. Salotti, 2012, "The Effects of Fiscal Decentralization on Household Income Inequality: some Empirical Evidence", Department of Roma University Working Paper.

[30] Sepulveda, C.F. and J. Martinez-Vasquez, 2011, "The Consequences of Fiscal Decentralization on Poverty and Income Inequality," *Environment and Planning C: Government and Policy*, vol. 29, 321-343.

[31] Westhoff, F., 1977, "Existence of Equilibria in Economies with a Local Public Good", *Journal of Economic Theory*, vol. 14, 84-112.

7 Appendix

7.1 Proof of Lemma 1

Proof. Let us consider (3).

(i) Using the implicit function theorem yields:

$$\frac{\partial \tau_j}{\partial y} = \frac{u'(c_j) + c_j u''(c_j)}{\frac{\partial^2 \mathcal{U}}{\partial {\tau_j}^2}}$$

Given Assumption 2a (2b), we deduce that $\partial \tau_j / \partial y > (<) 0$.

(ii) Using the implicit function theorem yields:

$$\frac{\partial \tau_j}{\partial \mu_j} = -\frac{\frac{n_j}{\zeta(n_j)}\beta \left[u'(q_j) + q_j u''(q_j)\right]}{\frac{\partial^2 \mathcal{U}}{\partial {\tau_j}^2}}$$

Given Assumption 2a (2b), we deduce that $\partial \tau_j / \partial \mu_j < (>) 0$.

(iii) Using the implicit function theorem and the envelope theorem yields:

$$\frac{\partial q_j}{\partial \mu_j} = -\frac{y \frac{n_j}{\zeta(n_j)}}{\frac{\partial^2 \mathcal{U}}{\partial {\tau_j}^2}} [-\tau_j y u''(c_j) + u'(c_j)].$$

Given the concavity of $u(.)$ we deduce that $\partial q_j / \partial \mu_j > 0$. ■

7.2 Proof of Lemma 2

Proof. (i) Using the implicit function theorem, we have

$$\frac{\partial \widehat{y}_j(\theta)}{\partial \theta} = \frac{\frac{-1}{\theta^2} v'(\frac{\widehat{y}_j(\theta)}{\theta}) - \frac{y}{\theta^3} v''(\frac{\widehat{y}_j(\theta)}{\theta})}{\frac{\partial^2 \mathcal{U}}{(\partial \widehat{y}_j(\theta))^2}} > 0.$$

(ii) Given that $\widehat{c}_j(\theta) = (1 - \tau_j^*)\widehat{y}_j(\theta)$ and using the implicit function theorem leads to

$$\frac{\partial \widehat{y}_j(\theta)}{\partial \tau_j^*} = -\frac{-u'\left(\widehat{c}_j(\theta)\right) - \Big((1-\tau_j^*)\widehat{y}_j(\theta)\Big) u''(\widehat{c}_j(\theta))}{\frac{\partial^2 \mathcal{U}}{(\partial \widehat{y}_j(\theta))^2}}.$$

Given Assumption 2a (2b), *i.e.* $-u''c/u' > (<) 1$, we deduce that $\partial \widehat{y}_j(\theta) / \partial \tau_j^* > (<) 0$. ■

7.3 Proof of Lemma 3

Proof. Let us express $V_{j'}(\theta) - V_j(\theta)$

$$\begin{aligned} V_{j'}(\theta) - V_j(\theta) \equiv\; & u((1-\tau_{j'})\widehat{y}_{j'}(\theta)) + \beta u(q_{j'}) - v\left(\frac{\widehat{y}_{j'}(\theta)}{\theta}\right) \\ & - \left[u((1-\tau_j)\widehat{y}_j(\theta)) + \beta u(q_j) - v\left(\frac{\widehat{y}_j(\theta)}{\theta}\right)\right] \end{aligned}$$

Differentiating with respect to θ and applying the envelope theorem yields

$$\frac{\partial}{\partial\theta}\left(V_{j'}(\theta)-V_j(\theta)\right)=\frac{\widehat{y}_{j'}(\theta)}{\theta^2}v'(\frac{\widehat{y}_{j'}(\theta)}{\theta})-\frac{\widehat{y}_j(\theta)}{\theta^2}v'(\frac{\widehat{y}_j(\theta)}{\theta}).$$

Given Corollary 1, with $\tau_{j'}>\tau_j$, under Assumption 2a (2b) $\widehat{y}_{j'}(\theta)>(<)\widehat{y}_j(\theta)$ and $v''(.)>0$, we deduce that $V_{j'}(\theta)-V_j(\theta)$ increases (decreases) with θ. ■

7.4 Proof of Proposition 2

Proof. Let us assume by contradiction that at equilibrium one jurisdiction j is formed of two productivity intervals, namely $\left[\underline{\theta}_1,\overline{\theta}_1\right]\cup\left[\underline{\theta}_2,\overline{\theta}_2\right]$. Thus we have for any $\theta\in\left[\underline{\theta}_1,\overline{\theta}_1\right]\cup\left[\underline{\theta}_2,\overline{\theta}_2\right]$, $V_j(\theta)>V_{j'}(\theta)$, whatever $j,j'=1,...,J$ and $j'\neq j$. Let us consider the individual θ', with $\overline{\theta}_1<\theta'<\underline{\theta}_2$, who belongs to jurisdiction $j'\neq j$, we thus have at equilibrium that $V_{j'}(\theta')>V_j(\theta')$. Let us assume without loss of generality that $(\tau_j,q_j)<<(\tau_{j'},q_{j'})$. Given Lemma 3, if for the individual θ', we have $V_{j'}(\theta')>V_j(\theta')$, under Assumption 2a (2b) it turns out that for any $\theta''>(<)\theta'$ and in particular for any $\theta''\in\left[\underline{\theta}_2,\overline{\theta}_2\right]$ $(\left[\underline{\theta}_1,\overline{\theta}_1\right])$, we have $V_{j'}(\theta'')>V_j(\theta'')$. Hence, a contradiction.

Further, since any individual lives in one and only one community, the equilibrium set of jurisdictions is a partition of $\left[\underline{\theta},\overline{\theta}\right]$. ■

7.5 Proof of Proposition 3

Proof. Item (i).

Let us consider two adjacent jurisdictions j and $j+1$ defined by $[\widetilde{\theta}_{j-1},\widetilde{\theta}_j]$ and $[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]$, respectively. Assume by contradiction that at equilibrium we have $V_j^*(\widetilde{\theta}_j)<V_{j+1}^*(\widetilde{\theta}_j)$. As $V_j^*(\theta)$ is a continuous function, for $\theta=\widetilde{\theta}_j-\varepsilon$ with ε an infinitesimal positive number, we have $V_j^*(\theta)<V_{j+1}^*(\theta)$. As $\theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j]$, he lives in jurisdiction j. He should thus move. Hence, a contradiction. ■

Proof. Item (ii).

Let us consider only the Assumption 2a case. Let us consider two adjacent jurisdictions j and $j+1$ defined by $[\widetilde{\theta}_{j-1},\widetilde{\theta}_j]$ and $[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]$, respectively. Assume by contradiction that at equilibrium we have $(\tau_{j+1},q_{j+1})<<(\tau_j,q_j)$ and $V_j(\theta)\leq(\geq)V_{j+1}(\theta)$ for any $\theta\in[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]\left([\widetilde{\theta}_{j-1},\widetilde{\theta}_j]\right)$. By Lemma 3, if for an individual θ' we have $V_j(\theta')\geq V_{j+1}(\theta')$ then for any $\theta''>\theta'$ and in particular any $\theta''\in[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]$ we also have $V_j(\theta'')>V_{j+1}(\theta'')$. Hence all individuals $\theta\in[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]$ have an incentive to move, leading to a contradiction.

Proof under Assumption 2b would be similar so we do not provide it. ■

Proof. Item (iii). Under either Assumption 2a or 2b, for any $\theta\in[\widetilde{\theta}_j,\widetilde{\theta}_{j+1}]$, and any $j=1,...,J-1$, $\widetilde{\theta}_0=\underline{\theta}$ and $\widetilde{\theta}_J=\overline{\theta}$, we know from Lemma 3 that

$$V_{j+1}(\theta)>V_{j'}\text{ for any }j'\neq j+1$$

Hence, $y^*(\theta) = \widehat{y}_j(\theta)$. ■

Proof. Item (iv). We aim to show that for any $\theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j]$ her best labor-supply-response is $y_j^*(\theta)$ then her best location is j.

Definition 2 *$\widehat{\theta}_{j,j}$ and $\widehat{\theta}_{j+1,j}$ are respectively such that $\widehat{y}_j(\widehat{\theta}_{j,j}) = \widetilde{y}_j$ and $\widehat{y}_{j+1}(\widehat{\theta}_{j+1,j}) = \widetilde{y}_j$.*

Given item (i) of Lemma 2 and that $\widetilde{y}_{j+1} > \widetilde{y}_j$ it turns out that $\widehat{\theta}_{j+1,j} < \widehat{\theta}_{j+1,j+1}$. We aim to show that item (iii) is satisfied only if $[\widetilde{\theta}_j, \widetilde{\theta}_{j+1}] \subseteq [\widehat{\theta}_{j+1,j}, \widehat{\theta}_{j+1,j+1}]$ for any $j = 1, ..., J-1$, $\widetilde{\theta}_0 = \underline{\theta}$ and $\widetilde{\theta}_J = \overline{\theta}$.

We then show that:

Lemma 4 *$y^*(\theta) = \widehat{y}_{j+1}(\theta) \in [\widetilde{y}_j, \widetilde{y}_{j+1}]$ for any $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$, any $j = 1, ..., J-1$, $\widetilde{\theta}_0 = \underline{\theta}$ and $\widetilde{\theta}_J = \overline{\theta}$.*

Take an individual θ such that $\in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$. Given item (iii), $y^*(\theta) = \widehat{y}_{j+1}(\theta)$. By Definition 2 and Lemma 2, we have $y^*(\theta) \in [\widehat{y}_{j+1}(\widetilde{\theta}_j), \widehat{y}_{j+1}(\widetilde{\theta}_{j+1})]$ for any $\theta \in [\widetilde{\theta}_j, \widetilde{\theta}_{j+1}]$.

We consider the two following cases.

(i) Assume by contradiction that $\widetilde{\theta}_j < \widehat{\theta}_{j+1,j}$. Hence, from item (i) of Proposition 3 and Lemma 3 for any agent $\theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]$, we have

$$V_{j+1}^*(\theta) - V_j(\theta) > 0. \tag{24}$$

From Lemma 2 we know that $\widehat{y}_{j+1}(\theta)$ increases with θ. Hence, we have $\widehat{y}_{j+1}(\theta) < \widehat{y}_{j+1}(\widehat{\theta}_{j+1,j}) = \widetilde{y}_j$ for all $\theta < \widehat{\theta}_{j+1,j}$. Given (6) we thus have

$$u((1-\tau_j)\widehat{y}_{j+1}(\theta)) + \beta u(q_j^*) \geq u((1-\tau_{j+1})\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*), \text{ for all } \theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]. \tag{25}$$

It is equivalent to

$$\begin{aligned} &u((1-\tau_j)\widehat{y}_{j+1}(\theta)) + \beta u(q_j^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right) \\ &\geq u((1-\tau_{j+1})\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]. \end{aligned}$$

As, for any j, $\widehat{y}_j(\theta) = \arg\max_y u((1-\tau_j)y) - v\left(\frac{y}{\theta}\right)$ we have

$$\begin{aligned} &u((1-\tau_j)\widehat{y}_j(\theta)) + \beta u(q_j^*) - v\left(\frac{\widehat{y}_j(\theta)}{\theta}\right) \\ &> u((1-\tau_j)\widehat{y}_{j+1}) + \beta u(q_j^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]. \end{aligned}$$

So we deduce that

$$u((1-\tau_j)\widehat{y}_j(\theta)) + \beta u(q_j^*) - v\left(\frac{\widehat{y}_j(\theta)}{\theta}\right)$$
$$> u((1-\tau_{j+1})\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]$$

leading to $V_{j+1}^*(\theta) - V_j(\theta) < 0$, for all $\theta \in [\widetilde{\theta}_j, \widehat{\theta}_{j+1,j}]$. Hence a contradiction.

(ii) By contradiction, assume that $\widetilde{\theta}_{j+1} > \widehat{\theta}_{j+1,j+1}$. Hence, from item (i) of Proposition 3 and Lemma 3 for any agent $\theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}]$, we have

$$V_{j+2}(\theta) - V_{j+1}^*(\theta) < 0. \tag{26}$$

From Lemma 2, we have $y_{j+1}^*(\theta) = \widehat{y}_{j+1}(\theta) > \widehat{y}_j(\widehat{\theta}_{j+1,j+1}) = \widetilde{y}_{j+1}$ for all $\theta > \widehat{\theta}_{j+1,j+1}$. Given (5), we thus have

$$u((1-\tau_{j+2}^*)\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+2}^*) \geq \tag{27}$$
$$u((1-\tau_{j+1}^*)\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*), \text{ for all } \theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}]. \tag{28}$$

It is equivalent to

$$u((1-\tau_{j+2}^*)\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+2}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right)$$
$$\geq u((1-\tau_{j+1}^*)\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}].$$

As, for any j, $\widehat{y}_j(\theta) = \arg\max_y u((1-\tau_j)y) - v\left(\frac{y}{\theta}\right)$ we have

$$u((1-\tau_{j+2})\widehat{y}_{j+2}(\theta)) + \beta u(q_{j+2}^*) - v\left(\frac{\widehat{y}_{j+2}(\theta)}{\theta}\right)$$
$$> u((1-\tau_{j+2})\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+2}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}].$$

So we deduce that

$$u((1-\tau_{j+2})\widehat{y}_{j+2}(\theta)) + \beta u(q_{j+2}^*) - v\left(\frac{\widehat{y}_{j+2}(\theta)}{\theta}\right)$$
$$> u((1-\tau_{j+1}^*)\widehat{y}_{j+1}(\theta)) + \beta u(q_{j+1}^*) - v\left(\frac{\widehat{y}_{j+1}(\theta)}{\theta}\right), \text{ for all } \theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}].$$

leading to $V_{j+2}(\theta) - V_{j+1}^*(\theta) > 0$, for all $\theta \in [\widehat{\theta}_{j+1,j+1}, \widetilde{\theta}_{j+1}]$. Hence a contradiction.

We thus have $[\widetilde{\theta}_j, \widetilde{\theta}_{j+1}] \subseteq [\widehat{\theta}_{j+1,j}, \widetilde{\theta}_{j+1,j+1}]$ leading to have $[\widehat{y}_{j+1}(\widetilde{\theta}_j), \widehat{y}_{j+1}(\widetilde{\theta}_{j+1})] \subseteq [\widetilde{y}_j, \widetilde{y}_{j+1}]$. Hence the result. ■

7.6 Existence of Equilibrium

Proposition 7 *Under Assumption 2a, there exists an equilibrium with J jurisdictions and endogenous income distribution.*

The proof relies on Westhoff (1977). The only departure from Westhoff (1977)'s model is that individuals choose their labor supply. We thus must check whether the income distribution that arises at equilibrium satisfies Westhoff's Assumption 3.

Proof. Our existence result entirely builds on Westhoff's proof. So this proof is valid only under Assumption 2a. Let us present the main lines of reasoning. We denote by $\overline{n} = (n_1, ..., n_j, ...n_J)$ the partition of $[\underline{\theta}, \overline{\theta}]$ into J jurisdictions. In order to apply Brouwer's fixed point Theorem, we need to define a compact and convex set to which the vector $\overline{n}$ belongs. First, we consider the two following sets $\mathcal{J}$ and $\mathcal{J}'$, $\mathcal{J}$ the set of all points on the interior of the $J-$dimensional unit simplex, and $\mathcal{J}'$ the set of all points on the $J-$dimensional unit simplex such that $n_j \geq m > 0$ for any $j = 1, ..., J$. $\mathcal{J}$ is an open set. $\mathcal{J}'$ is compact and convex. Westhoff shows that under assumptions analog to our Assumptions 1, 2a and 3 all jurisdictions must have a population size greater than m at the free mobility equilibrium. Moreover, Westhoff considers an income distribution such that each individual has a strictly positive income. In our set-up where the income distribution is endogenous, let us mention that $y^*(\theta) > 0$ for any θ as given that $0 < \tau_j^* < 1$ and $\lim_{c \longrightarrow 0} u'(0) = +\infty$, $y_j^*(\theta)$ is an interior solution of (8). Second, following Westhoff, denoting by $\lambda_j(\overline{n})$ the number of individuals in jurisdiction j who want to move into another jurisdiction, $0 \leq \lambda_j(\overline{n}) \leq n_j$ for $j = 1, ..., J$, we can use $g \circ f$ with

$$f : \mathcal{J}' \longrightarrow \mathcal{J} \text{ where } f(n_j) = \frac{n_j + \lambda_j(\overline{n})}{1 + \sum\limits_{i=1}^{J} \lambda_i(\overline{n})} \text{ over all } j = 1, ..., J$$

and

$$g : \mathcal{J} \longrightarrow \mathcal{J}' \text{ where}$$

$$(g \circ f)(n_j) = f(n_j) + \left(\frac{1}{J} - f(n_j)\right)\left(\max_{j=1,...,J}\{h_j(f(n_1), ..., f(n_J))\}\right)$$

$$\text{with } h_j(f(n_1), ..., f(n_J)) = \begin{cases} \frac{m - f(n_j)}{(1/J) - f(n_j)} \text{ if } f(n_j) < m, \\ 0 \text{ if } f(n_j) \geq m \end{cases} \text{ over all } j = 1, ..., J$$

which is a continuous mapping from $\mathcal{J}$ into itself. We can thus apply Brouwer's fixed point theorem. There thus exists $\overline{n}^*$ such that $\overline{n}^* = (g \circ f)(\overline{n}^*)$. Theorem 4 in Westhoff (1977) shows that if $\overline{n}^* = (g \circ f)(\overline{n}^*)$ then $\overline{n}^* = f(\overline{n}^*)$. The final step of Westhoff's proof is to show that for the vector $\overline{n}^*$ no indiviudal has an incentive to move, *i.e.* $\lambda_j(\overline{n}) = 0$ over all $j = 1, ..., J$. This ends the proof. ■

7.7 Construction Figure 1

The existence of a kink at point $\widetilde{y}_j$ characterizing the upper envelope is proved in the following Lemma.

Lemma 5 *Given Assumption 2a (2b), when* $\tau_{j+1} > (<)\tau_j$ *we have for any* y $(1-\tau_{j+1})u'((1-\tau_{j+1})y) > (1-\tau_j)u'((1-\tau_j)y)$.

Proof. Assumption 2a (2b) can be written as follows

$$-(1-\tau_j)yu''((1-\tau_j)y) - u'((1-\tau_j)y) > (<)0, \ \forall \tau_j, y.$$

It is equivalent to

$$\frac{\partial}{\partial \tau_j}\left[(1-\tau_j)u'((1-\tau_j)y)\right] > (<)0.$$

With $\tau_{j+1} > (<)\tau_j$, we easily deduce that

$$(1-\tau_{j+1})u'((1-\tau_{j+1})y) > (1-\tau_j)u'((1-\tau_j)y).$$

■

7.8 Proof of Proposition 4

Proof. In order to solve the planner's program, we write Khun-Tucker conditions as follows

$$u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) = u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right) + s_{jk}(\theta)^2$$
$$\text{for any } j,k = 1,...,J,\ k \neq j \text{ and any } \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j]$$

with $s_{jk}(\theta) \in \mathbb{R}$.

The Lagrangian of the optimization program can be written:

$$\begin{aligned}
\pounds = {} & \sum_{j=1}^{J} \int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} \omega(\theta)\left[u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right)\right] h(\theta)d\theta \\
& + \lambda\left(\sum_{j=1}^{J} \int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} y_j(\theta)h(\theta)d\theta - \sum_{j=1}^{J} \int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} c_j(\theta)h(\theta)d\theta - \sum_{j=1}^{J}\zeta(n_j)q_j\right) \\
& + \sum_{j=1}^{J} \sum_{k=1,k\neq j}^{J} \int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} \chi_{jk}(\theta)\left[u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right)\right. \\
& \left. - \left(u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right) + s_{jk}(\theta)^2\right)\right] h(\theta)d\theta
\end{aligned}$$

with λ the Lagrange multiplier associated to the resource constraint, $\chi_{jk}(\theta)$ the Lagrange multiplier associated to (10). Let us now study the first order conditions[13].

First, we have

$$\frac{\partial \pounds}{\partial s_{jk}(\theta)} = 0 \Leftrightarrow \chi_{jk}(\theta) s_{jk}(\theta) = 0, \ \forall \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j].$$

As at the optimum, we must have

$$\chi_{jk}(\theta) \left[u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) - \left(u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right) + s_{jk}(\theta)^2 \right) \right] = 0,$$
$$\forall \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j, k, k \neq j$$

For a given θ, we thus have

$$\text{if } u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) > (=) u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_{jk}(\theta)}{\theta}\right)$$
$$\text{then } s_{jk}(\theta) \neq (=) 0 \text{ implying that} \chi_{jk}(\theta) = (\geq) 0. \tag{29}$$

It is important to emphasize that $\chi_{jk}(\theta)$ could be equal to zero when

$$u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) = u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_{jk}(\theta)}{\theta}\right).$$

Second, the FOC with respect to $c_j(\theta)$ gives

$$\frac{\partial \pounds}{\partial c_j(\theta)} = 0 \Leftrightarrow \left(\omega(\theta) + \sum_{k=1, k\neq j}^{J} \chi_{jk}(\theta) \right) u'(c_j(\theta)) = \lambda, \ \forall \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j. \tag{30}$$

Third, the FOC with respect to $y_j(\theta)$ gives

$$\frac{\partial \pounds}{\partial y_j(\theta)} = 0 \Leftrightarrow \left(\omega(\theta) + \sum_{k=1, k\neq j}^{J} \chi_{jk}(\theta) \right) \frac{1}{\theta} v'\left(\frac{y_j(\theta)}{\theta}\right) = \lambda, \ \forall \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j. \tag{31}$$

Equations (30) and (31) imply

$$u'(c_j(\theta)) = \frac{1}{\theta} v'\left(\frac{y_j(\theta)}{\theta}\right), \ \forall \theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j. \tag{32}$$

Fourth, let us study the FOC with respect to $c_k(\theta)$ and $y_k(\theta)$ for any $\theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j, k, k \neq j$. From (29), if

$$u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) > u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right) \text{ then } \chi_{jk}(\theta) = 0, \tag{33}$$

it turns out that

$$\frac{\partial \pounds}{\partial c_k(\theta)} = -\chi_{jk}(\theta) u'(c_k(\theta)) = 0 \text{ and } \frac{\partial \pounds}{\partial y_k(\theta)} = \chi_{jk}(\theta) \frac{1}{\theta} v'\left(\frac{y_k(\theta)}{\theta}\right) = 0.$$

[13] For notational purpose, we omit to mention that with $\widetilde{\theta}_0 = \underline{\theta}$ and $\widetilde{\theta}_J = \overline{\theta}$.

From (29), if

$$u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right) = u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right) \text{ then } \chi_{jk}(\theta) \geq 0.$$

It turns out that an infinitesimal increase in $c_k(\theta)$ generates a migration of $h(\theta)$ agents of productivity θ from community j to community k. Hence,

$$\begin{aligned}\frac{\partial \pounds}{\partial c_k(\theta)} = &- \left[u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right)\right]\omega(\theta)h(\theta)\\ &+ \left[u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right)\right]\omega(\theta)h(\theta)\\ &- \lambda h(\theta)\left[y_j(\theta) - c_j(\theta) - \zeta'(n_j)q_j - y_k(\theta) + c_k(\theta) + \zeta'(n_k)q_k\right]\\ &- \chi_{jk}(\theta)u'(c_k(\theta))h(\theta) = 0 \Rightarrow\end{aligned}$$

$$-y_j(\theta) + c_j(\theta) + y_k(\theta) - c_k(\theta) + \zeta'(n_k)q_k - \zeta'(n_k)q_k = \frac{\chi_{jk}(\theta)}{\lambda}u'(c_k(\theta)). \tag{34}$$

Further, an infinitesimal decrease in $y_k(\theta)$ generates a migration of $h(\theta)$ agents of productivity θ from community j to community k. Hence,

$$\begin{aligned}\frac{\partial \pounds}{\partial y_k(\theta)} = &- \left[u(c_j(\theta)) + \beta u(q_j) - v\left(\frac{y_j(\theta)}{\theta}\right)\right]\omega(\theta)h(\theta)\\ &+ \left[u(c_k(\theta)) + \beta u(q_k) - v\left(\frac{y_k(\theta)}{\theta}\right)\right]\omega(\theta)h(\theta)\\ &- \lambda h(\theta)\left[y_j(\theta) - c_j(\theta) - \zeta'(n_j)q_j - y_k(\theta) + c_k(\theta) + \zeta'(n_k)q_k\right]\\ &- \chi_{jk}(\theta)\frac{1}{\theta}v'\left(\frac{y_k(\theta)}{\theta}\right)h(\theta) = 0 \Rightarrow\end{aligned}$$

$$-y_j(\theta) + c_j(\theta) + y_k(\theta) - c_k(\theta) + \zeta'(n_j)q_j - \zeta'(n_k)q_k = \frac{\chi_{jk}(\theta)}{\lambda}\frac{1}{\theta}v'\left(\frac{y_k(\theta)}{\theta}\right). \tag{35}$$

Fifth, let us study the FOC with respect to q_j. We have

$$\frac{\partial \pounds}{\partial q_j} = 0 \Longleftrightarrow \beta u'(q_j)\left(\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\omega(\theta) + \sum_{k=1,k\neq j}^{J}\chi_{jk}(\theta)\right)h(\theta)d\theta\right) = \lambda\zeta(n_j),\ \forall\theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j, \tag{36}$$

Equations (30) and (36) imply

$$\beta u'(q_j)\left(\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\frac{1}{u'\left(c_j(\theta)\right)}h(\theta)d\theta\right) = \zeta(n_j),\ \forall\theta \in [\widetilde{\theta}_{j-1}, \widetilde{\theta}_j], \forall j, \tag{37}$$

Finally, let us consider the optimal allocations of individuals in the different jurisdictions. On the one hand, the government can't increase social welfare by moving any agent θ from community j

to k. Formally, we have

$$\omega(\theta)\left[u(c_k(\theta))+\beta u(q_k)-v\left(\frac{y_k(\theta)}{\theta}\right)\right]-\omega(\theta)\left[u(c_j(\theta))+\beta u(q_j)-v\left(\frac{y_j(\theta)}{\theta}\right)\right]$$
$$+\lambda\left[y_k(\theta)-y_j(\theta)+c_j(\theta)-c_k(\theta)\right]-\lambda\left[\zeta'(n_k)q_k-\zeta'(n_j)q_j\right]\leq 0$$
$$\forall\theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j],\forall j,k\neq j$$

which is equivalent to

$$\omega(\theta)\left[u(c_j(\theta))+\beta u(q_j)-v\left(\frac{y_j(\theta)}{\theta}\right)\right]-\omega(\theta)\left[u(c_k(\theta))+\beta u(q_k)-v\left(\frac{y_k(\theta)}{\theta}\right)\right]$$
$$\geq\lambda\left[y_k(\theta)-y_j(\theta)+c_j(\theta)-c_k(\theta)-\zeta'(n_k)q_k+\zeta'(n_j)q_j\right],\ \forall\theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j],\forall j,k\neq j. \tag{38}$$

On the other hand, the government can't increase the social utility by moving any agent θ' from community k to j. Therefore

$$\omega(\theta')\left[u(c_j(\theta'))+\beta u(q_j)-v\left(\frac{y_j(\theta')}{\theta'}\right)\right]-\omega(\theta')\left[u(c_k(\theta'))+\beta u(q_k)-v\left(\frac{y_k(\theta')}{\theta'}\right)\right]$$
$$+\lambda\left[y_j(\theta')-y_k(\theta')-c_j(\theta')+c_k(\theta')\right]-\lambda\left[\zeta'(n_j)q_j-\zeta'(n_{j+1})q_{j+1}\right]\leq 0,$$

which is equivalent to

$$\omega(\theta')\left[u(c_j(\theta'))+\beta u(q_j)-v\left(\frac{y_j(\theta')}{\theta'}\right)\right]-\omega(\theta')\left[u(c_k(\theta'))+\beta u(q_k)-v\left(\frac{y_k(\theta')}{\theta'}\right)\right]$$
$$\leq\lambda\left[y_k(\theta')-y_j(\theta')+c_j(\theta')-c_k(\theta')-\zeta'(n_k)q_k+\zeta'(n_j)q_j\right]. \tag{39}$$

From (29), if

$$u(c_j(\theta))+\beta u(q_j)-v\left(\frac{y_j(\theta)}{\theta}\right)=u(c_k(\theta))+\beta u(q_k)-v\left(\frac{y_k(\theta)}{\theta}\right)\text{ then }\chi_{jk}(\theta)\geq 0.$$

Given (34) and (38), this implies

$$0\geq\lambda\left[y_k(\theta)-y_j(\theta)+c_j(\theta)-c_k(\theta)-\zeta'(n_k)q_k+\zeta'(n_j)q_j\right]=\chi_{jk}(\theta)u'(c_k(\theta)),\ \forall\theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j],\forall j,k\neq j \tag{40}$$

As $u'(c_k(\theta))>0$, equation (40) implies that

$$\chi_{jk}(\theta)=0,\ \forall\theta\in[\widetilde{\theta}_{j-1},\widetilde{\theta}_j],\forall j,k\neq j\text{ and }y_k(\theta)-y_j(\theta)+c_j(\theta)-c_k(\theta)-\zeta'(n_k)q_k+\zeta'(n_j)q_j=0. \tag{41}$$

From (38) and (39), we have

$$\omega(\widetilde{\theta}_j)\left[u(c_j(\widetilde{\theta}_j))+\beta u(q_j)-v\left(\frac{y_j(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]-\omega(\widetilde{\theta}_j)\left[u(c_{j+1}(\widetilde{\theta}_j))+\beta u(q_{j+1})-v\left(\frac{y_{j+1}(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]$$
$$\geq\lambda\left[y_{j+1}(\widetilde{\theta}_j)-y_j(\widetilde{\theta}_j)+c_j(\widetilde{\theta}_j)-c_{j+1}(\widetilde{\theta}_j)-\zeta'(n_{j+1})q_{j+1}+\zeta'(n_j)q_j\right]$$
$$\geq\omega(\widetilde{\theta}_j)\left[u(c_j(\widetilde{\theta}_j))+\beta u(q_j)-v\left(\frac{y_j(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]-\omega(\widetilde{\theta}_j)\left[u(c_{j+1}(\widetilde{\theta}_j))+\beta u(q_{j+1})-v\left(\frac{y_{j+1}(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]$$

implying that

$$\lambda\left[y_{j+1}(\widetilde{\theta}_j)-y_j(\widetilde{\theta}_j)+c_j(\widetilde{\theta}_j)-c_{j+1}(\widetilde{\theta}_j)-\zeta'(n_{j+1})q_{j+1}+\zeta'(n_j)q_j\right]$$
$$=\omega(\widetilde{\theta}_j)\left[u(c_j(\widetilde{\theta}_j))+\beta u(q_j)-v\left(\frac{y_j(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]-\omega(\widetilde{\theta}_j)\left[u(c_{j+1}(\widetilde{\theta}_j))+\beta u(q_{j+1})-v\left(\frac{y_{j+1}(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)\right]. \tag{42}$$

From (41) we deduce that

$$u(c_j(\widetilde{\theta}_j))+\beta u(q_j)-v\left(\frac{y_j(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right)=u(c_{j+1}(\widetilde{\theta}_j))+\beta u(q_{j+1})-v\left(\frac{y_{j+1}(\widetilde{\theta}_j)}{\widetilde{\theta}_j}\right).$$

■

7.9 Proof of Proposition 5

Proof. The aim of the proof is to characterize taxes so that equations (11)-(13) and (9) coincide with (16)-(19).

First, we can see that equations (11) and (16) are equivalent when for any j:

$$(1-\tau_j-T_1(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**}))=1\Leftrightarrow T_1(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**})=-\tau_j.$$

We thus deduce that $T(y,\tau_j,\mu_j,n_j,I)$ can be expressed as follows:

$$T(y,\tau_j,\mu_j,n_j,I)=-\tau_j y+\varphi(\tau_j,\mu_j,n_j,I) \tag{43}$$

Second, in order for the median voter most preferred level of local public good to satisfy the Bowen-Lindhal-Samuelson condition (see equations 17 and 12), taxes must be such that for any j

$$\left[y_j^{**}(\theta^m)+T_2(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**})\right]u'(c_j^{**}(\theta^m))=\left(\frac{\int_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} y_j^{**}(\theta)h(\theta)d\theta}{\int_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}\left(u'(c_j^{**}(\theta))\right)^{-1}h(\theta)d\theta}\right)$$

Given (43), we have

$$\frac{\partial\varphi(\tau_j^m,\mu_j^{**},n_j^{**},I^{**})}{\partial\tau_j}=\frac{\left(u'(c_j^{**}(\theta^m))\right)^{-1}\int_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}} y_j^{**}(\theta)h(\theta)d\theta}{\int_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}\left(u'(c_j^{**}(\theta))\right)^{-1}h(\theta)d\theta}$$

which is equivalent to

$$\frac{\partial\varphi(\tau_j^m,\mu_j^{**},n_j^{**},I^{**})}{\partial\tau_j}=\mu_j^{**}\frac{\left(u'(c_j^{**}(\theta^m))\right)^{-1}}{\left(\frac{1}{n_j^{**}}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}\left(u'(c_j^{**}(\theta))\right)^{-1}h(\theta)d\theta\right)}.$$

Third, given the "free-mobility" constraints, we know that optimal allocations satisfy the equilibrium equation (18). Moreover, given (43) and (15), equation (13) is satisfied when taxes are such that

$$\begin{aligned}&\varphi(\tau_j^{**},\mu_j^{**},n_j^{**},I^{**})-\varphi(\tau_{j+1}^{**},\mu_{j+1}^{**},n_{j+1}^{**},I^{**})\\&=q_j^{**}\zeta'(n_j^{**})-q_{j+1}^{**}\zeta'(n_{j+1}^{**})\ \forall j=1,...,J-1.\end{aligned}$$

Fourth, equations (9) and (19) are equivalent. Indeed let us consider (9):

$$\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}y_j^{**}(\theta)h(\theta)d\theta=\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}c_j^{**}(\theta)h(\theta)d\theta+\sum_{j=1}^{J}\zeta(n_j^{**})q_j^{**}.$$

Given (15) and (2), we have

$$\begin{aligned}\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}y_j^{**}(\theta)h(\theta)d\theta=&\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}\left((1-\tau_j^m)y_j^{**}(\theta)-T(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**})\right)h(\theta)d\theta\\&+\sum_{j=1}^{J}\tau_j^m\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}y_j^{**}(\theta)h(\theta)d\theta.\end{aligned}$$

It is equivalent to

$$\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}y_j^{**}(\theta)h(\theta)d\theta=\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}\left(y_j^{**}(\theta)-T(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**})\right)h(\theta)d\theta$$

and leads to

$$\sum_{j=1}^{J}\int\limits_{\widetilde{\theta}_{j-1}^{**}}^{\widetilde{\theta}_j^{**}}T(y_j^{**}(\theta),\tau_j^m,\mu_j^{**},n_j^{**},I^{**})h(\theta)d\theta=0.$$

Sorting Condition At Stage 2, the slope of the individual's indifference curve in space (q, τ) become:

$$S_T(q_j, \tau_j, y) = \frac{\beta u'(q_j)}{yu'((1-\tau_j)y - T(y, \overline{\tau}_j, \mu_j, n_j, I, \theta))}$$

Hence,

$$\frac{\partial S_T(q_j, \tau_j, y)}{\partial y} = -\frac{\beta u'(q_j)\left[u'(c_j) + y\left(1-\tau_j - T_1(y, \overline{\tau}_j, \mu_j, n_j, I, \theta)\right)u''(c_j)\right]}{\left[yu'((1-\tau_j)y - T(y, \overline{\tau}_j, \mu_j, n_j, I, \theta))\right]^2}$$

Given that $T_1(y, \overline{\tau}_j, \mu_j, n_j, I) = -\tau_j$, we have

$$\frac{\partial S_T(q_j, \tau_j, y)}{\partial y} = -\frac{\beta u'(q_j)\left[u'(c_j) + yu''(c_j)\right]}{\left[yu'(c_j)\right]^2}.$$

Hence,

$$\frac{\partial S_T(q_j, \tau_j, y)}{\partial y} \gtrless 0 \Leftrightarrow u'(c_j) + yu''(c_j) \lesseqgtr 0$$

leading to

$$\frac{\partial S_T(q_j, \tau_j, y)}{\partial y} \gtrless 0 \Leftrightarrow \frac{c_j}{y} \lesseqgtr -\frac{c_j u''(c_j)}{u'(c_j)}.$$

Let us stress that Assumption 2a (2b) does not guarantee that $\frac{\partial S_T(q_j,\tau_j,y)}{\partial y} > (<) 0$ as c_j may be higher (lower) than y.

Intermediate Preferences Let us express $V_{j'}(\theta) - V_j(\theta)$

$$\begin{aligned} V_{j'}(\theta) - V_j(\theta) \equiv\ & u((1-\tau_{j'})\widehat{y}_{j'}(\theta) - T(\widehat{y}_{j'}(\theta), \overline{\tau}_{j'}, \mu_{j'}, n_{j'}, I))) + u(q_{j'}) - v(\frac{\widehat{y}_{j'}(\theta)}{\theta}) \\ & - \left[u((1-\tau_j)\widehat{y}_j(\theta) - T(\widehat{y}_j(\theta), \overline{\tau}_j, \mu_j, n_j, I))) + u(q_j) - v(\frac{\widehat{y}_j(\theta)}{\theta})\right] \end{aligned}$$

Differentiating with respect to θ and applying the envelope theorem

$$\frac{\partial}{\partial \theta}\left(V_{j'}(\theta) - V_j(\theta)\right) = \frac{\widehat{y}_{j'}(\theta)}{\theta^2} v'(\frac{\widehat{y}_{j'}(\theta)}{\theta}) - \frac{\widehat{y}_j(\theta)}{\theta^2} v'(\frac{\widehat{y}_j(\theta)}{\theta}).$$

Considering (16) and applying the implicit function theorem leads to have $\widehat{y}_{j'}(\theta) > (<) \widehat{y}_j(\theta)$ when $\varphi_{j'} > (<) \varphi_j$. Given the convexity of $v(.) > 0$, we deduce that $V_{j'}(\theta) - V_j(\theta)$ increases (decreases) with θ. This ends the proof. ■

7.10 Stochastic Model, Numerical Values and Simulation Results

7.10.1 Stochastic Model and Proof of Proposition 6

Stochastic Model. We consider individuals who are characterized by a productivity parameter θ such that

$$\log \theta \rightsquigarrow \mathcal{N}\left(\widehat{\theta}, \varrho_\theta^2\right) \tag{44}$$

We normalize the mass of individuals to one. They earn a gross income now defined as follows

$$y(\theta) = \theta l \varepsilon$$

with l denoting the labor supply and ε an income shock that agents observe once individual effort, the jurisdiction are chosen and the vote on local taxation is made. We assume that

$$\log \varepsilon \rightsquigarrow \mathcal{N}\left(\psi, \varrho_{\varepsilon}^{2}\right) \tag{45}$$

$$\text{with } \mathbb{E}\left[\varepsilon\right] = 1. \tag{46}$$

Denoting by $\mathbb{E}[.]$ the expectation operator, an individual θ who lives in jurisdiction j has the following utility function:

$$\mathcal{U}_j(\theta) = \mathbb{E}\left[\frac{\left(c_j\left(\theta\right)\right)^{1-\sigma}}{1-\sigma} + \beta\frac{q^{1-\sigma}}{1-\sigma}\right] - \frac{\Phi}{\eta}\left(l_j\left(\theta\right)\right)^{\eta}, \text{ with } \Phi, \sigma > 0 \text{ and } \eta > 1.$$

The consumption of an agent θ who live in jurisdiction j is equal to

$$c_j(y(\theta)) = (1-\tau_j)y_j(\theta) = (1-\tau_j)\theta l_j\left(\theta\right)\varepsilon, \tag{47}$$

The public good of the jurisdiction j is equal to

$$q_j = \frac{\tau_j n_j \mu_j}{\zeta\left(n_j\right)} \text{ with } \zeta\left(n_j\right) \equiv n_j^{\nu},\ \nu \geq 0. \tag{48}$$

Let us mention that the properties of a lognormal distribution imply that

$$\mathbb{E}\left[\varepsilon\right] = e^{\psi + \frac{\varrho_{\varepsilon}^{2}}{2}}. \tag{49}$$

Using equations (46) and (49) we obtain $\psi = -\varrho_{\varepsilon}^{2}/2$.

The properties of the lognormal distribution also imply

$$\begin{aligned}\mathbb{E}\left[\varepsilon^{1-\sigma}\right] &= \mathbb{E}\left[e^{(1-\sigma)\log\varepsilon}\right] \\ &= e^{(1-\sigma)\psi + \frac{(1-\sigma)^2}{2}\varrho_{\varepsilon}^{2}} \\ &= e^{-\frac{\sigma(1-\sigma)\varrho_{\varepsilon}^{2}}{2}}.\end{aligned} \tag{50}$$

Equations characterizing the equilibrium are the following:

(i) For any j, any individual θ who lives in jurisdiction j, the choice of effort must satisfy the following first order condition:

$$\left((1-\tau_j)\theta\right)^{1-\sigma}\left(l_j\left(\theta\right)\right)^{-\sigma}\mathbb{E}\left[\varepsilon^{1-\sigma}\right] = \Phi\left(l_j\left(\theta\right)\right)^{\eta-1}$$

given (50), we have

$$((1-\tau_j)\theta)^{1-\sigma}\left(l_j\left(\theta\right)\right)^{-\sigma}e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2}}=\Phi\left(l_j\left(\theta\right)\right)^{\eta-1}$$

which leads to the following effort decision

$$l_j^*\left(\theta\right)=\left(\frac{(1-\tau_j)\theta e^{-\frac{\sigma\varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{1-\sigma}{(\sigma+\eta-1)}}. \tag{51}$$

The income earned by an agent θ is now such that

$$y_j^*\left(\theta\right)=\theta l_j^*\left(\theta\right)\varepsilon=(1-\tau_j)^{\frac{1-\sigma}{\sigma+\eta-1}}\left(\frac{\theta}{\Phi^{\frac{1}{\eta}}}\right)^{\frac{\eta}{\sigma+\eta-1}}e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2(\sigma+\eta-1)}}\varepsilon, \tag{52}$$

and its expected income can be expressed as follows

$$\mathbb{E}[y_j^*\left(\theta\right)]=\theta l_j^*\left(\theta\right)\mathbb{E}[\varepsilon]=(1-\tau_j)^{\frac{1-\sigma}{\sigma+\eta-1}}\left(\frac{\theta}{\Phi^{\frac{1}{\eta}}}\right)^{\frac{\eta}{\sigma+\eta-1}}e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2(\sigma+\eta-1)}}. \tag{53}$$

The expected median income of community j is

$$\mathbb{E}[y_j^*\left(\theta^m\right)]=\theta^m l_j^*(\theta^m)=(1-\tau_j)^{\frac{1-\sigma}{\sigma+\eta-1}}\left(\frac{\theta}{\Phi^{\frac{1}{\eta}}}\right)^{\frac{\eta}{\sigma+\eta-1}}e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2(\sigma+\eta-1)}}. \tag{54}$$

The community j average income becomes

$$\mu_j^*=\left(\frac{(1-\tau_j)}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{1-\sigma}{\sigma+\eta-1}}e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2(\sigma+\eta-1)}}\left(\frac{\int\limits_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\int_\varepsilon\left(\theta\right)^{\frac{\eta}{\sigma+\eta-1}}\varepsilon f\left(\varepsilon\right)d\varepsilon\right)h\left(\theta\right)d\theta}{n_j}\right). \tag{55}$$

Using the independence of θ and ε and the law of large numbers, we can deduce that

$$\begin{aligned}\frac{\int\limits_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\int_\varepsilon\left(\theta\right)^{\frac{\eta}{\sigma+\eta-1}}\varepsilon f\left(\varepsilon\right)d\varepsilon\right)h\left(\theta\right)d\theta}{n_j}&=\left(\frac{\int\limits_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\theta\right)^{\frac{\eta}{\sigma+\eta-1}}h\left(\theta\right)d\theta}{n_j}\right)\left(\int_\varepsilon\varepsilon f\left(\varepsilon\right)d\varepsilon\right)\\&=\left(\frac{Int_j\theta}{n_j}\right)\mathbb{E}[\varepsilon]=\left(\frac{Int_j\theta}{n_j}\right),\end{aligned}$$

where

$$Int_j\theta\equiv\int\limits_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\theta\right)^{\frac{\eta}{\sigma+\eta-1}}h\left(\theta\right)d\theta$$

hence, (55) can be written as follows:

$$\mu_j^* = \left(\frac{(1-\tau_j)\, e^{-\frac{\sigma \varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}} \right)^{\frac{1-\sigma}{\sigma+\eta-1}} \frac{Int_j \theta}{n_j}. \tag{56}$$

(ii) For any $j = 1, ..., J$, the tax rate is chosen by the median voter as follows:

$$\tau_j = \arg\max_{\tau_j} \mathbb{E}\left[\frac{((1-\tau_j)\theta^m l_j(\theta^m)\varepsilon)^{1-\sigma}}{1-\sigma} \right] + \beta \mathbb{E}\left[\frac{\left((n_j)^{1-\nu} \tau_j \mu_j^* \right)^{1-\sigma}}{1-\sigma} \right] - \frac{1}{\eta}(l_j(\theta^m))^\eta .$$

Applying the law of large numbers yields

$$\tau_j = \arg\max_{\tau_j} \left\{ \frac{((1-\tau_j)\theta^m l_j(\theta^m))^{1-\sigma} \mathbb{E}\left[\varepsilon^{1-\sigma}\right]}{1-\sigma} + \beta \frac{\left((n_j)^{1-\nu} \tau_j \mu_j^* \right)^{1-\sigma}}{1-\sigma} \right\}.$$

From (50) the following maximization program can be written as follows

$$\max_{\tau_j} \left\{ \frac{((1-\tau_j)\theta^m l_j(\theta^m))^{1-\sigma} e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2}}}{1-\sigma} + \beta \frac{\left((n_j)^{1-\nu} \tau_j \mu_j^* \right)^{1-\sigma}}{1-\sigma} \right\}.$$

The first-order condition gives

$$(1-\tau_j))^{-\sigma} (\theta^m l_j(\theta^m))^{1-\sigma} e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2}} = \beta \left((n_j)^{1-\nu} \mu_j^* \right)^{1-\sigma} (\tau_j)^{-\sigma},$$

which is equivalent to

$$\tau_j = \frac{\beta^{\frac{1}{\sigma}} \left(\frac{\theta^m l_j(\theta^m)}{(n_j)^{1-\nu} \mu_j^*} \right)^{\frac{\sigma-1}{\sigma}} e^{\frac{(1-\sigma)\varrho_\varepsilon^2}{2}}}{1 + \beta^{\frac{1}{\sigma}} \left(\frac{\theta^m l_j(\theta^m)}{(n_j)^{1-\nu} \mu_j^*} \right)^{\frac{\sigma-1}{\sigma}} e^{\frac{(1-\sigma)\varrho_\varepsilon^2}{2}}}$$

which can be written as follows

$$\tau_j = \frac{1}{1 + \beta^{-\frac{1}{\sigma}} \left(\frac{(n_j)^{1-\nu} \mu_j^*}{\theta^m l_j(\theta^m)} \right)^{\frac{\sigma-1}{\sigma}} e^{\frac{(\sigma-1)\varrho_\varepsilon^2}{2}}}. \tag{57}$$

Using the fact that from equations (56) and (54) we have

$$\frac{\mu_j^*(\theta)}{\theta^m l_j^*(\theta^m)} = \frac{\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j} \left(\frac{\theta}{\theta_j^m} \right)^{\frac{\eta}{\sigma+\eta-1}} h(\theta)\, d\theta}{n_j}$$

leads to

$$\tau_j^* = \frac{1}{1+\beta^{-\frac{1}{\sigma}}\left(\left(n_j\right)^{1-\epsilon}\frac{\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}\left(\frac{\theta}{\theta_j^m}\right)^{\frac{\eta}{\sigma+\eta-1}}h(\theta)d\theta}{n_j}\right)^{\frac{\sigma-1}{\sigma}} e^{\frac{(\sigma-1)\varrho_\varepsilon^2}{2}}}. \tag{58}$$

(iii) Boundary individuals $\widetilde{\theta}_j$ are defined by

$$\mathbb{E}\left[V_j^*(\widetilde{\theta}_j)\right] - \mathbb{E}\left[V_{j+1}^*(\widetilde{\theta}_j)\right] = 0, \text{ for } j = 1, ..., J-1 \tag{59}$$

with

$$\mathbb{E}\left[V_j^*(\widetilde{\theta}_j)\right] \equiv \mathbb{E}\left[\frac{\left((1-\tau_j^*)\widetilde{\theta}_j l_j^*\left(\widetilde{\theta}_j\right)\varepsilon\right)^{1-\sigma}}{1-\sigma}\right] + \beta\frac{\left(\tau_j^*\left(n_j^*\right)^{1-\nu}\mu_j^*\right)^{1-\sigma}}{1-\sigma} - \frac{1}{\eta}\left(l_j^*\left(\widetilde{\theta}_j\right)\right)^{\eta}$$

Replacing (50), (56) and (51) in this equation we find for any individual θ in community j

$$\begin{aligned}\mathbb{E}\left[V_j^*(\theta)\right] \equiv\ & \frac{(\sigma+\eta-1)}{(1-\sigma)\eta}\left(\left(1-\tau_j^*\right)\theta\right)^{\frac{\eta(1-\sigma)}{\sigma+\eta-1}} e^{-\frac{\sigma(1-\sigma)\eta\varrho^2}{2(\sigma+\eta-1)}} \\ & + \frac{\beta}{1-\sigma}\left(\tau_j^*\left(n_j^*\right)^{1-\nu}\left(1-\tau_j^*\right)^{\frac{1-\sigma}{\sigma+\eta-1}} e^{-\frac{\sigma(1-\sigma)\varrho^2}{2(\sigma+\eta-1)}}\left(\frac{\int_{\widetilde{\theta}_{j-1}}^{\widetilde{\theta}_j}(\theta)^{\frac{\eta}{\sigma+\eta-1}}h(\theta)\,d\theta}{n_j}\right)\right)^{1-\sigma}.\end{aligned}$$

Let us first consider that effort is exogenous and $l_j(\theta) = \bar{l} = 1$. We then have

$$y_j^*(\theta) = y^*(\theta) = \theta\varepsilon.$$

The whole income distribution is independent of the equilibrium partition. The log of the income of an agent is equal to

$$\ln y^{exo}(\theta) = \ln y_j^*(\theta) = \ln\theta + \ln\varepsilon.$$

Then, the variance is given by

$$var\left[\ln y^{exo}(\theta)\right] = \varrho_\theta^2 + \varrho_\varepsilon^2. \tag{60}$$

It turns out that in such a case, the income variance does not depend on the equilibrium set of jurisdictions. Although jurisdictions can be ranked according to their mean income due to the income sorting result, there will be overlapping between the income distributions of two subsequent jurisiditions. This is due to a transitory income shock that occurs after the formation of jurisdictions (see Figure A5).

Let us now endogenize effort. Applying the ln operator to both sides of the equation (52), we obtain:

$$\ln y_j^*\left(\theta,\varepsilon\right) = -\frac{\sigma\left(1-\sigma\right)\varrho_\varepsilon^2}{2\left(\sigma+\eta-1\right)} + \frac{1-\sigma}{\sigma+\eta-1}\ln\left(1-\tau_j\right) + \frac{\eta}{\sigma+\eta-1}\ln\theta - \frac{1}{\sigma+\eta-1}\ln\Phi + \ln\varepsilon. \quad (61)$$

Now the income of an agent is jurisdiction-dependent as it depends on $\ln\left(1-\tau_j\right)$. It then turns out that the whole income distribution is a function of the equilibrium pattern of jurisdictions.

The variance of the whole income distribution then equals

$$\begin{aligned} var\left[\ln y^*\left(\theta,\varepsilon\right)\right] &= \left(\frac{1-\sigma}{\eta+\sigma-1}\right)^2 var\left[\ln\left(1-\tau_j\right)\right] + \left(\frac{\eta}{\eta+\sigma-1}\right)^2 \varrho_\theta^2 + \varrho_\varepsilon^2 \\ &+ 2\frac{\left(1-\sigma\right)\eta}{\left(\sigma+\eta-1\right)^2}\sum_{j=1,\theta\in j}^{N} cov(\ln\left(1-\tau_j\right),\ln\theta). \end{aligned} \quad (62)$$

We can thus see whether the possibility for individuals to exert effort may increase inequality in this setup. From (60) and (62), we have:

$$\begin{aligned} var\left[\ln y^*\left(\theta,\varepsilon\right)\right] - var\left[\ln y^{exo}\left(\theta\right)\right] &= \left(\frac{1-\sigma}{\eta+\sigma-1}\right)^2 var\ln\left(1-\tau_j\right) \\ &+ \left(\left(\frac{\eta}{\eta+\sigma-1}\right)^2-1\right)\varrho_\theta^2 + 2\frac{\left(1-\sigma\right)\eta}{\left(\sigma+\eta-1\right)^2}\sum_{j=1,\theta\in j}^{N} cov(\ln\left(1-\tau_j\right),\ln\theta) \end{aligned}$$

which has an ambiguous. If for instance $\sigma<1$ then $\left(\left(\frac{\eta}{\eta+\sigma-1}\right)^2-1\right)<0$ but it is possible that the local taxation component $\left(\frac{1-\sigma}{\eta+\sigma-1}\right)^2 var\ln\left(1-\tau_j\right)$ and the covariance component may counterbalance this effect.

Proof of Proposition 6. We can also compare inequality characterizing the centralized equilibrium with inequality in the decentralized case. Given (62), we have

$$\begin{aligned} var\left[\ln y^*\left(\theta,\varepsilon\right)\right]_{J>1} - var\left[\ln y^*\left(\theta,\varepsilon\right)\right]_{J=1} &= \left(\frac{1-\sigma}{\eta+\sigma-1}\right)^2 var\ln\left(1-\tau_j\right) \\ &+ 2\frac{\left(1-\sigma\right)\eta}{\left(\sigma+\eta-1\right)^2}\sum_{j=1,\theta\in j}^{N} cov(\ln\left(1-\tau_j\right),\ln\theta). \end{aligned}$$

Given Proposition 3, we have

$$\sigma \gtrless 1 \Longrightarrow cov(\ln\left(1-\tau_j\right),\ln\theta) \lesseqgtr 0$$

implying that

$$\frac{(1-\sigma)\,\eta}{(\sigma+\eta-1)^2}\sum_{j=1,\theta\in j}^{N} cov(\ln(1-\tau_j)\,,\ln\theta)>0$$

and leading to

$$var\left[\ln y^*\left(\theta,\varepsilon\right)\right]_{J>1}-var\left[\ln y^*\left(\theta,\varepsilon\right)\right]_{J=1}>0.$$

7.10.2 Numerical Values

The average income of the society is equal to μ^a and the median income is μ^m. If the income follows a lognormal distribution, we have

$$\mu^a=e^{\mathbb{E}[\ln y(\theta,\varepsilon)]+\frac{1}{2}(var(\ln y(\theta,\varepsilon)))}$$

equivalently,

$$\ln\mu^a=\mathbb{E}\left[\ln y\left(\theta,\varepsilon\right)\right]+\frac{1}{2}\left(var\left(\ln y\left(\theta,\varepsilon\right)\right)\right). \tag{63}$$

While median income equals:

$$\mu^m=e^{\mathbb{E}[\ln y(\theta,\varepsilon)]}\Rightarrow\ln\mu^m=\mathbb{E}\left[\ln y\left(\theta,\varepsilon\right)\right]. \tag{64}$$

Using (63) and (64) we can deduce the following equation

$$var\left(\ln y\left(\theta,\varepsilon\right)\right)=2\ln\frac{\mu^a}{\mu^m}. \tag{65}$$

In order to set ϱ_ε^2, we consider that ε can be interpreted as transitory income and use the ratio of transitory to total earnings in U.S. calculated by Moffitt and Gottschalk (2001). Denoting by ϑ the ratio of the variance of $\ln\varepsilon$ to the variance of $\ln y\left(\theta,\varepsilon\right)$, Moffitt and Gottschalk (2001) find that

$$\vartheta=0.5.$$

Hence it is expressed as follows

$$\vartheta=\frac{\varrho_\varepsilon^2}{var\ln y\left(\theta,\varepsilon\right)} \tag{66}$$

we deduce that

$$\varrho_\varepsilon=\left(\vartheta var\left(\ln y\right)\right)^{\frac{1}{2}}=0.6267,$$

and using the properties of lognormal distribution and given that $\mathbb{E}\left[\varepsilon\right]=1$ we deduce that

$$\psi=-\frac{\varrho_\varepsilon^2}{2}=-0.1963$$

and

$$\varsigma_y=\frac{\sigma+\eta-1}{\eta}\left(\left(1-\vartheta\right)*var\left(\ln y\right)\right)^{\frac{1}{2}}=0.6643,$$

where ς_y is the standard deviation of the log of income.

Parameters Φ and β are calibrated to obtain an average effort equal to 1, when there is only one community.

From equation (51), we obtain the average labor supply, denoted by $\widehat{l^a}(\theta)$

$$\widehat{l^a}(\theta) = \left(\frac{(1-\tau)e^{-\frac{\sigma\varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{1-\sigma}{(\sigma+\eta-1)}} \int_{\underline{\theta}}^{\overline{\theta}} (\theta)^{\frac{1-\sigma}{(\sigma+\eta-1)}} h(\theta)\, d\theta$$

It is normalized to one. Hence,

$$1 = \left(\frac{(1-\tau)\, e^{-\frac{\sigma\varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{(1-\sigma)}{\sigma+\eta-1}} \int_{\underline{\theta}}^{\overline{\theta}} \theta^{\frac{1-\sigma}{\sigma+\eta-1}} h(\theta)\, d\theta$$

which is equivalent to

$$1 = \left(\frac{(1-\tau)\, e^{-\frac{\sigma\varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{(1-\sigma)}{\sigma+\eta-1}} e^{\frac{(1-\sigma)}{\sigma+\eta-1}\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right)}$$

Isolating Φ in this equation leads to

$$\Phi = (1-\tau)^{1-\sigma}\, e^{-\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2}}\, e^{(1-\sigma)\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right)}. \tag{67}$$

Applying the ln function yields

$$\begin{aligned}\ln\Phi = {} & (1-\sigma)\ln(1-\tau) \\ & -\frac{\sigma(1-\sigma)\varrho_\varepsilon^2}{2} \\ & +(1-\sigma)\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right).\end{aligned} \tag{68}$$

In order to give a value to Φ, we need to compute ϱ_ε^2, ϱ_θ^2, $\widehat{\theta}$.

The equation (56) implies that if we are in a centralized economy, then

$$\mu^a = \left(\frac{(1-\tau)\, e^{-\frac{\sigma\varrho_\varepsilon^2}{2}}}{\Phi^{\frac{1}{1-\sigma}}}\right)^{\frac{1-\sigma}{\sigma+\eta-1}} \left(\int_{\underline{\theta}}^{\overline{\theta}} (\theta)^{\frac{\eta}{\sigma+\eta-1}} h(\theta)\, d\theta\right). \tag{69}$$

Replacing (67) in this equation leads to

$$\mu^a = \frac{\int_{\underline{\theta}}^{\overline{\theta}} (\theta)^{\frac{\eta}{\sigma+\eta-1}} h(\theta)\, d\theta}{e^{\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right)\frac{1-\sigma}{\sigma+\eta-1}}}$$

which is equivalent to

$$\mu^a = \frac{e^{\frac{\eta}{\sigma+\eta-1}\left(\widehat{\theta}+\frac{\eta}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right)}}{e^{\frac{(1-\sigma)}{\sigma+\eta-1}\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right)}}$$

hence,

$$\mu^a = e^{\widehat{\theta}+\left(\eta^2-(1-\sigma)^2\right)\frac{\varrho_\theta^2}{2(\sigma+\eta-1)^2}}$$

and,

$$\mu^a = e^{\widehat{\theta}+\frac{(\eta+1-\sigma)\varrho_\theta^2}{2(\sigma+\eta-1)}}.$$

When there is only one community, equation (61) can be written as follows

$$\begin{aligned}\ln y\left(\theta,\varepsilon\right) = {} & \frac{1-\sigma}{\sigma+\eta-1}\ln\left(1-\tau\right) \\ & + \frac{\eta}{\sigma+\eta-1}\ln\left(\theta\right) - \frac{1}{\sigma+\eta-1}\ln\left(\Phi\right) \\ & - \frac{\sigma\left(1-\sigma\right)\varrho_\varepsilon^2}{2\left(\sigma+\eta-1\right)} + \ln\varepsilon.\end{aligned}$$

Replacing (68) in this equation, we obtain

$$\begin{aligned}\ln y\left(\theta,\varepsilon\right) = {} & \frac{\eta}{\sigma+\eta-1}\ln\left(\theta\right) \\ & - \frac{(1-\sigma)}{\sigma+\eta-1}\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right) \\ & + \ln\varepsilon.\end{aligned}$$

From this equation, we deduce the expected value of the logarithm of income:

$$\begin{aligned}\mathbb{E}\left[\ln y\left(\theta,\varepsilon\right)\right] = {} & \frac{\eta}{\sigma+\eta-1}\widehat{\theta} \\ & - \frac{(1-\sigma)}{\sigma+\eta-1}\left(\widehat{\theta}+\frac{(1-\sigma)}{\sigma+\eta-1}\frac{\varrho_\theta^2}{2}\right) \\ & - \frac{\varrho_\varepsilon^2}{2}\end{aligned}$$

which is equivalent to

$$\mathbb{E}\left[\ln y\left(\theta,\varepsilon\right)\right] = \widehat{\theta} - \left(\frac{(1-\sigma)}{\sigma+\eta-1}\right)^2\frac{\varrho_\theta^2}{2} - \frac{\varrho_\varepsilon^2}{2}. \tag{70}$$

The standard deviation equals

$$var\ln y\left(\theta,\varepsilon\right) = \left(\frac{\eta}{\sigma+\eta-1}\right)^2\varrho_\theta^2 + \varrho_\varepsilon^2. \tag{71}$$

Replacing (65) in (66) leads to

$$\varrho_\varepsilon^2 = 2\vartheta\ln\left(\frac{\mu^a}{\mu^m}\right). \tag{72}$$

From equations (65) and (71), the standard deviation can be written as follows

$$var\ln y(\theta,\varepsilon)=\left(\frac{\eta}{\sigma+\eta-1}\right)^2\varrho_\theta^2+\vartheta\left(var\ln y(\theta,\varepsilon)\right)$$

implying that

$$\varrho_\theta^2=\frac{(1-\vartheta)\left(var\ln y(\theta,\varepsilon)\right)}{\left(\frac{\eta}{\sigma+\eta-1}\right)^2}$$

given (66) we have

$$\varrho_\theta^2=\frac{2(1-\vartheta)}{\left(\frac{\eta}{\sigma+\eta-1}\right)^2}\ln\left(\frac{\mu^a}{\mu^m}\right). \tag{73}$$

From (70), we have

$$\widehat{\theta}=\mathbb{E}\left[\ln y(\theta,\varepsilon)\right]+\frac{\varrho_\varepsilon^2}{2}+\left(\frac{1-\sigma}{\sigma+\eta-1}\right)^2\frac{\varrho_\theta^2}{2}$$

and from (64), (72) and (73) we obtain

$$\widehat{\theta}=\ln\mu^m+\left(\vartheta+(1-\vartheta)\left(\frac{1-\sigma}{\eta}\right)^2\right)\ln\left(\frac{\mu^a}{\mu^m}\right). \tag{74}$$

Integrating in (68) expressions in (72), (73) and (74) imply

$$\begin{aligned}\frac{\ln\Phi}{(1-\sigma)}&=\ln(1-\tau)+\ln\mu^m\\&-\sigma\vartheta\left(\ln\frac{\mu^a}{\mu^m}\right)\\&+\left(\vartheta+(1-\vartheta)\left(\frac{1-\sigma}{\eta}\right)^2\right)\ln\frac{\mu^a}{\mu^m}\\&+(1-\vartheta)\left(\frac{1-\sigma}{\eta}\right)^2\frac{(\sigma+\eta-1)}{(1-\sigma)}\ln\frac{\mu^a}{\mu^m}\end{aligned}$$

leading to

$$\begin{aligned}\ln\Phi&=(1-\sigma)\ln(1-\tau)\,\mu^m\\&+\left(\vartheta+\frac{(1-\vartheta)}{\eta}\right)(1-\sigma)^2\ln\frac{\mu^a}{\mu^m}.\end{aligned} \tag{75}$$

Using (75) to give a value to the Φ parameter, we obtain $\Phi=0.5493$.

When there is only one community, (58) becomes

$$\tau=\frac{1}{1+\beta^{-\frac{1}{\sigma}}\left((\theta^m)^{-\frac{\eta}{\sigma+\eta-1}}\int\limits_{\underline{\theta}}^{\overline{\theta}}(\theta)^{\frac{\eta}{\sigma+\eta-1}}h(\theta)\,d\theta\right)^{\frac{\sigma-1}{\sigma}}e^{\frac{(\sigma-1)\varrho_\varepsilon^2}{2}}} \tag{76}$$

Applying the properties of the lognormal distribution, we have

$$\tau = \frac{1}{1+\beta^{-\frac{1}{\sigma}}\left(e^{-\frac{\eta}{\sigma+\eta-1}\ln\theta^m+\frac{\eta}{\sigma+\eta-1}\left(\mathbb{E}[\ln\theta]+\frac{\eta}{2(\sigma+\eta-1)}var(\ln\theta)\right)}\right)^{\frac{\sigma-1}{\sigma}}e^{\frac{(\sigma-1)\varrho_{\varepsilon}^2}{2}}}$$

which can be simplified as follows

$$\tau = \frac{1}{1+\beta^{-\frac{1}{\sigma}}e^{\frac{(\sigma-1)}{2\sigma}\left(\left(\frac{\eta}{\sigma+\eta-1}\right)^2\varrho_\theta^2+\sigma\varrho_\varepsilon^2\right)}}$$

leading to the following expression of β

$$\beta = \left(\frac{\tau}{1-\tau}\right)^{\sigma}e^{\frac{(\sigma-1)}{2}\left(\left(\frac{\eta}{\sigma+\eta-1}\right)^2\varrho_\theta^2+\sigma\varrho_\varepsilon^2\right)}. \tag{77}$$

Given (77), we obtain

$$\beta = \left(\frac{\tau^{cen}}{1-\tau^{cen}}\right)^{\sigma}e^{\frac{(\sigma-1)}{2}\left(\left(\frac{\eta}{\sigma+\eta-1}\right)^2\varrho_\theta^2+\sigma\varrho_\varepsilon^2\right)}.$$

7.10.3 Simulation Results

We focus on the following endogenous variables: n_j, $\widetilde{\theta}_j$, μ_j, τ_j, q_j, $l_j(\theta)$, $y_j(\theta)$, the level of individual utility $V_j(\theta_j)$ and the level of individual utility measured in equivalent consumption units, denoted by $V_j^{eq}(\theta_j)$. We also consider aggregate variables such as population average income, denoted by μ^a, total welfare, denoted by V^a, total utility in equivalent consumption units, denoted by V^{eq}. We finally examine some inequality measures such as $var\left[\ln y^*(\theta,\varepsilon)\right]$, the Lorenz curve of $\ln y^*(\theta,\varepsilon)$, the share of total inequality generated by the difference in the average income between communities used by Calabrese, Epple and Romano (2012), denoted by SHARE-INEQ, $var\left[\ln V^{eq}\right]$ and the Lorenz curve V^{eq}. Variables μ^a, V^a, V^{eq} and SHARE-INEQ are defined as follows:

$$\begin{aligned}
\mu^a &= \int_{\underline{\theta}}^{\overline{\theta}}\int_0^{\infty} y_j(\theta,\varepsilon)\,f(\varepsilon)\,h(\theta)\,d\varepsilon d\theta,\\
V^a &= \int_{\underline{\theta}}^{\overline{\theta}} V_j(\theta_j)h(\theta)\,d\theta,\\
V^{eq} &= \int_{\underline{\theta}}^{\overline{\theta}} V_j^{eq}(\theta_j)h(\theta)\,d\theta \text{ with } V_j^{eq}(\theta_j) \text{ such that } V_j(\theta_j) = \frac{\left(V_j^{eq}(\theta_j)\right)^{1-\sigma}}{1-\sigma}\\
SHARE-INEQ &= \frac{\sum_j n_j\left(\mu_j-\mu^a\right)^2}{var\left(\ln y(\theta,\varepsilon)\right)}.
\end{aligned}$$

The following Tables describe the simulations results for the centralized case, the 2, 3, 4 or 5 jurisdictions cases:

	Size					Boundary Individual				
J	n_1	n_2	n_3	n_4	n_5	$\widetilde{\theta}_1$	$\widetilde{\theta}_2$	$\widetilde{\theta}_3$	$\widetilde{\theta}_4$	$\widetilde{\theta}_5$
1	1					726.9035				
2	0.7899	0.2101				20.0833	726.9035			
3	0.6420	0.2437	0.1143			10.6136	35.6171	726.9035		
4	0.5474	0.2358	0.1429	0.0739		7.4563	19.4301	50.5655	726.9035	
5	0.4816	0.2225	0.1466	0.0966	0.0526	5.8750	13.6082	28.1145	64.7511	726.9035

Table A1: Jurisdictions' size and boundary individuals.

	Jurisdiction' Average Income					Local tax				
J	μ_1	μ_2	μ_3	μ_4	μ_5	τ_1	τ_2	τ_3	τ_4	τ_5
1						0.3004				
2	4.1330	13.9563				0.3517	0.5338			
3	3.6374	9.0010	19.0111			0.3853	0.5258	0.6250		
4	3.3716	7.4700	12.2676	23.6566		0.4102	0.5325	0.6051	0.6855	
5	3.2002	6.6702	10.0816	15.2717	28.0105	0.4300	0.5418	0.6027	0.6600	0.7285

Table A2: Jurisdictions' average income and local tax rates.

	Local Public Good				
J	q_1	q_2	q_3	q_4	q_5
1	1.6390				
2	1.1483	1.5651			
3	0.8999	1.1532	1.3583		
4	0.7571	0.9381	1.0607	1.1977	
5	0.6628	0.804	1 08911	0.9740	1.0737

Table A3: Local public goods.

J	μ^a	V^a	V^{eq}
1	5.4567	-0.9931	1.1193
2	6.1965	-1.1086	1.0034
3	6.7016	-1.2138	0.9164
4	7.1074	-1.3088	0.8510
5	7.4532	-1.3968	0.7992

Table A4: Average income μ^a, total welfare V^a and total welfare in equivalent consumption units, V^{eq}.

J	$\mathrm{Var}(\log y)$	$\mathrm{Var}\mu/\mathrm{Var}y$	$\mathrm{Var}(\log V^{eq})$
1	0.7818		0.2876
2	0.8541	0.2927	0.2596
3	0.8911	0.3598	0.2362
4	0.9167	0.3925	0.2178
5	0.9351	0.4136	0.2028

Table A5: Inequality measures.

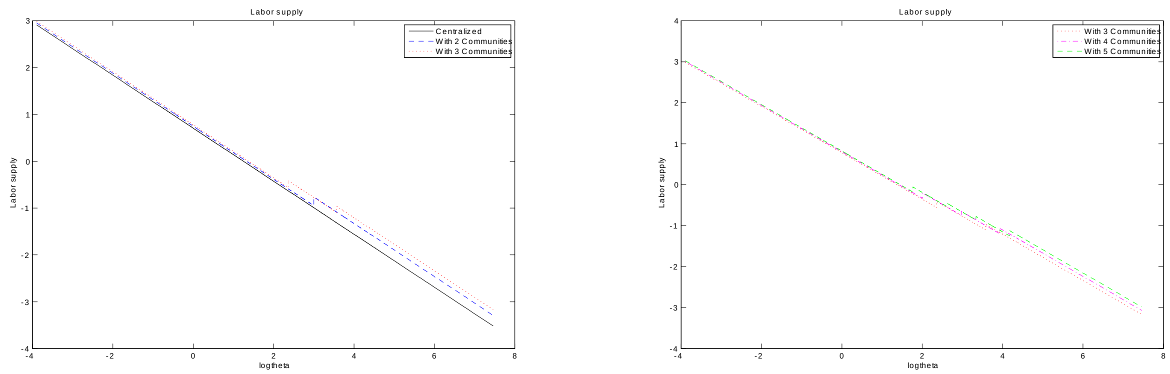

Figure A1: Equilibrium individual labor supply with respect to productivity θ, for $J = 1, 2, 3$ (left) and $J = 3, 4, 5$ (right).

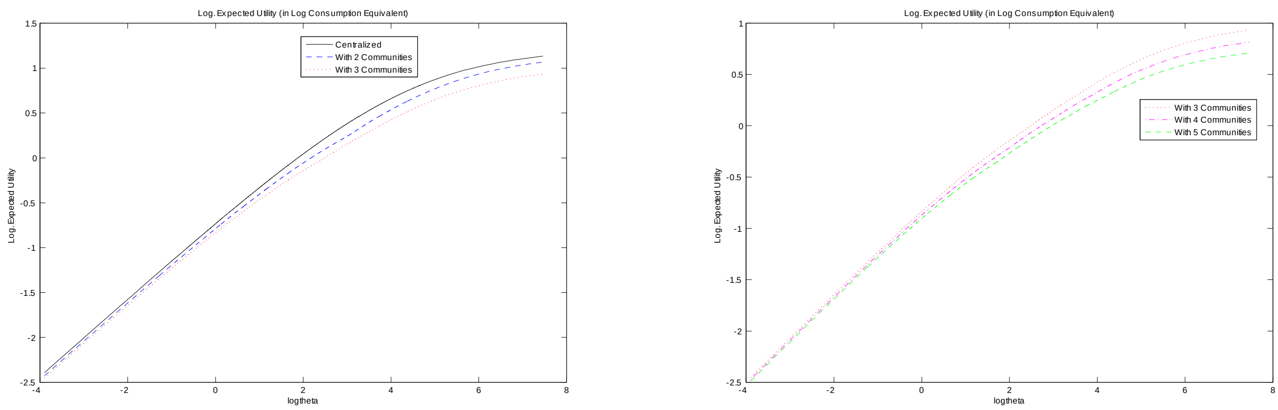

Figure A2: Level of individual utility in equivalent consumption units, for $J = 1, 2, 3$ (left) and $J = 3, 4, 5$ (right).

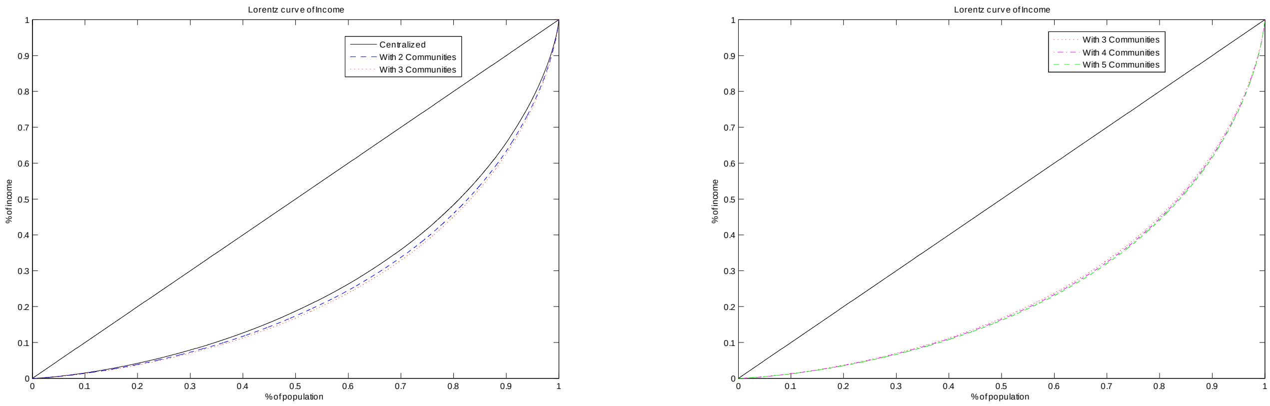

Figure A3: Lorenz curves of income distribution for $J = 1, 2, 3$ (left) and $J = 3, 4, 5$ (right).

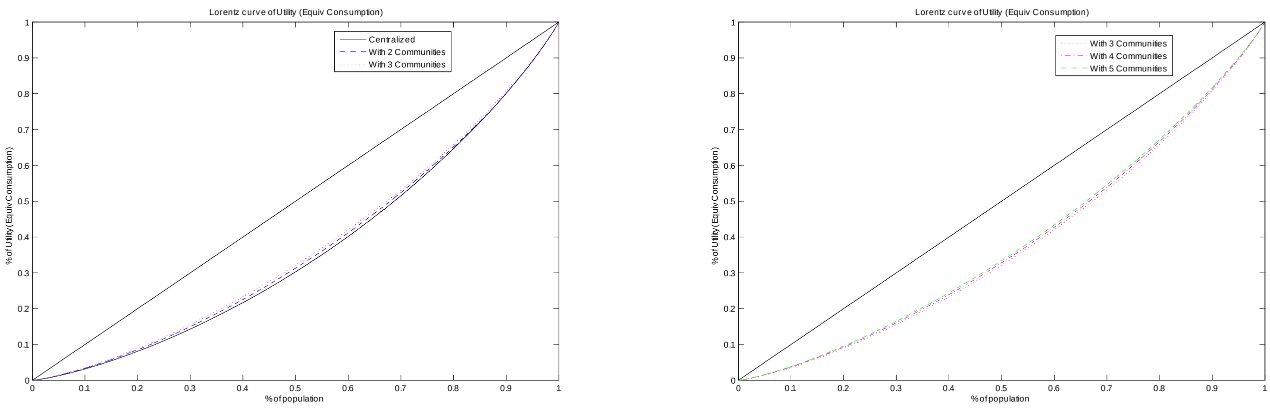

Figure A4: Lorenz curves of V^{eq} for $J = 1, 2, 3$ (left) and $J = 3, 4, 5$ (right).

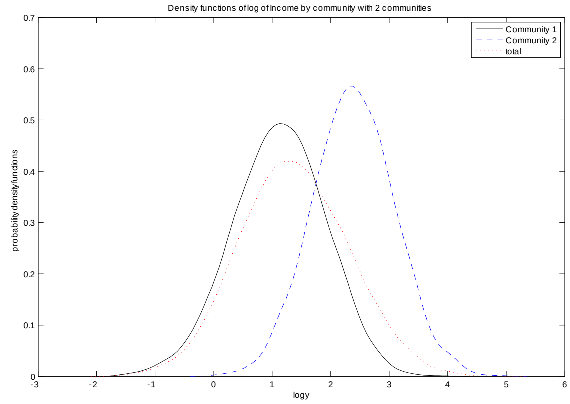

Figure A5: Overlapping of income distributions.