Group Formation, Free Mobility and Social Optimality

Antoine LOEPER[1]

December 2009

[1]Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, a-loeper@kellogg.northwestern.edu

Abstract

We consider a game in which a continuum of heterogeneous individuals partition themselves into communities. Communities can be of any size and group externalities can exhibit economies of scale or congestion effects. When group externalities are anonymous, we show that free mobility equilibria are socially optimal when group externalities increase or decrease logarithmically with community size. When they increase less (resp. more) than logarithmically, the equilibrium exhibits excessive agglomeration (resp. excessive fragmentation). These results hold irrespective of the distribution of preferences and the set of available communities. When group externalities are not anonymous, the optimality of free mobility equilibria requires an additional symmetry condition. We characterize conditions under which free mobility equilibria are excessively or insufficiently segregated. We apply these results to local public goods economies and translate the logarithmic and the symmetry conditions into conditions on the public good technology and the progressivity of taxation.

Keywords: Group Formation, Economies of Scale, Congestion, Local Public Goods Economy.

1 Introduction

Numerous economic and social activities are conducted within groups. The distinctive feature of these social communities is that the benefits they generate to their members depend both on the characteristics of the community (their internal rules or the services they provides to their members) and on their membership. Larger groups exploit economies of scale, indivisibilities, network effects or risk sharing while smaller groups minimize congestion effects and can better match the preferences of their members. For this reason, the size of these communities depends both on the distribution of preferences and the group externalities they generate. National defense is typically provided at the national level while garbage collection can be done at a more local level. Technologies with little scope for customization and strong network effects have millions of users while some software are designed for a single firm. Likewise, the size of religious communities varies widely among faiths.

Ideally, a partition of society into communities should balance the cost of heterogeneous groups with the benefits of larger groups. However, in practice, group membership is determined by individual decisions. Freedom to migrate is constitutionally guaranteed in many developed countries. The central question of this paper is whether free mobility leads to a partition of society into communities of optimal size and composition.

To do so, we analyze the Nash equilibria of a group formation game with a continuum of players in which each player's strategy is the community she chooses to join. The set of members of a community with characteristics c is simply the set of individuals whose strategy is c. Individuals have idiosyncratic preferences over community characteristics. The latter can be interpreted as the tax and spending scheme of a local jurisdiction, the legal code of a state, the beliefs of a religious community or the properties of a technology. The welfare of a member of a community depends also on the set of individuals who choose the same community. In the baseline model, we assume that these group externalities are anonymous but we do not make any hypothesis about their shape. In particular, our results apply both to the case of economies of scale and congestion externalities.

Our main contribution is to provide conditions on the shape of group externalities under which Nash equilibria are socially optimal. We first show that when group externalities increase (or decrease) logarithmically in group size, a socially optimal Nash equilibrium exists. We then prove that the logarithmic condition is essentially necessary and characterize the nature of the inefficiency when group externalities are non-logarithmic. If they increase "more than logarithmically" with group size, free mobility leads to too much fragmentation while if they increase "less than logarithmically", it generates excessive agglomeration. These results hold irrespective of the distribution of preferences and the set of available communities. Hence, our model suggests that it is the shape of economies of scale or congestion costs rather than the distribution of preferences which determines the social cost of free mobility. When applied to local public good economies, the logarithmic condition can be readily transposed into conditions on the public good technology and the tax-spending scheme.

The intuition behind the main result is a simple Pigouvian argument. When an individual leaves a group to join another one, she imposes two kinds of externality: an *emigration externality* on the members of the community she leaves, and an *immigration externality* on the members of the community she joins. When returns from group size are logarithmic, these two effects cancel out so migration has no aggregate externality. When returns are more than logarithmic, then the gap between the private cost and the social cost of emigration is increasing in the size of the origin community. Hence, free mobility will lead too much migration from large groups to small groups and the equilibrium will be too fragmented.

We then consider non-anonymous group externalities and show that the optimality of free mobility requires an additional symmetry condition: in each community, the benefit for the members of type a of an additional member of type a must be equal to the benefit for the members of type a of an additional member of type a. Furthermore, we show that when group externalities exhibit “more than logarithmic” (resp. less) homophily, then the Nash equilibria are insufficently (resp. excessively) segregated.

The paper is organized as follows. Section 2 discusses the literature. Section 3 lays out the basic model. Section 4 derives some preliminary lemmas. Section 5 considers the case of anonymous group effects and section 6 extends the model to heterogeneous externalities. In section 7, we endogenize the group policy. Section 8 concludes.

2 Related Literature

In his seminal paper, Tiebout (1956) argued that by voting with their feet, citizens reveal truthfully their preferences for public goods and form groups of optimal size. To show this insight formally, the literature on club theory and free mobility makes assumptions which basically amount to assume away any scarcity or mismatch in the individual allocation problem. Roughly speaking, if s is the optimal size of a group that some individual i prefers, then s is finite and there is always $s-1$ unmatched individuals with identical preferences willing form a group with i. (Bewley 1981) Our model departs from the club theory literature in that we do not assume that group size is negligible or that there is as many clubs as types of individuals.

The way we model the group formation process is similar to Konishi, Lebreton and Weber (1997a, 1997b, 1998) and Milchtaich (1996). These papers study the existence of Nash equilibria with a finite population and either economies of scale or congestion externalities. Our specification encompasses both economies of scale and congestion effects and avoids the existence problem by assuming a continuum of individuals. We focus instead on the welfare analysis.

The cooperative game theory literature (Greenberg and Weber 1986, 1993, Demange 1994, Haimanko, Lebreton and Weber 2004) has analyzed the trade-off between economies of scale and the cost of heterogeneity in large groups. These papers identify restrictions on the distribution of idiosyncratic preferences which guarantee that the group formation game has a non-empty core (or some variant of it). Our model focuses on individual mobility and shows that what matters for social welfare is the shape of group externalities.

Finally, a few papers have analyzed the welfare consequences of various rules of secession, integration and immigration. It may be socially optimal to restrict mobility by requiring the unanimous consent of the destination jurisdiction (Jehiel and Scotchmer 2001). When secessions are decided by majority rule, they can lead to inefficiently small jurisdictions (Bolton and Roland 1997, Alesina and Spolaore 1997) . The rationale is that the pivotal voter of a seceding group does not internalize the diseconomies of scale she imposes on the jurisdiction she is seceding from. Our results show that this intuition can be misleading in the case of individual mobility because an individual migrating from a community c to a community c' exerts both an emigration externality on c and an immigration externality on c'. Contrary to the aforementioned papers, we allow for various forms of group externalities and show that their shape determines the relative magnitude of the immigration and emigration externalities.

3 The Model

3.1 Communities and Preferences

The set of individuals is indexed by I. Each community is identified with a set of characteristics indexed by $c \in C$ which are independent of its membership. The welfare of individual $i \in I$ depends on the characteristics c of the community she chooses and the set of people J who choose the same community c:

$$U_i\left(c, J\right) = V_i\left(c\right) + W\left(\mu\left(J\right)\right), \tag{1}$$

where μ is a measure on I. In the sequel, we shall refer to $\mu(J)$ as the size or the mass of J. For all $\varepsilon > 0$ and $V \in \mathbb{R}^C$, we denote

$$B(V,\varepsilon) = \{i \in I : \forall c \in C, |V_i(c) - V(c)| < \varepsilon\}. \tag{2}$$

To guarantee the existence of equilibria, we assume that μ is finite and atomless:

Assumption 1 *$\mu(I) < \infty$ and for all $\varepsilon > 0$, $V \in \mathbb{R}^C$, $B(V,\varepsilon)$ is μ-measurable and*

$$\begin{aligned} \textit{for all } u &\in \mathbb{R}, c, c' \in C, c \neq c', \ \mu(\{i \in I : V_i(c) - V_i(c') = u\}) = 0, \\ \textit{for all } u &> 0, c \in C, \ \mu(\{i \in I : \sup V_i - V_i(c) = u\}) = 0. \end{aligned}$$

The function V_i embodies idiosyncratic preferences over community characteristics. If communities are religious groups, then c can be thought of as the belief of a community. If communities are networks of users of a technology, c can be thought as the properties of the technology, e.g. the functionalities of a software. In the case of political jurisdictions, c can be the tax and spending scheme of a community, its zoning laws, legal code or cultural policy.

The term W in (1) determines the group externalities. Our specification implicitly assumes that they are anonymous and uniform across players. This assumption is relaxed in section 6. The additive separability of V_i and W in (1) means that individuals rank communities of a same size m independently of m. Throughout, we will make the following technical assumptions:

Assumption 2 *$\lim_{m\to 0} W(m)$ exists but can be $\pm\infty$, W is differentiable on $]0, \mu(I)]$, and is differentiable at 0 if $\lim_{m\to 0} W(m)$ is finite.*

Either $\lim_{m \to 0} W(m)$ *is a lower bound of* W *on* $]0, \mu(I)]$ *or* $|C| < \infty$.
For any individual i, $C^*(i) = \arg\max_{c \in C} V_i(c)$ *is non empty and* $V_i(c)$ *is bounded over* $I \times C$.

Our results apply indifferently to the case of *economies of scale* (i.e. W increasing) or *congestion externalities* (i.e. W decreasing). In the latter case, assumption 2 requires C to be finite so that the congestion problem is non trivial. Some of our results will require the following assumption:

Definition 1 *The distribution of preferences has a connected support if for all* $\varepsilon > 0$, $\mu(B(V_i, \varepsilon)) > 0$ *and* $\{V_i : i \in I\}$ *is path connected for the topology of the uniform convergence on* $\mathbb{R}^C$.[1].

The following example describes a standard local public good economy which satisfies the above requirement. Communities are local jurisdictions which provide excludable services to their residents (e.g. pools, public parks, schools or police protection) and finance them via head taxes,[2] for instance because individual characteristics are unobservable or because of fiscal competition.[3] This example is analyzed in greater details in subsection 5.5, and is extended to non uniform taxation in subsection 6.3.

[1] The topology of uniform convergence is the topology generated by the sets $B(V, \varepsilon)$ for all $V \in \mathbb{R}^C$ and $\varepsilon > 0$. It coincides with the Euclidean topology when C is finite.

[2] Equal share is a common assumption in the literature on group formation (Buchanan 1965, McGuire 1974, Greenberg and Weber 1986, Konishi 1996, Alesina and Spolaore 1997, Konishi, Lebreton and Weber 1998, Jehiel and Scotchmer 2001).

[3] If jurisdictions can raise capital taxes and head taxes, capital is mobile and investment decisions are made after tax are set, the standard "race to the bottom" argument states that local jurisdictions will levy only head taxes.

Example 1 *A jurisdiction is characterized by a level of public goods* $g \in \mathbb{R}^n_+$, *a location* $l \in L$ *and a tax* t. *The cost of providing* g *in a jurisdiction of mass* m *is* $C(g,m)$; l *can be interpreted as the geographical location of a public facility or more generally as a characteristic with respect to which individuals have heterogeneous preferences. The welfare of individual* i *with type* $\theta_i \in \mathbb{R}^m$ *in a jurisdiction* (l,g,t) *of size* m *is:*

$$U_i(l,g,t) = V(l,\theta_i) + H(g) - t. \tag{3}$$

The following three cases are examples of tax and spending schemes under which (3) is a special case of (1):

(i) g *is fixed and uniform across communities and the tax balances the budget:* $t = \frac{C(g,m)}{m}$,

(ii) the tax t *is fixed and uniform across communities and* g *maximizes community welfare under the budget constraint:*

$$W(m) = \sup_{g \in \in \mathbb{R}^n_+ : C(g,m) \le mt} (H(g) - t),$$

(iii) g *and* t *are chosen in each jurisdiction so as to maximize the welfare of its residents:*

$$W(m) = \sup_{g \in \in \mathbb{R}^n_+, t \ge 0 : C(g,m) \le mt} (H(g) - t).$$

The reader can check that if H and C are smooth, C is strictly increasing in g, $V(l,\theta) - V(l',\theta)$ is continuous and nowhere locally constant in θ for all $l \neq l'$, and if the distribution of types μ^θ is absolutely continuous with compact support, then assumptions 2 and 1 are satisfied.[4] If furthermore

[4] A measure μ on $\mathbb{R}^m$ is absolutely continuous if all sets of Lebesgue measure 0 are of μ-measure 0.

μ^θ has a density function with a connected support, then the distribution of preferences has a connected support in the sense of definition 1.

3.2 Strategies

A strategy profile σ is a profile of community choices: $(\sigma_i)_{i \in I} \in C^I$. It is said to be measurable if $V_i(\sigma_i)$ is measurable. Throughout the paper, we do not distinguish between any two strategies which are equal μ-almost everywhere.[5] The distance between two strategy profiles is given by $d(\sigma, \sigma') = \mu(\{i : \sigma_i = \sigma'_i\})$.

We use the following notations: $M(c, \sigma) = \{i \in I : \sigma_i = c\}$ is the set of members of the community c at the strategy profile σ and $m(c, \sigma) = \mu(M(c, \sigma))$ is its size. A community $c \in \Gamma$ is said to be active at σ if $m(c, \sigma) > 0$. The set of active communities is denoted $A(\sigma)$. We shall refer to $M(c_\emptyset, \sigma)$ as the set of individuals in inactive communities and $m(c_\emptyset, \sigma) = \mu(M(c_\emptyset, \sigma))$.

A Nash equilibrium is a measurable strategy profile σ such that for almost all $i \in I$ and all $c \in C$, $U_i(\sigma_i, M(\sigma_i, \sigma)) \geq U_i(c, M(c, \sigma))$.[6] This notion of equilibrium implicitly assumes that individuals are free to migrate to any community but the characteristics of each community are fixed before membership decisions are made. In other words, citizens vote only with their feet. In section 7, we discuss the alternative equilibrium notion in which individ-

[5] With a slight abuse of notation, C^I will denote the quotient space of the set of measurable strategies with respect to the equivalence relation "equal μ-almost everywhere".

[6] An equivalent requirement is that for almost all $i \in I$, there exists $\varepsilon > 0$ such that it is not possible to relocate all individuals in $B(V_i, \varepsilon) = \{j \in I : \forall c \in C, |V_i(c) - V_j(c)| < \varepsilon\}$ so as to make them all better-off.

uals vote on their community characteristics after membership decisions are taken.

3.3 Social Welfare

Strategy profiles are ranked according to the utilitarian social welfare function:[7]

$$S(\sigma) = \int_i U_i(\sigma_i, M(\sigma_i, \sigma))\, d\mu(i)\,. \tag{4}$$

A measurable strategy profile σ is socially optimal if for all measurable σ', $S(\sigma) \geq S(\sigma')$. It is locally optimal if $S(\sigma) > -\infty$ and if there exists $\epsilon > 0$ such that for all measurable σ' with $d(\sigma, \sigma') < \epsilon$, $S(\sigma) \geq S(\sigma')$. For technical reasons, we will occasionally use a weaker notion of optimality which allows us to characterize the effect of free mobility more accurately:

Definition 2 *For all $\varepsilon > 0$, let μ_i^ε be the measure such that $\mu(\{i\}) = \varepsilon$ and for all μ-measurable $J \subset I$, $\mu_i^\varepsilon(J \setminus \{i\}) = \mu(J)$. A strategy profile σ is ε-optimal if for almost all $i \in I$, there exists $\epsilon > 0$ such that for all c and all $\varepsilon \in]0, \epsilon[$, under the measure μ_i^ε, $S(\sigma) \geq S(c, \sigma_{-i})$.*

In words, ε-optimality means that even if i had a small but positive measure, i could not increase social welfare on its own by migrating to a different community. As we shall see, it is slightly weaker than local optimality.

[7] One can check that from assumptions 1 and 2, for all measurable strategy profile σ, $U_i(\sigma)$ is either integrable or (4) is $-\infty$.

4 Preliminary Lemmas

In this section, we introduce the notion of potential consistency and derive important intermediate results. The proofs of the lemmas are relegated to the appendix.

Definition 3 *A potential* p *is a map from* C *to* $\mathbb{R} \cup \{\pm\infty\}$ *such that for all but a countable number of communities,* $p(c) = p_\emptyset$ *for some* $p_\emptyset \in \mathbb{R} \cup \{\pm\infty\}$*. A strategy profile* σ *is consistent with a potential* p *if for all* $c, c' \in C$*, for almost all* $i \in I$*,* $\sigma_i \in C(p, V_i)$*, where*

$$C(p, V_i) \equiv \arg\max_{c \in C} (V_i(c) + p(c)), \tag{5}$$

and for all $c \notin A(\sigma)$*,* $p(c) = p_\emptyset$ *for some* $p_\emptyset \in \mathbb{R} \cup \{-\infty\}$*.*

It follows readily from definition 3 that a strategy profile σ is a Nash equilibrium if and only if σ is consistent with the potential $p_\sigma^{NE}(c) = W(m(c, \sigma))$. The following two remarks show that a potential defines a unique strategy profile and that, reciprocally, a strategy can be consistent with a unique potential up to a constant:

Remark 1 *If* σ *is consistent with a potential* p*, then* $C(p, V_i)$ *is single-valued for almost all* i*, and thus* σ *is the only strategy profile consistent with* p*.*

Remark 2 *If the distribution of preferences has a connected support in the sense of definition 1 and if* σ *is consistent with two potentials* p *and* p'*, then there exists* $\delta \in \mathbb{R}$ *such that for all* $c, c' \in A(\sigma)$*,* $p'(c) = p(c) + \delta$*. Moreover, if* $m(c_\emptyset, \sigma) > 0$*,* $p'_\emptyset = p_\emptyset + \delta$*.*

Lemma 1 shows that any strategy which is consistent with a potential achieves an optimal sorting of individuals between communities conditional on a size constraints for each community.

Lemma 1 *Let p be a potential and σ^p be a strategy profile consistent with p, then σ^p is the unique solution of*

$$\max_{\sigma \in C^I : \forall c \in C,\ m(c,\sigma) = m(c,\sigma^p)} S(\sigma).$$

As argued earlier, a Nash equilibrium σ is consistent with the potential $W(m(c,\sigma))$, so lemma 1 implies that free mobility automatically achieves an optimal sorting of individuals conditional on the size of the groups that σ assigns to each community. Hence, Nash equilibria are suboptimal only if communities have non optimal sizes. Contrary to the typical setup in club theory, the latter kind of inefficiency is a genuine concern even with a continuum of individuals: because the optimal size of communities is not necessarily negligible, membership is a scarce resource which has to be allocated optimally. The next lemma shows that for a profile of community choices σ to generate communities of optimal size, it must be consistent with the potential p^*_σ defined as:

$$\begin{aligned} \forall c &\in A(\sigma), \ p^*_\sigma(c) = W(m(c,\sigma)) + m(c,\sigma) W'(m(c,\sigma)) \\ \forall c &\notin A(\sigma), \ p^*_\sigma(c) = \lim_{m \to 0} W(m). \end{aligned} \tag{6}$$

Lemma 2 *A strategy profile σ is ε-optimal if and only if it is consistent with the potential p^*_σ defined in (6).*

In other words, if individuals were choosing their community on the basis of the following utility function:

$$U_i(\sigma) = V_i(\sigma_i) + W(m(\sigma_i, \sigma)) + m(\sigma_i, \sigma) W'(m(\sigma_i, \sigma)),$$

rather than $U_i(\sigma) = V_i(\sigma_i) + W(m(\sigma_i, \sigma))$ then free mobility would be equivalent to ε-optimality. By comparing p^{NE} and p^*, lemma 2 implies that if a community-contingent tax scheme equal to $m(c, \sigma) W'(m(c, \sigma))$ was imposed on all individuals, privately and socially optimal community choices would coincide. The intuition behind this result is a simple Pigouvian argument: as a small mass ε of individuals leave a community c_1 of mass m_1 to join another community c_2 of mass m_2, they impose an emigration externality on the members of community c_1 which is given by

$$(m_1 - \varepsilon)(W(m_1 - \varepsilon) - W(m_1)) = m_1 W'(m_1)\varepsilon + o(\varepsilon).$$

They impose an immigration externality on the members of community c_2 which is given by

$$m_2 (W(m_2 + \varepsilon) - W(m_2)) = m_2 W'(m_2)\varepsilon + o(\varepsilon).$$

Hence, the marginal externality imposed by some migrant on the members of some community c for joining c rather than another community is $m(c, \sigma) W'(m(c, \sigma))$.

The next lemma shows that in general, ε-optimality is weaker than local optimality. It implies global optimality under some convexity condition and it implies local optimality only in a “first order sense”.

Lemma 3 *If σ is locally optimal, it is ε-optimal. Reciprocally, if σ is ε-optimal and $mW(m)$ is concave, σ is optimal. If σ is ε-optimal and* $\frac{\partial[mW(m)]}{\partial m}$

is Lipschitz continuous, then for all σ', $S(\sigma') \leq S(\sigma) + O\left((d(\sigma', \sigma))^2\right)$ *as* $d(\sigma', \sigma) \to 0$.[8] *If* $\frac{\partial^2[mW(m)]}{\partial m^2} > 0$ *for some* $m_o > 0$, *then there exists a distribution of preferences* $\{V_i : i \in I\}$ *and a strategy profile* σ *which is* ε-*optimal but not locally optimal.*

5 Free Mobility and Social Welfare

5.1 Equilibria and Social Optima: Existence

We denote $C^* = \{c \in C : \mu(i \in I : c \in C^*(i)) > 0\}$ the set of communities which are the most preferred community of a non-negligible set of individuals. The proofs of the next two propositions can be found in the appendix.

Proposition 1 *If* C^* *is finite or if* $\lim_{m \to 0} W(m) = -\infty$, *a Nash equilibrium exists.*

It should be clear from the proof of proposition 1 that when $\lim_{m \to 0} W(m) = -\infty$, for any set $D \subset S$ one can find a Nash equilibrium whose support is in D, hence there will typically be many Nash equilibria.

Proposition 2 *If* C *is finite, a social optimum exists. If* C *is compact,* $V_i(c)$ *is continuous in* c *for all* i, W' *is bounded below on* $]0, \mu(I)]$ *and* $\lim_{m \to 0} W(m) = -\infty$, *a social optimum exists.*

[8] A function $f(x)$ is a $O(x)$ as $x \to 0$ if there exists M and $\delta > 0$ such that for all $x \in [0, \delta]$, $|f(x)| < Mx$.

5.2 Logarithmic Externalities and Social Optimality

The following definition will play a central role in our argument:

Definition 4 *Group externalities are logarithmic if for all* $m \in [0, \mu(I)]$, $W(m) = \alpha \ln(m)$ *for some* $\alpha \in \mathbb{R}$*. They are more (resp. less) than logarithmic if for all* $m \in [0, \mu(I)]$, $mW'(m)$ *is increasing (resp. decreasing).*

Since $\frac{d[mW(m)]}{dm} = W(m) + mW'(m)$, group externalities are more (resp. less) than logarithmic if the marginal aggregate group externality $\frac{d[mW(m)]}{dm}$ increases more (resp. less) rapidly than the per capita externality $W(m)$.

Observe that when W is logarithmic, the potential $p^*_\sigma(c)$ as defined in (6) is given by $p^*_\sigma(c) = W(m(c, \sigma)) + \alpha$ for some $\alpha \in \mathbb{R}$. From lemma 2, σ is ε-optimal if and only if it consistent with the potential $p^*_\sigma(c)$, which means that σ is a Nash equilibrium. Hence, we immediately have the following:

Proposition 3 *If* W *is logarithmic, then* σ *is a Nash equilibrium if and only if it is* ε*-optimal.*

The basic intuition is the following: as a small mass ε of individuals leave a community c_1 of mass m_1 to join another community c_2 of mass m_2, the aggregate utilitarian externality on these two communities cancel out: the emigration externality imposed on the members of community c_1 is $-m_1 \frac{d[\alpha \ln(m)]}{dm}(m_1)\varepsilon = -\alpha\varepsilon$ while the immigration externality imposed on the members of community c_2 is $m_2 \frac{d\alpha \ln(m)}{dm}(m_2)\varepsilon = \alpha\varepsilon$. Hence, when group externalities increase (or decrease) logarithmically, individual mobility has no aggregate externality on the rest of society.

Corollary 1 *Suppose W is logarithmic. If W is increasing, C is compact and $V_i(c)$ is continuous in c for all i, there exists a socially optimal Nash equilibrium. If W is weakly decreasing, the set of social optima coincides with the set of Nash equilibria and is non-empty.*

Proof. Suppose $W = \alpha \ln$. The existence of a social optimum comes from the second part of proposition 2 for $\alpha > 0$ and from assumption 2 and the first part of proposition 2 for $\alpha < 0$. The case $\alpha = 0$ is obvious from assumption 2. From lemma 3, optimality implies ε-optimality and is equivalent for $\alpha \leq 0$ so proposition 3 completes the proof. ■

The next proposition shows that the logarithmic condition in proposition 3 is essentially necessary:

Proposition 4 *Suppose the distribution of preferences has a connected support in the sense of definition 1. If $m(c_1,\sigma) W'(m(c_1,\sigma)) \neq m(c_2,\sigma) W'(m(c_2,\sigma))$ for some $c_1, c_2 \in A(\sigma)$ or if $m(c_\emptyset,\sigma) > 0$ and $m(c,\sigma) W'(m(c,\sigma)) \neq 0$ for some $c \in A(\sigma)$, then σ is not locally optimal.*

Proof. Since σ is a Nash equilibrium, it is consistent with the potential $W(m(c,\sigma))$. If $W(m(c,\sigma))$ and $p^*_\sigma(c) = W(m(c,\sigma)) + m(c,\sigma) W'(m(c,\sigma))$ are not equal up to a constant, remark 1 implies that σ cannot be consistent with the potential $p^*_\sigma(c)$. Therefore, from lemma 2, σ is not ε-optimal. From lemma 3, it is not locally optimal. ■

As a direct corollary, if $mW'(m)$ is one to one, then a Nash equilibrium can be locally optimal only if all groups have the same size in equilibrium, which can happen only under non generic distribution of preferences.[9]

[9]This remark is reminiscent of the stringent conditions made in finite club economy models to guarantee the existence of an optimal equilibrium. In our model, they would

5.3 Non Logarithmic Externalities and Fragmentation

A simple heuristic reasoning sheds some light on the nature of the inefficiency when W is more or less than logarithmic (c.f. definition 4): if σ is a Nash equilibrium, c_1 and c_2 are two neighbors communities (in the sense that some individual is indifferent between them) of size m_1 and m_2, then by moving a mass ε of people "close to the border" between c_1 and c_2, total welfare changes by:

$$-m_1 W'(m_1)\,\varepsilon + m_2 W'(m_2)\,\varepsilon,$$

which is negative if W is more than logarithmic: joining a smaller group imposes a negative aggregate externality on society. This suggests that when group externalities increase more than logarithmically, too few individuals will join large communities, too many will migrate to small communities and the free mobility equilibrium will be excessively fragmented. By the same token, if W is less than logarithmic, free mobility will lead to excessive agglomeration. To make this statement precise, we need to define formally the notion of agglomeration and fragmentation:

Definition 5 *Let $(m_n^\sigma)_{n\in\mathbb{N}}$ denote the mass of the active communities for the strategy profile σ ranked in decreasing order, with the convention that $m_n = 0$ for $n > |A(\sigma)|$. We say that σ is more concentrated than σ' if*

---

translate as follows: W is maximized at some finite m^*, there exists a finite number of types, a mass nm^* of individuals of each type for some $n \in \mathbb{N}$ (no scarcity on the demand side), and for all i, $|C^*(i)| \geq n$ (no scarcity on the supply side). In this case, the optimal partition is such that each individual i is in a community in $C^*(i)$ of size m^* and such a partition is trivially stable under free mobility. Consistently with proposition 4, all clubs are such that $mW'(m) = 0$.

$m(c_\emptyset,\sigma) \le m(c_\emptyset,\sigma')$ *and for all* $p \in \mathbb{N}$,[10]

$$\sum_{n \le p} m_n^\sigma \ge \sum_{n \le p} m_n^{\sigma'}.$$

A strategy profile σ *exhibits excessive fragmentation (resp. agglomeration) if there exists* $\varepsilon > 0$ *such that for all* σ'*, if* $d(\sigma,\sigma') \le \varepsilon$ *and* σ *is more concentrated than* σ' *(resp* σ' *is more concentrated than* σ*), then* $S(\sigma) \ge S(\sigma')$*.*

In words, a partition of society is excessively fragmented if it is not possible to increase welfare via small migration movements from larger to smaller communities. In particular, a local optimum is neither excessively fragmented nor excessively agglomerated.

Proposition 5 *If* W *is more (resp. less) than logarithmic, then all Nash equilibria* σ *such that* $A(\sigma)$ *is finite exhibit excessive fragmentation (resp. excessive agglomeration).*

Proof. Let σ be a Nash equilibrium such that $A(\sigma)$ is finite and let σ' be such that $d(\sigma,\sigma') = \varepsilon$. If we denote $J = \{i \in I : \sigma_i \ne \sigma_i'\}$, then

$$\begin{aligned} S(\sigma') - S(\sigma) &= \int_{i \in J} \left(V_i(\sigma_i') + W(m(\sigma_i',\sigma)) - V_i(\sigma_i) - W(m(\sigma_i,\sigma))\right) d\mu(i) \\ &\quad + \sum_{c \in A(\sigma)} m(c,\sigma')\left(W(m(c,\sigma')) - W(m(c,\sigma))\right) \end{aligned} \tag{7}$$

[10] This condition is equivalent to the requirement that $(m_n^\sigma - m(c_\emptyset,\sigma'))_{n\in\mathbb{N}}$ majorizes $\left(m_n^{\sigma'} - m(c_\emptyset,\sigma)\right)_{n\in\mathbb{N}}$. See e.g. Hardy, Littlewood and Polya 1988. It means that σ' obtains from σ by a finite sequence of relocation from smaller to larger communities (see e.g. Arnold 1987).

Since σ is a Nash equilibrium, the first term in the right hand-side of (7) is negative. Since W is differentiable and $A(\sigma)$ is finite, (7) can be approximated by

$$\sum_{c\in A(\sigma)} m(c,\sigma) W'(m(c,\sigma))(m(c,\sigma')-m(c,\sigma))+o(\mu(I)) \tag{8}$$

$$= \sum_{c\in A(\sigma)} m(c,\sigma) W'(m(c,\sigma))\begin{pmatrix} m(c,\sigma')-m(c_\emptyset,\sigma') \\ -m(c,\sigma)+m(c_\emptyset,\sigma)\end{pmatrix} \tag{9}$$

$$+(m(c_\emptyset,\sigma')-m(c_\emptyset,\sigma))\sum_{c\in A(\sigma)} m(c,\sigma) W'(m(c,\sigma))+o(\mu(I)) \tag{10}$$

If W is more than logarithmic, then $m(c,\sigma) W'(m(c,\sigma))$ is increasing in $m(c,\sigma)$. If σ' is more concentrated than σ, then $(m(c,\sigma')-m(c_\emptyset,\sigma'))_{c\in A(\sigma)}$ majorizes $(m(c,\sigma)-m(c_\emptyset,\sigma))_{c\in A(\sigma)}$ (see e.g. Arnold 1987) so the first term in the right hand-side of (8) is negative. Moreover, in absolute value it must be greater than

$$\min_{\substack{c,c'\in A(\sigma): \\ m(c,\sigma)W'(m(c,\sigma))\neq m(c',\sigma)W'(m(c',\sigma))}} \mu(J)\left|m(c,\sigma) W'(m(c,\sigma))-m(c',\sigma) W'(m(c',\sigma))\right|,$$

so it dominates the term $o(\mu(I))$ in (8).

Finally, if $W'\geq 0$, the last term in the right hand-side of (8) since $m(c_\emptyset,\sigma')\leq m(c_\emptyset,\sigma)$, which proves that $S(\sigma')\leq S(\sigma)$. If $W'(m)<0$ for some m, then since $mW'(m)$ is increasing, this means that $\limsup_{m\to 0} mW'(m)$ is negative. This implies that $\lim_{m\to 0} W(m)$ is not a lower bound on W, so from assumption 2, C must be finite and with a slight abuse of notation, we can take $A(\sigma)=C$ and $m(c_\emptyset,\sigma)=0$ in the preceding reasoning. The case W less than logarithmic can be treated similarly. ■

For instance, if group externalities are increasing and linear as in Konishi, Lebreton and Weber 1997a, then $mW'(m)=\alpha m$ for some $\alpha>0$ so free mobility leads to too much fragmentation. If each community provides a fixed, excludable public good (called a "government" in Alesina and Spolaore 1997 and Lebreton and Weber 2003) the cost of which is affine in community size and shared equally then $W(m) = -\frac{\alpha}{m} - \beta$ for some $\alpha, \beta > 0$ and $mW'(m) = \frac{\alpha}{m}$. From proposition 5, the free mobility equilibria will exhibit excessive agglomeration.

Observe that the results in subsection 5.2 and 5.3 hold irrespective of the distribution of preferences $\{V_i : i \in I\}$ and the set of available communities C. Moreover, proposition 4 shows that the alternative exercise of characterizing distributions of preferences for which free mobility is compatible with social optimality for a given W would lead to very restrictive conditions whenever W is not logarithmic. This suggests that in our setup, the social cost of free mobility depends on the shape of economies of scale/congestion externalities rather than on the distribution of preferences.

5.4 Local Public Good Economies

5.5 Local Public Good Economies

In this section, we apply our results to the local public good economies described in example 1.

5.5.1 Exogenous Local Public Goods

Let us first assume that the public good bundle is fixed and uniform across jurisdictions and financed by head taxes (i.e. case (i)). In this case, we can omit the reference to g and $W(m) = H - c(m)$ where $H \in \mathbb{R}$ and $c(m)$ is the per capita cost of providing g to a jurisdiction of mass m. Ideally, the size of each jurisdiction should trade-off the cost of preferences heterogeneity and the gains from economies of scale. From proposition 3, free mobility achieves such a trade-off if $c(m)$ is logarithmic in m, that is if the total cost of providing g in a community of size m is $C(m) = m(\alpha \ln(m) + \beta)$ for some $\alpha, \beta \in \mathbb{R}$. More generally, since $\frac{dC}{dm} = c + m\frac{dc}{dm}$, W is more (resp. less) than logarithmic in m if and only if the marginal cost $\frac{dC}{dm}$ increases less (resp. more) rapidly than the per capita cost c. So proposition 5 implies the following:

Corollary 2 *If the marginal cost of providing the local public good increases more (resp. less) rapidly with the jurisdiction size than the per capita cost, Nash equilibria will be excessively agglomerated (resp. excessive fragmented).*

It should be noticed that in this setup, the case in which $\frac{dC}{dm}$ increases more rapidly than c is probably the more realistic scenario. Indeed, if $C(m) = F + \alpha m^{\rho}$ for some $F, \alpha, \rho \in \mathbb{R}^{+}$, then $\frac{dC}{dm} - c = \alpha(\rho - 1)m^{\rho - 1} - \frac{F}{m}$ which is increasing in m. So in this case, independently of the economies of scale ρ and the fixed costs F, free mobility will lead to excessive agglomeration. This result provides an interesting counterpart to Alesina and Spolaore 1997, Botlon and Roland 1997 or Lebreton and Weber 2003 who argue that the possibility of secession can lead to inefficiently fragmented partitions of society.

5.5.2 Optimal Local Public Goods

Let us now consider the case in which each jurisdiction chooses g and t so as to maximize the welfare of its residents (i.e. case (iii)). If $g^{*}(m)$ denotes the maximizer of $H(g) - c(g, m)$, free mobility leads to an optimal partition of society if $H(g^*(m)) + c(g^*(m), m)$ increases logarithmically in m. Using the envelope theorem, this condition is satisfied if and only if $m\frac{\partial c}{\partial m}(g^*(m), m)$ is constant in m.

To illustrate the applicability of our results, let us consider the case in which the cost of provision is multiplicative in g and m, i.e. $C(g, m) = s(m)\, c(g)$, and g is expressed in money metric utils:

$$U_i(g, l) = V(l, \theta_i) + g - \frac{s(m)}{m} c(g).$$

The welfare of individual i in a community located at l of size m is:

$$U_i^*(l, m) = V(l, \theta_i) + \frac{s(m)}{m} \max_{g \geq 0} \left[\frac{mg}{s(m)} - c(g) \right] = V(l, \theta_i) + \frac{s(m)}{m} c^* \left(\frac{m}{s(m)} \right),$$

where c^* is the convex conjugate of c (see e.g. Rockafellar 1970). One can easily check that if $c(x) = \exp(\beta x)$, then $c^*(y) = \frac{y}{\beta} \ln\left(\frac{y}{\beta}\right) - \frac{y}{\beta}$, so in the parametric case $C(g, m) = \alpha m^\rho \exp(\beta x)$,

$$\begin{aligned} U_i^*(l) &= V(l, \theta_i) + \alpha m^{\rho-1} \left(\frac{1}{\alpha \beta m^{\rho-1}} \ln\left(\frac{1}{\alpha \beta m^{\rho-1}} \right) - \frac{1}{\alpha \beta m^{\rho-1}} \right) \\ &= V(l, \theta_i) + \frac{1-\rho}{\beta} \ln(m) - \frac{1}{\beta} - \frac{\ln(\alpha\beta)}{\beta}, \end{aligned}$$

which shows that group externalities are logarithmic. Hence, proposition 3 implies:

Corollary 3 *If $C(g, m) = \alpha m^\rho c(g) \exp(\beta g)$ for some $\alpha, \rho, \beta \in \mathbb{R}$, where g is measured in utils, and each local jurisdiction chooses g so as to maximize the welfare of its constituents, then independently of the economies of scale ρ, there exists a socially optimal Nash equilibrium.*

6 Relational Heterogeneity

In this subsection, we allow preferences to exhibit not only *idiosyncratic heterogeneity* (i.e. heterogeneity in preferences on group characteristics through the V_i term) but also *relational heterogeneity* (i.e. heterogeneity in preferences over community membership through the W term). This type of heterogeneity arises for instance if group externalities are not anonymous, that is if they depend not only on the number of members but also on their characteristics (social manners, religious beliefs, wealth...). In the local public good economies described in example 1, this is the case if public goods are financed by income taxes and there is income inequalities. Relational heterogeneity can arise also with anonymous group externalities if individuals have different preferences over community size. In example 1, this is the case if head taxes are fixed, the level of public good varies with the jurisdiction size (case (ii)) and the willingness to pay for the public goods (i.e. H) is heterogeneous. Different individuals would then trade off differently the satisfaction of their idiosyncratic preferences V_i with the size of the tax base.

For simplicity, we assume that society is composed of a set I^a of individuals of type a and a set I^b of individuals of type b. The welfare of each type of individual in a community c with a set J^a of individuals of type a and a set J^b of individuals of type b is given by:

$$\begin{aligned} U_i^a &= V_i(c) + W^a\left(\mu^a(J^a), \mu^b\left(J^b\right)\right), \\ U_i^b &= V_i(c) + W^a\left(\mu^a(J^a), \mu^b\left(J^b\right)\right). \end{aligned}$$

Throughout this subsection, we omit the type superscript for vector notations. On top of assumption 1 and 2, we assume that W is twice differentiable on $]0,\mu^a(I^a)]\times\left]0,\mu^b\left(I^b\right)\right]$ and $\lim_{m\to\left(m^a,m^b\right)} W(m)$ exists but can be infinite when $m^a=0$ or $m^b=0$. The set of members of type $t\in\{a,b\}$ in community c is denoted $M^t(c,\sigma)$ and its mass $m^t(c,\sigma)=\mu^t(M^t(c,\sigma))$. A community c is inactive if $m(c,\sigma)=(0,0)$.

6.1 Symmetry and Social optimality

Definition 3 becomes:

Definition 6 *A potential p is a pair of maps from C to $\mathbb{R}\cup\{\pm\infty\}$ such that for all but a countable number of communities, $p(c)=\left(p^a_\emptyset,p^b_\emptyset\right)$ for some $p^a_\emptyset,p^b_\emptyset\in\mathbb{R}\cup\{\pm\infty\}$.*
A strategy profile σ is consistent with a potential p if for all $t\in\{a,b\}$, for all $c,c'\in C$, for almost all $i\in I^t$, $\sigma_i\in\arg\max V_i(c)+p^t(c)$ and for all $c\notin A(\sigma)$, $p^t(c)=p^t_\emptyset$ for some $p^t_\emptyset\in\mathbb{R}\cup\{-\infty\}$.

One can easily adapt the proof of proposition 1 to show the existence of a Nash equilibrium. As in the case of anonymous group externalities, a strategy profile σ is a Nash equilibrium if and only if σ is consistent with the potential $p^{NE}_\sigma(c)=W(m(c,\sigma))$. By analogy with lemma 2, the potential that characterizes ε-optimality is determined by the migration externalities: by joining a community c of mass $\left(m^a,m^b\right)$, a small mass ε of individuals of type a imposes an externality on the residents of c of type a and b which is given by:

$$\begin{aligned} m^a\left[W^a\left(m^a+\varepsilon,m^b\right)-W^a\left(m^a,m^b\right)\right] &\simeq m^a\frac{\partial W^a}{\partial m^a}\varepsilon, \\ m^b\left[W^b\left(m^a+\varepsilon,m^b\right)-W^b\left(m^a,m^b\right)\right] &\simeq m^b\frac{\partial W^b}{\partial m^a}\varepsilon, \end{aligned}$$

Hence, one can easily adapt the proof of lemma 2 to show that σ is ε-optimal if and only if σ os consistent with the potential

$$p_\sigma^* = W\left(m\left(c,\sigma\right)\right) + \left(m^a \frac{\partial W^a}{\partial m^a} + m^b \frac{\partial W^b}{\partial m^a}, m^a \frac{\partial W^a}{\partial m^b} + m^b \frac{\partial W^b}{\partial m^b}\right). \tag{11}$$

The same argument as in proposition 3 and 4 implies the following:

Proposition 6 *If the migration externalities $\left(E^a, E^b\right)$ defined as:*

$$E^a\left(m\right) = m^a \frac{\partial W^a}{\partial m^a} + m^b \frac{\partial W^b}{\partial m^a},\ E^b\left(m\right) = m^b \frac{\partial W^b}{\partial m^b} + m^a \frac{\partial W^a}{\partial m^b},$$

are constant in m, then for all distribution of preferences, σ is a Nash equilibrium if and only if it is ε-optimal.
Reciprocally, if $\{V_i : i \in I_a\}$ and $\{V_i : i \in I_b\}$ are connected in the sense of definition 1, and if σ is a Nash equilibrium such that $E^a\left(m\left(c,\sigma\right)\right)$ or $E^b\left(m\left(c,\sigma\right)\right)$ is not the same for all active communities, then σ is not locally optimal.

The condition that E^a and E^b are constant in $\left(m^a, m^b\right)$ is more restrictive than the logarithmic condition in the anonymous externality case. The reason is that when group externalities are not anonymous, social optimality requires communities to be not only of optimal size but also of optimal composition. The following proposition decompose the condition

Proposition 7 *The migration externalities $\left(E^a, E^b\right)$ defined in proposition 6 are constant in m if and only if there exists $\alpha, \beta \in \mathbb{R}$ such that for all $m \in \mathbb{R}_+^2$ and all $\gamma > 0$,*

$$E\left(\gamma m\right) = \left(\alpha, \beta\right) \ln\left(\gamma\right) + E\left(m\right), \tag{12}$$

$$\frac{\partial E^a}{\partial m^b} = \frac{\partial E^b}{\partial m^a}. \tag{13}$$

Proof. Suppose $E(m) \equiv (\alpha, \beta)$ for some $\alpha, \beta \in \mathbb{R}$. Substracting $\frac{\partial E^a}{\partial m^b} = 0$ to $\frac{\partial E^b}{\partial m^a} = 0$ we get (13). Substituting in $\left(E^a, E^b\right)$, we get

$$E(m) = m^a \frac{\partial E}{\partial m^a}(m) + m^b \frac{\partial E}{\partial m^b}(m). \tag{14}$$

So for all $\gamma > 0$,

$$\frac{E(\gamma m)}{\gamma} = m^a \frac{\partial E}{\partial m^a}(\gamma m) + m^b \frac{\partial E}{\partial m^b}(\gamma m) = \frac{\partial E(\gamma m)}{\partial \gamma} = \frac{(\alpha, \beta)}{\gamma}.$$

Integrating the last equation above with respect to γ, we get that for all $\gamma > 0$,

$$E(\gamma m) = (\alpha, \beta) \ln(\gamma) + \left(f^a(m), f^b(m)\right).$$

Substituting $\gamma = 1$, we get (12).

Reciprocally, if E satisfies (12), then the right hand-side of (14) is constant. If it satisfies (13) as well, then equality (14) holds and E is constant. ■

Condition (12) generalizes the logarithmic condition of proposition 3. Roughly speaking, this condition corresponds to the requirement that groups are of optimal size. Condition (13) imposes some form of symmetry: in each community, the benefit for the members of type a of an additional member of type b must be equal to the benefit for the members of type b of an additional member of type a. Intuitively, this condition corresponds to the additional requirement that the composition of each group is optimal.

To illustrate the conditions of proposition 7, consider the following subclass of heterogeneous group externality:

Definition 7 *A profile of group externalities W is additively separable in group size and composition if there exists functions f^a, f^b, g^a and g^b such* *that:*

$$W(m) = \left(f_a(m^s) + g_a(r^a), f_b(m^s) + g_b\left(r^b\right)\right), \qquad (15)$$

where $m^s = m^a + m^b$, $r^a = \frac{m^a}{m^a+m^b}$ *and* $r^b = \frac{m^b}{m^a+m^b}$

If W is additively separable, simple calculus yields:

$$\begin{aligned} E^a(m) &= m^s\left[r^a f_a'(m^s) + r^b f_b'(m^s)\right] + r^a r^b g_a'(r^a) - \left(r^b\right)^2 g_b'\left(r^b\right), \quad (16) \\ E^b(m) &= m^s\left[r^a f_a'(m^s) + r^b f_b'(m^s)\right] + r^a r^b g_b'\left(r^b\right) - (r^a)^2 g_a'(r^a). \end{aligned}$$

Since $E(\gamma m)$ must be constant in γ, (16) implies that f_a and f_b must be logarithmic. Substituting in (16), this implies in turn that for all r, $rg_a'(r) = (1-r)\,g_b'(1-r)$. This condition is satisfied for instance if the homophily terms are logarithmic and symmetric, i.e. $g(r) = \left(\alpha \ln(r^a), \beta \ln\left(r^b\right)\right)$ and $\alpha = \beta$.

6.2 Homophily and Segregation

In line with proposition 5, one can derive conditions on the migration externality functions under which Nash equilibria will be too stratified or on contrary excessively mixed. To do so, we focus on the case in which group externalities group externalities are additively separable in group size and composition (c.f. definition 7).

Definition 8 *Under the notations of definition 7, group externalities exhibit more (resp. less) than logarithmic homophily if* $rg_a'(r)$ *and* $rg_b'(r)$ *are increasing (resp. decreasing) in* r*.*

A simple heuristic reasoning sheds some light on the nature of the inefficiency when group externalities are not logarithmically homophylic: if σ is a

Nash equilibrium, c_1 and c_2 are two neighbors communities (in the sense that some individual of both types are indifferent between them) of composition m_1 and m_2 such that $r_1^a > r_2^a$, then by moving a mass ε of individual of type a "close to the border" from c_1 to c_2 and by moving the same mass ε of individual of type b "close to the border" from c_2 to c_1, using (16), total welfare changes by:

$$\begin{aligned} & \varepsilon\left[-E^a\left(m_1\right)+E^b\left(m_1\right)\right]+\varepsilon\left[E^a\left(m_2\right)-E^b\left(m_2\right)\right] \\ = & \varepsilon\left[-g_a'\left(r_1^a\right)r_1^a+r_2^a g_a'\left(r_2^a\right)\right]-\varepsilon\left[-r_1^b g_b'\left(r_1^b\right)+r_2^b g_b'\left(r_2^b\right)\right] \end{aligned}$$

The expression above is negative whenever $rg_a'(r)$ and $rg_b'(r)$ are increasing: migrating to a community with a smaller proportion of individual of one's type generates a negative aggregate externality on society. This suggests that when homophily is more than logarithmic, too many individuals of type a will join communities with a minority of type b and vice versa so the free mobility equilibrium will be insufficiently segregated. By the same token, when homophily is less than logarithmic, free mobility will lead to excessive segregation. To make this statement precise, we need to define formally the notion of segregation

Definition 9 *Let σ and σ' be two strategies such that for all $c \in C$,*

$$\left(m^a+m^b\right)(c,\sigma)=\left(m^a+m^b\right)(c,\sigma').$$

We say that σ' is more segregated than σ if σ' can be obtained from σ by a sequence of migrations in which individuals of type a migrate to communities with a greater share of type a while individuals of type b migrate to communities with a greater share of type b.

A strategy profile σ exhibits excessive (resp. insufficient) segregation if there exists $\varepsilon > 0$ such that for all σ', if $d(\sigma, \sigma') \leq \varepsilon$ and σ is more (resp. less) segregated than σ', then $S(\sigma) \geq S(\sigma')$.

In words, a partition of society is excessively segregated if it is not possible to increase welfare by relocating individuals so as to form more segregated groups. Observe that the notion of excessive segregation, as the notion of excessive fragmentation, is a local one. In particular, a local optimum is neither excessively nor insufficiently stratified.

Proposition 8 *If group externalities exhibit more (resp. less) than logarithmic homophily, then all Nash equilibria σ such that $A(\sigma)$ is finite are insufficently (resp. excessively) segregated.*

Proof. Available soon... ■

6.3 Applications

6.3.1 Preferences Heterogeneity for Local Public Goods

To illustrate the implications of condition (13), consider the local public good economy of example 1 in which members pay type-dependent community fees (τ^a, τ^b) and communities use their revenue from membership fees to provide community services (i.e. case (ii)). If the local public goods are non rival, group externalities are then given by

$$W(m) = (\theta^a, \theta^b) H(\tau^a m^a + \tau^b m^b) - (\tau^a, \tau^b), \tag{17}$$

for some function H. The type θ parametrizes the willingness to pay for the community goods. In this setup, the symmetry condition (13) boils down to $\theta^a/\tau^a = \theta^b/\tau^b$, i.e. that each individual is taxed in proportion to her willingness to pay for local public goods. This condition is reminiscent of the Lindhal condition for the revelation of preferences and the provision of public good absent any mobility issue. Hence, although the optimal allocation of individuals between jurisdictions and the optimal provision of public goods within a given jurisdiction are two independent economic problems, it is interesting to notice that they lead to similar conclusions.

6.3.2 Mobility and redistribution

The literature on federalism has long argued that free mobility imposes a constraint on local redistribution (see e.g. Brown and Oates 1987, Wildasin 1991, Epple and Romer 1991 or Konishi, Lebreton and Weber 1998). In this subsection, we explore this issue and consider the case in which different types have different incomes and are taxed for public good financing and redistribution purposes. The examples we derive below are admittedly stylized and are meant to illustrate the applicability of our results.

Consider again the last example in which group preferences are given by (17). The preceding reasoning shows that unless income is perfectly and positively correlated with the willingness to pay for the local public good, redistribution will lead to inefficient location decisions. To qualify the distortion in community choices, consider for simplicity the scenario in which rich and poor have the same willingness to pay for the public good $\theta^a = \theta^b = 1$, and $H = \ln$, so that community size is not an issue. Then it is straightforward to see that proposition 3 holds in this case if we change the utilitarian social welfare function S by weighting each individual by its tax contribution:

$$S_\tau(\sigma) = \tau^a \int_{i \in I^a} U_i^a(\sigma_i, M(\sigma_i, \sigma))\, d\mu(i) + \tau^b \int_{i \in I^b} U_i^b(\sigma_i, M(\sigma_i, \sigma))\, d\mu(i)\,.$$

Hence, free mobility equilibria will be the ε-optima of a regressive social welfare ordering in which richer individuals are assigned a greater social weight. Individuals will tend to agglomerate in the communities with richer members while the poor communities will attract fewer members and enjoy less economies of scale. Moreover, the more redistributive the tax scheme, i.e. the greater τ^a/τ^b and the more regressive S_τ, and the more "regressive" will be the allocations of individuals into communities.

Suppose now that the tax progressivity $\pi = \tau^a/\tau^b$ is fixed and the tax rate t is endogenized in each community. For simplicity, suppose that t is chosen so as to maximize the utilitarian welfare in each community:

$$t(m) = \arg\max \begin{bmatrix} m^a\left(\ln\left(\pi t m^a + t m^b\right) - \pi t\right) \\ +m^b\left(\ln\left(\pi t m^a + t m^b\right) - t\right) \end{bmatrix} = \frac{m^a + m^b}{\pi m^a + m^b}.$$

The effect of mobility can be seen by looking at the migration externalities as defined in proposition 6. Simple calculus gives:

$$\left(E^a, E^b\right)(m) = (1,1) + (\pi - 1)\left(\frac{\pi}{1 + \pi\frac{m^a}{m^b}}, -\frac{1}{1 + \pi\frac{m^b}{m^a}}\right), \qquad (18)$$

The migration externalities E depends only on the ratio of rich and poor m^a/m^b in a given community. In particular, migration externalities are independent of community size so there will not be any distortion between communities with similar ratio of poor and rich. However, E is constant at some equilibrium only if m^a/m^b is constant across communities. One can see from (18) that whenever $\pi > 1$, E^a and E^b are both increasing in m^a/m^b:

when the tax rate is progressive, joining poorer communities generates positive externalities and is thus under-provided in equilibrium. From (18), the distortion will be greater the greater the progressivity π of the tax and the more variations there are in m^a/m^b across communities.

7 Endogenous Policy Making

As argued in section 3, the Nash equilibrium requirement implicitly assumes that the characteristics c of each community are fixed while membership is endogenously determined in equilibrium. For instance, in the case of example 1, we assumed that the location l of each community was exogenous. Alternatively, we could assume that both jurisdiction location/characteristics and membership are jointly determined in equilibrium. In other words, individuals vote with their feet and their ballot (see e.g. Konishi 1996). In this subsection, we will refer to the characteristic subject to a vote as the location.

In this scenario, the population is partitioned into (measurable) groups $(J_p)_{p\in P}$ and each group J_p chooses the location l of its jurisdiction according to an exogenously specified location function $\Lambda : 2^I \rightarrow L$ (e.g. Condorcet winner or social optimum). We can define a partition equilibrium as a partition $(J_p)_{p\in P}$ such that for all $p \in P$, all individuals in J_p are members of a community located at $l = \Lambda(J_p)$ and no individual wants to join a different group. In our atomistic setup, it is reasonable to assume that no single individual has an effect on the location function $\Lambda(J)$ if $\mu(J) > 0$. If we assume furthermore that Λ respects unanimity, i.e. $\Lambda(J) \in \cap_{j\in J} C^*(j)$ whenever $\cap_{j\in J} C^*(j) \neq \emptyset$, then one can easily see that for all partition equilibrium $(J_p)_{p\in P}$, one can construct a strategy profile σ as follows: $i \in J_p \Longrightarrow \sigma_i = \Lambda(J_p)$, and this strategy profile is a Nash equilibrium. More precisely, partition equilibria are Nash equilibria which are consistent with the location function Λ.

However, our welfare analysis cannot be simply transposed to this equilibrium concept because the notion of local optimality has to be redefined to integrate the location function Λ. Adapting the welfare analysis would warrant a paper in itself. Nevertheless, we mention one important case in which our analysis can be easily transposed: suppose that $\Lambda(J)$ is the socially optimal location for every $J \subset I$. Then an optimal strategy profile in the previous sense is consistent with Λ by construction, and as such, it is an optimal partition. If furthermore W is logarithmic, then there exists a socially optimal Nash equilibrium, and since it is consistent with Λ, it is a partition equilibrium:

Proposition 9 *If the location function Λ selects a socially optimal location and if W is logarithmic, under the condition of corollary 1, there exists a socially optimal partition equilibrium. Conversely, if the distribution of preferences is connected in the sense of definition 1, and if $(J_p)_{p\in P}$ is a partition equilibrium such that $\mu(J_p)\, W'(\mu(J_p))$ is not constant across groups, then $(J_p)_{p\in P}$ is not socially optimal.*

Proof. If $(J_p)_{p\in P}$ is a partition equilibrium, then as argued above, it is equivalent to a Nash equilibrium σ. If $\mu(J_p)\, W'(\mu(J_p))$ is not constant across groups, then $m(c,\sigma)\, W'(m(c,\sigma))$ is not constant across communities and from proposition 4, it is not socially optimal. ■

Notice that the case of a socially optimal location functions encompass the voting model in which $C = [0, 1]$, Λ selects the condorcet winner of the group, the population is uniformly distributed on $[0, 1]$ and $V_i(l) = V(|l - i|)$ for some concave function V with a peak at 0 (Alesina and Spolaore 1997, Jehiel and Schotchmer 2001).

8 Concluding Remarks

This paper builds a model of group formation in which individuals trade off joining a community which corresponds to their idiosyncratic preferences with the benefits from being in a community of the right size. We show that if group externalities are anonymous and if individual vote with their feet but take the characteristics of the community they join as given, then the social optimality of the equilibria depends on the shape of group externalities. Whether the latter increase faster or slower than logarithmically determines whether free mobility leads to excessive fragmentation or agglomeration. These results hold for any distribution of preferences and any set of available communities. Our results can be explained by a simple Pigouvian argument which amounts to compare the emigration and the immigration externalities an individual imposes on society.

The migration externality on which our analysis is based is essentially an individual concept which has no canonical equivalent for group behavior. An interesting direction for further research would be to analyze how the shape of economies of scale affect the set of cooperative equilibria in group formation games since the existing game theoretic literature has explored a somewhat orthogonal direction by focusing on the set of available communities and the distribution of preferences.[11]

As shown in section 7, our welfare analysis cannot be readily transposed to the case in which citizens vote on the characteristics/location of their community once membership is determined. Indeed, the nature of migration externalities is radically different in this case: by moving to a different community, an individual affects not only the size of the community but also their characteristics/location, which has an impact on all other members. The latter effect is likely to depend both on the location function Λ and on the distribution of preferences within the community.

9 Appendix

We first prove remark 1 and 2. From definition 3, if there exists $c \in C(p, V_i)$ such that $c \notin A(\sigma)$ for a positive mass of individuals, then $p_\emptyset > -\infty$. From assumption 2, for all such individuals, $c \in C^*(i)$. Hence, if $C(p, V_i)$ is not single-valued, at least one of the following two equalities must hold:

$$
\begin{aligned}
V_i(c_1) + p(c_1) &= V_i(c_2) + p(c_2), \\
V_i(c_1) + p(c_1) &= \max V_i + p_\emptyset,
\end{aligned}
$$

for some $c_1, c_2 \in A(\sigma)$ with $c_1 \neq c_2$. From assumption 1, for a given c_1 and c_2, the set of individuals for which either of the two equalities above is

[11] An intermediate preferences condition on V_i in Greenberg and Weber 1986 and Demange 1994 and the unidimensionality of C in Greenberg and Weber 1993 and Haimanko, Lebreton and Weber 2004.

satisfied is of measure 0. Since $A(\sigma)$ is countable, $C(p, V_i)$ is single valued for almost all i, which proves remark 1.

To prove remark 2, suppose σ is consistent with p and p'. For all $V \in \mathbb{R}^C, \delta \in \mathbb{R}$ and $\varepsilon > 0$, we denote $E(V, \delta, \varepsilon) = \{i \in B(V, \varepsilon) : p'(\sigma_i) - p(\sigma_i) = \delta\}$. From definition 3, p can take only a countable number of values, so for almost all $i \in B(V, \varepsilon)$, there exists $\delta \in \mathbb{R}$ such that $i \in E(V, \delta, \varepsilon)$ and $\mu(E(V, \delta, \varepsilon)) > 0$. Let $V \in \{V_i : i \in I\}$. Since $\mu(B(V, \varepsilon)) > 0$ for all $\varepsilon > 0$, from what precedes, for all $\varepsilon > 0$ there exists $\delta_\varepsilon, \delta'_\varepsilon$ such that $\mu(E(V, \delta_\varepsilon, \varepsilon)) > 0$ and $\mu(E(V, \delta'_\varepsilon, \varepsilon)) > 0$ so for almost all $i \in E(V, \delta_\varepsilon, \varepsilon)$ and almost all $j \in E(V, \delta'_\varepsilon, \varepsilon)$,

$$\begin{aligned} V_j(\sigma_i) + p(\sigma_i) + \varepsilon &\geq V_i(\sigma_i) + p(\sigma_i) \geq V_i(\sigma_j) + p(\sigma_j) \geq V_j(\sigma_j) + p(\sigma_j) - \varepsilon, \\ V_j(\sigma_j) + p'(\sigma_j) &\geq V_j(\sigma_i) + p'(\sigma_i), \end{aligned}$$

which implies

$$p(\sigma_i) - p'(\sigma_i) \geq p(\sigma_j) - p'(\sigma_j) - 2\varepsilon.$$

A symmetric argument gives

$$p(\sigma_i) - p'(\sigma_i) \leq p(\sigma_j) - p'(\sigma_j) + 2\varepsilon,$$

which shows that $\lim_{\varepsilon \to 0} |\delta_\varepsilon - \delta'_\varepsilon| = 0$. Moreover, since V_i is bounded over $I \times C$ and σ is consistent with p and p', δ_ε is bounded as $\varepsilon \to 0$ so there exists a sequence $\varepsilon_n \to 0$ such that δ_{ε_n} converges.

Let $\Phi : \mathbb{R}^C \rightrightarrows \mathbb{R}$ be defined as $\delta \in \Phi(V)$ if there exists a sequence $\varepsilon_n \to 0$ and $\delta_n \to \delta$ such that $\mu(E(V, \delta_n, \varepsilon_n)) > 0$. From what precedes, for all $V \in \{V_i : i \in I\}$, $\Phi(V)$ is single-valued. By construction, it is continuous for the topology of uniform convergence on $\{V_i : i \in I\}$. To conclude the proof, let $\Gamma : [0,1] \rightarrow \{V_i : i \in I\}$ be a continuous path. Then $\Phi \circ \Gamma$ defines a continuous function. Since $\Phi(V)$ can take at most a countable number of values, $\Phi \circ \Gamma$ must be constant.

9.1 Proof of lemma 1

Since $\mu(I)$ is finite, $A(\sigma)$ is countable so $A(\sigma) = \{c_n : 1 \leq n < N\}$ for some $N \in \mathbb{N} \cup \{+\infty\}$. For all i, we can assimilate V_i to a vector $\left(V_i(c_\emptyset), (V_i(c_n))_{1 \leq n < N}\right)$ with the convention that the first coordinate is $-\infty$ if $\sigma_i \in A(\sigma)$ and $V_i(\sigma_i)$ otherwise. Let $\Delta(\{c_\emptyset\} \cup A(\sigma))$ be the set of probabilities over $\{c_\emptyset\} \cup A(\sigma)$. Likewise, each strategy σ_i can be assimilated as a degenerate element of $\Delta(\{c_\emptyset\} \cup A(\sigma))$ with the convention that the probability of $c_\emptyset$, denoted $\sigma_i(c_\emptyset)$, is 1 if $\sigma_i \in C^*(i) \setminus A(\sigma)$ and the probability of c_n, denoted $\sigma_i(c_n)$, is 1 if $\sigma_i = c_n$. If we allow each σ_i to be any element of $\Delta(\{c_\emptyset\} \cup A(\sigma))$ (i.e. possibly non degenerate) instead, the relaxed maximization program becomes

$$\max_{\substack{\sigma \in (\Delta(\{c_\emptyset\} \cup A(\sigma)))^I : \\ \forall n : 1 \leq n < N, \ \int_{i \in I} \sigma_i(c_n) d\mu(i) = m(c_n, \sigma^p)}} \int_{i \in I} \left(\sum_{c \in \{c_\emptyset\} \cup A(\sigma)} \sigma_i(c) V_i(c) \right) d\mu(i) + \tag{19}$$

$$\sum_{1 \leq n < N} m(c_n, \sigma^p) W(m(c_n, \sigma^p)) + \left[\mu(I) - \sum_{1 \leq n < N} m(c_n, \sigma^p)\right] \lim_{m \to 0} W(\tag{20}$$

which is a a standard linear program since the term in (20) is constant.

By linearity, the solution set of the relaxed program is a convex subset of $(\Delta(\{c_\emptyset\} \cup A(\sigma)))^I$ whose extreme points (see Rockafellar 1970) σ are such that for almost all i, σ_i is a degenerate probability. These points are solutions to the original problem. If $\lim_{m \to 0} W(m)$ is a lower bound of W, then under our notation, $\sigma_i = c_\emptyset$ implies $\sigma_i \in C^*(i)$. In this case, if the solution is not unique, it must be that for a positive mass of individuals and for some n, m, $V_i(c_n) - V_i(c_m)$ or $V_i(c_\emptyset) - V_i(c_n)$ is the same, which is impossible by assumption 1. If $\lim_{m\to 0} W(m)$ is not a lower bound of W, then from assumption 2, $|C|$ is finite and the preceding argument hold since $c_\emptyset$ can represent only a finite number of communities.

Let σ be a strategy consistent with a potential p. The scalar $p(c_n)$ is the multiplier of the linear constraint $\int_{i\in I}(e_n.\sigma_i)\,d\mu(i) = m(c_n, \sigma^p)$ and $p_\emptyset$ is the multiplier for the linear constraint $\int_{i\in I}(e_\emptyset.\sigma_i)\,d\mu(i) = m\left(c_\emptyset, \sigma^p\right)$. The complementary slackness are satisfied by construction, and since σ is consistent with p, each σ_i maximizes the Lagrangian within $\Delta(\{c_\emptyset\} \cup A(\sigma))$. Since the problem is convex, a solution of the Kuhn and Tucker conditions is a maximum (see e.g. theorem 1 in chapter 8 of Luenberger 1969).

9.2 Proof of lemma 2

Let $i \in I$ and suppose that σ_i violates the definition of ε-optimality at σ. We want to show that for almost all such i, $\sigma_i \notin C(p^*_\sigma, V_i)$. From remark 1, we can assume that $C(p^*_\sigma, V_i)$ is single-valued. Then there exists $\varepsilon_n \to 0$ and $c \in C$ such that $S(c, \sigma_{-i}) > S(\sigma)$ under $\mu_i^{\varepsilon_n}$. Suppose first that $c \in A(\sigma)$ and $\sigma_i \in A(\sigma)$, then implies

$$\begin{aligned} & \varepsilon_n V_i(c) + (m(c,\sigma) + \varepsilon_n) W(m(c,\sigma) + \varepsilon_n) + m(\sigma_i, \sigma) W(m(\sigma_i, \sigma)) \\ > \; & \varepsilon_n V_i(\sigma_i) + m(c,\sigma) W(m(c,\sigma)) + (m(\sigma_i,\sigma) + \varepsilon_n) W(m(\sigma_i, \sigma) + \varepsilon_n). \end{aligned}$$

which can be rewritten as

$$V_i(c) + m(c,\sigma)\frac{W(m(c,\sigma)+\varepsilon_n) - W(m(c,\sigma))}{\varepsilon_n} + W(m(c,\sigma)+\varepsilon_n) > \tag{21}$$
$$V_i(\sigma_i) + m(\sigma_i,\sigma)\frac{W(m(\sigma_i,\sigma)+\varepsilon_n) - W(m(\sigma_i,\sigma))}{\varepsilon_n} + W(m(\sigma_i,\sigma)+\varepsilon_n).$$

Taking the limit, by definition of p^*_σ we get

$$V_i(c) + p^*_\sigma(c) \geq V_i(\sigma_i) + p^*_\sigma(\sigma_i). \tag{22}$$

From (21), $c \neq \sigma_i$, and since $C(p^*_\sigma, V_i)$ is single-valued, (22) implies that $\sigma_i \notin C(p^*_\sigma, V_i)$.

If $c \notin A(\sigma)$ or $\sigma_i \notin A(\sigma)$ and $\lim_{m\to 0} W(m)$ is finite, one can easily check that the reasoning above holds unchanged. If $\lim_{m\to 0} W(m) = -\infty$ and $c \notin A(\sigma)$, then $\sigma_i \notin A(\sigma)$. If $\sigma_i \notin A(\sigma)$, then $p^*_\sigma(\sigma_i) = -\infty$ so $\sigma_i \notin C(p^*_\sigma, V_i)$. Finally, if $\lim_{m\to 0} W(m) = +\infty$, then for all $c' \notin A(\sigma)$, $p^*_\sigma(c') = +\infty$. So from assumption 2, for σ to be consistent with p^*_σ, it must be that $A(\sigma) = C$.

Reciprocally, let $i \in I$ be such that there exists $\epsilon > 0$ such that for all $c \in C$ and all $\varepsilon \in]0,\epsilon[$, under the measure μ_i^ε, $S(\sigma) \geq S(c, \sigma_{-i})$. We want to show that for all $c \in C$,

$$V_i(\sigma_i) + p^*_\sigma(\sigma_i) \geq V_i(c) + p^*_\sigma(c). \tag{23}$$

If $c \in A(\sigma)$ and $\sigma_i \in A(\sigma)$, then by assumption, for all $\varepsilon \in]0,\epsilon[$:

$$\varepsilon V_i(c) + (m(c,\sigma)+\varepsilon) W(m(c,\sigma)+\varepsilon) + m(\sigma_i,\sigma) W(m(\sigma_i,\sigma)) \leq$$
$$\varepsilon V_i(\sigma_i) + m(c,\sigma) W(m(c,\sigma)) + (m(\sigma_i,\sigma)+\varepsilon) W(m(\sigma_i,\sigma)+\varepsilon),$$

which can be rewritten as

$$
\begin{aligned}
V_i(c) + m(c,\sigma)\frac{W(m(c,\sigma)+\varepsilon) - W(m(c,\sigma))}{\varepsilon} + W(m(c,\sigma)+\varepsilon) \leq & \quad (24) \\
V_i(\sigma_i) + m(\sigma_i,\sigma)\frac{W(m(\sigma_i,\sigma)+\varepsilon) - W(m(\sigma_i,\sigma))}{\varepsilon} + W(m(\sigma_i,\sigma)+\varepsilon).
\end{aligned}
$$

By letting $\varepsilon \to 0$, we get (23).

If $c \notin A(\sigma)$ or $\sigma_i \notin A(\sigma)$ and $\lim_{m\to 0} W(m)$ is finite, one can easily check that the reasoning above holds unchanged. If $\lim_{m\to 0} W(m) = -\infty$ and $\sigma_i \notin A(\sigma)$, the left hand-side of (24) goes to $-\infty$ as $\varepsilon \to 0$ so (24) is violated as $\varepsilon \to 0$ for any $c \in A(\sigma)$. Therefore, $\sigma_i \in A(\sigma)$ and (23) trivially holds for all $c \notin A(\sigma)$. Finally if $\lim_{m\to 0} W(m) = +\infty$, then ε-optimality together with assumption 2 implies that $A(\sigma) = C$.

9.3 Proof of lemma 3

Let σ be a local optimum. If σ is not an ε-optimum, there must be a set $J \subset I$ such that $\mu(J) > 0$ and for all $i \in J$, σ_i violates ε-optimality at σ. From lemma 2, this means that for all $i \in J$, there exists $c_i \in C$ such that

$$V_i(c_i) + p^*_\sigma(c_i) > V_i(\sigma_i) + p^*_\sigma(\sigma_i). \quad (25)$$

Necessarily, there exists $J_o \subset J$ such that $\mu(J_o) > 0$ and

$$
\begin{aligned}
(i) \text{ either } \forall i &\in J_o, \sigma_i \in A(\sigma) \text{ or } \forall i \in J_o, \sigma_i \notin A(\sigma), \\
(ii) \text{ either } \forall i &\in J_o, c_i \in A(\sigma) \text{ or } \forall i \in J_o, c_i \notin A(\sigma).
\end{aligned}
$$

Moreover, since $A(\sigma)$ is countable, we can assume that if $\sigma_i \in A(\sigma)$ (resp. $c_i \in A(\sigma)$), for all $i \in J_o$, $\sigma_i = \sigma_o$ (resp. $c_i = c_o$) for some fixed $\sigma_o \in A(\sigma)$ (resp. $c_o \in A(\sigma)$).

Suppose first that for all $i \in J_o$, $\sigma_i = \sigma_o$ and $c_i = c_o$ for some $\sigma_o, c_o \in A(\sigma)$. Let $J_\varepsilon \subset J_o$ be such that $\mu(J_\varepsilon) = \varepsilon$. Since σ is a local optimum, if all individuals in J_ε migrate from σ_o to c_o, social welfare must weakly decrease, i.e.

$$\begin{aligned} & \int_{i \in J_\varepsilon} V_i(c_o)\, d\mu(i) + (m(c_o,\sigma) + \varepsilon) W(m(c_o,\sigma) + \varepsilon) \\ & + (m(\sigma_o,\sigma) - \varepsilon) W(m(\sigma_o,\sigma) - \varepsilon) \\ \leq & \int_{i \in J_\varepsilon} V_i(\sigma_o)\, d\mu(i) + m(c_o,\sigma) W(m(c_o,\sigma)) + m(\sigma_o,\sigma) W(m(\sigma_o,\sigma)), \end{aligned}$$

which can be rewritten as

$$\begin{aligned} & \int_{i \in J_\varepsilon} \begin{pmatrix} V_i(c_o) + W(m(c_o,\sigma) + \varepsilon) \\ +m(c_o,\sigma) \left[\frac{W(m(c_o,\sigma)+\varepsilon) - W(m(c_o,\sigma))}{\varepsilon}\right] \end{pmatrix} d\mu(i) \\ \leq & \int_{i \in J_\varepsilon} \begin{pmatrix} V_i(\sigma_o) + W(m(\sigma_o,\sigma) - \varepsilon) \\ +m(\sigma_o,\sigma) \left[\frac{W(m(\sigma_o,\sigma)) - W(m(\sigma_o,\sigma)-\varepsilon)}{\varepsilon}\right] \end{pmatrix} d\mu(i), \end{aligned}$$

and finally

$$\int_{i \in J_\varepsilon} \left[V_i(\sigma_o) + p^*_\sigma(\sigma_o) - V_i(c_o) - p^*_\sigma(c_o)\right] d\mu(i) \geq f(\varepsilon), \tag{26}$$

where

$$\begin{aligned} f(\varepsilon) = & \left[W(m(c_o,\sigma) + \varepsilon) - W(m(c_o,\sigma)) - \varepsilon W'(m(c_o,\sigma))\right] \\ & - \left[W(m(\sigma_o,\sigma) + \varepsilon) - W(m(\sigma_o,\sigma)) - \varepsilon W'(m(\sigma_o,\sigma))\right], \end{aligned}$$

and since W is differentiable, $f(\varepsilon) = o(\varepsilon)$.[12] Observe that the set J_s of individuals in J_o such that (25) is satisfied with some slack s is of positive measure for some $s > 0$. Hence, for ε small enough one can choose $J_\varepsilon \subset J_s$,

[12] The notation $f(\varepsilon) = o(\varepsilon)$ means that $\frac{f(\varepsilon)}{\varepsilon} \to 0$ as $\varepsilon \to 0$.

so the integrand in (26) is bounded above by $-s$ for all $i \in J_\varepsilon$. From what precedes, (26) is violated as $\varepsilon \to 0$, a contradiction.[13]

Suppose now that for all $i \in J_o$, $\sigma_i \notin A(\sigma)$. Necessarily, $|C|$ is not finite so from assumption 2, $\lim_{m\to 0} W(m) < +\infty$. Local optimality implies that $\lim_{m\to 0} W(m) > -\infty$ (otherwise $S(\sigma) = -\infty$) and for almost all $i \in I$ such that σ_i is inactive, $\sigma_i \in C^*(i)$. Observe then that there is no $c_i \notin A(\sigma)$ which satisfies (25). Thus for all $i \in J_o$, $c_i = c_o$ for some $c_o \in A(\sigma)$. Let $J_\varepsilon \subset J_o$ be such that $\mu(J_\varepsilon) = \varepsilon$. Since σ is a local optimum, if all individuals i in J_ε migrate from their respective community σ_i to c_o, total welfare must decrease, i.e.

$$\begin{aligned} & \int_{i\in J_\varepsilon} V_i(c_o)\, d\mu(i) + (m(c_o,\sigma)+\varepsilon) W(m(c_o,\sigma)+\varepsilon) \\ \leq & \int_{i\in J_\varepsilon} V_i(\sigma_i)\, d\mu(i) + \varepsilon W(0) + m(c_o,\sigma) W(m(c_o,\sigma)), \end{aligned}$$

which can be rewritten as

$$\begin{aligned} & \int_{i\in J_\varepsilon} \begin{pmatrix} V_i(c_o) + W(m(c_o,\sigma)+\varepsilon) \\ +m(c_o,\sigma)\frac{W(m(c_o,\sigma)+\varepsilon)-W(m(c_o,\sigma))}{\varepsilon} \end{pmatrix} d\mu(i) \\ \leq & \int_{i\in J_\varepsilon} (V_i(\sigma_i) + W(0))\, d\mu(i), \end{aligned}$$

and finally

$$\int_{i\in J_\varepsilon} (V_i(\sigma_i) + W(0) - V_i(c_o) - p^*_\sigma(c_o))\, d\mu(i) \geq f(\varepsilon), \tag{27}$$

where as explained earlier, since W is differentiable, $f(\varepsilon) = o(\varepsilon)$. For the same reason as in the first case, for ε small enough one can choose J_ε such

[13] This comes from the fact that the set J_s of individuals in J_o such that (25) is satisfied with some slack s is of positive measure for some $s > 0$. For $\varepsilon > 0$, one can choose J_ε as a subset of J_s.

that the integrand in (27) is bounded above by some negative constant for all $i \in J_\varepsilon$, a contradiction.

Finally, if for all $i \in J_o$, $c_i \notin A(\sigma)$ and $\sigma_i = \sigma_o$ for some $\sigma_o \in A(\sigma)$, (25) implies that $\lim_{m \to 0} W(m) > -\infty$. If $\lim_{m \to 0} W(m) = +\infty$, local optimality and assumption 2 implies that $A(\sigma) = C$, a contradiction. If $\lim_{m \to 0} W(m)$ is finite, the same reasoning as before shows that local optimality implies

$$\int_{i \in J_\varepsilon} \left(V_i(\sigma_o) + p^*_\sigma(\sigma_o) - V_i(c_i) - W(0)\right) d\mu(i) \geq f(\varepsilon), \tag{28}$$

for some $f(\varepsilon) = o(\varepsilon)$ and for some J_ε such that $\mu(J_\varepsilon) = \varepsilon$ and for all $i \in J_\varepsilon$, the integrand of (28) is bounded above by some negative constant, a contradiction.

Reciprocally, let σ be ε-optimal and let $\sigma' \in A(\sigma)^I$. For all $c, c' \in A(\sigma) \cup A(\sigma')$, if we denote $J(c, c')$ the set of individuals i such that $\sigma_i = c$ and $\sigma'_i = c'$, $J(c, c_\emptyset)$ the set of individuals such that $\sigma_i = c$ and $\sigma'_i \notin A(\sigma) \cup A(\sigma')$, $J(c_\emptyset, c)$ the set of individuals such that $\sigma_i \notin A(\sigma) \cup A(\sigma')$ and $\sigma'_i = c$, we have[14]

$$\begin{aligned} m(c, \sigma') - m(c, \sigma) &= \mu(J(c_\emptyset, c)) - \mu(J(c, c_\emptyset)) \\ &\quad + \sum_{c' \in A(\sigma'), c' \neq c} \left(\mu(J(c', c)) - \mu(J(c, c'))\right), \end{aligned} \tag{29}$$

$$m(c, \sigma') = \mu(J(c, c)) + \mu(J(c_\emptyset, c)) + \sum_{c' \in A(\sigma'), c' \neq c} \mu(J(c', c)) \tag{30}$$

In what follows, we use the notational convention $mW(m) = 0$ if $m = 0$

[14] Notice that we can ignore the individuals who are in active communities both under σ and under σ' because ε-optimality implies that $\sigma_i \in C^*(i)$ and since we want to majorate $S(\sigma') - S(\sigma)$, we can assume w.l.o.g. that $\sigma'_i \in C^*(i)$.

even if $\lim_{m\to 0} W(m) = \pm\infty$.[15] We have

$$\begin{aligned}
& S(\sigma') - S(\sigma) && (31)\\
= & \sum_{c,c'\in A(\sigma)\cup A(\sigma')} \int_{i\in J(c,c')} \left[V_i(c') + W(m(c',\sigma')) - V_i(c) - W(m(c,\sigma))\right] d\mu(i) \\
& + \sum_{c\in A(\sigma)\cup A(\sigma')} \int_{i\in J(c,c_\emptyset)} \left[V_i(\sigma_i') + \lim_{m\to 0} W(m) - V_i(c) - W(m(c,\sigma))\right] d\mu(i) \\
& + \sum_{c'\in A(\sigma)\cup A(\sigma')} \int_{i\in J(c_\emptyset,c')} \left[V_i(c') + W(m(c',\sigma')) - V_i(\sigma_i) - \lim_{m\to 0} W(m)\right] d\mu(i)
\end{aligned}$$

For all $c' \in A(\sigma) \cup A(\sigma')$, the summand of the last term of the right hand-side of (31) can be rewritten ass:

$$\begin{aligned}
& \int_{i\in J(c_\emptyset,c')} \left[V_i(c') + W(m(c',\sigma')) - V_i(\sigma_i) - \lim_{m\to 0} W(m)\right] d\mu(i) \quad (32)\\
= & \int_{i\in J(c_\emptyset,c')} \left[V_i(c') + W(m(c',\sigma)) - V_i(\sigma_i) - \lim_{m\to 0} W(m)\right] d\mu(i) \\
& + \mu(J(c_\emptyset,c'))\left[W(m(c',\sigma')) - W(m(c',\sigma))\right]
\end{aligned}$$

Using (29), the summand in (31) for $c = c'$ can be rewritten as:

$$\begin{aligned}
& \int_{i\in J(c,c)} \left[W(m(c,\sigma')) - W(m(c,\sigma))\right] d\mu(i) && (33)\\
= & \ \mu(J(c,c))\left[W(m(c,\sigma')) - W(m(c,\sigma))\right] \\
& m(c,\sigma) \begin{pmatrix} \mu(J(c_\emptyset,c)) - \mu(J(c,c_\emptyset)) + \\ \sum_{c'\in A(\sigma'),c'\neq c} \mu(J(c',c)) - \mu(J(c,c')) \end{pmatrix} W'(m(c,\sigma)) \\
& - m(c,\sigma)\left[m(c,\sigma') - m(c,\sigma)\right] W'(m(c,\sigma)).
\end{aligned}$$

[15] Notice that if $\lim_{m\to 0} W(m) = \pm\infty$, ε-optimality (together with assumption 2 in the $+\infty$ case) implies that all individuals are in active communities. Since we want to majorate $S(\sigma') - S(\sigma)$, we can ignore deviations to empty groups so in what follows, $\lim_{m\to 0} W(m)$ is finite.

For $c \neq c'$, the summand in (31) can be rewritten as follows,

$$\begin{aligned}
&\int_{i \in J(c,c')} \left[V_i(c') + W(m(c',\sigma')) - V_i(c) - W(m(c,\sigma))\right] d\mu(i) \quad (34) \\
= &\int_{i \in J(c,c')} \left[V_i(c') + W(m(c',\sigma)) - V_i(c) - W(m(c,\sigma))\right] d\mu(i) \\
&+\mu(J(c,c'))\left[W(m(c',\sigma')) - W(m(c',\sigma))\right].
\end{aligned}$$

Substituting (32), (33) and (34) in (31), and regrouping the terms in $J(c, c_\emptyset)$, $J(c_\emptyset, c')$ and $J(c, c')$, $S(\sigma') - S(\sigma)$ can be rewritten as

$$\begin{aligned}
&\sum_{c \in A(\sigma) \cup A(\sigma')} \int_{i \in J(c,c_\emptyset)} \begin{bmatrix} V_i(\sigma_i') + \lim_{m \to 0} W(m) - V_i(c) \\ -W(m(c,\sigma)) - m(c,\sigma) W'(m(c,\sigma)) \end{bmatrix} d\mu(i) \qquad (35) \\
&+ \sum_{c' \in A(\sigma) \cup A(\sigma')} \int_{i \in J(c_\emptyset,c')} \begin{bmatrix} V_i(c') + W(m(c',\sigma)) + m(c,\sigma) W'(m(c,\sigma)) \\ -V_i(\sigma_i) - \lim_{m \to 0} W(m) \end{bmatrix} d\mu(i) \\
&+ \sum_{\substack{c,c' \in A(\sigma) \cup A(\sigma') \\ c' \neq c}} \int_{i \in J(c,c')} \begin{bmatrix} V_i(c') + W(m(c',\sigma)) + m(c',\sigma) W'(m(c',\sigma)) \\ -V_i(c) - W(m(c,\sigma)) - m(c,\sigma) W'(m(c,\sigma)) \end{bmatrix} d\mu(i) \\
&+ \sum_{c' \in A(\sigma) \cup A(\sigma')} \mu(J(c_\emptyset, c'))\left[W(m(c',\sigma')) - W(m(c',\sigma))\right] \\
&\sum_{c,c' \in A(\sigma) \cup A(\sigma'), c' \neq c} \mu(J(c,c'))\left(W(m(c',\sigma')) - W(m(c',\sigma))\right) \\
&+ \sum_{c \in A(\sigma) \cup A(\sigma')} \mu(J(c,c))\left[W(m(c,\sigma')) - W(m(c,\sigma))\right] \\
&- \sum_{c \in A(\sigma) \cup A(\sigma')} \left[m(c,\sigma') - m(c,\sigma)\right] m(c,\sigma) W'(m(c,\sigma)).
\end{aligned}$$

The integrands in the first three sums above is equal to $V_i(\sigma_i) + p^*_\sigma(\sigma_i) - V_i(\sigma_i') - p^*_\sigma(\sigma_i')$, which is non positive by ε-optimality. Using (30), the other four terms simplify as follows:

$$\begin{aligned}
& \sum_{c\in A(\sigma)\cup A(\sigma')} \begin{pmatrix} \mu\left(J\left(c_{\emptyset},c\right)\right)+\mu\left(J\left(c,c\right)\right) \\ +\sum_{c'\in A(\sigma),c'\neq c}\mu\left(J\left(c',c\right)\right) \end{pmatrix} \left(W\left(m\left(c,\sigma'\right)\right)-W\left(m\left(c,\sigma\right)\right)\right) \\
& -\sum_{c\in A(\sigma)\cup A(\sigma')} \left[m\left(c,\sigma'\right)-m\left(c,\sigma\right)\right] m\left(c,\sigma\right) W'\left(m\left(c,\sigma\right)\right) \\
= & \sum_{c\in A(\sigma)\cup A(\sigma')} \begin{pmatrix} m\left(c,\sigma'\right)\left(W\left(m\left(c,\sigma'\right)\right)-W\left(m\left(c,\sigma\right)\right)\right) \\ -\left(m\left(c,\sigma'\right)-m\left(c,\sigma\right)\right) m\left(c,\sigma\right) W'\left(m\left(c,\sigma\right)\right) \end{pmatrix} \\
= & \sum_{c\in A(\sigma)\cup A(\sigma')} \Phi\left(m\left(c,\sigma\right),m\left(c,\sigma'\right)\right),
\end{aligned}$$

where $\Phi\left(m_1,m_2\right)=m_2\left(W\left(m_2\right)-W\left(m_1\right)\right)-m_1\left(m_2-m_1\right)W'\left(m_1\right)$. Simple calculus yields that

$$\frac{\partial\Phi}{\partial m_2}=W\left(m_2\right)+m_2W'\left(m_2\right)-W\left(m_1\right)-m_1W'\left(m_1\right).$$

Since $\Phi\left(m,m\right)=0$, this shows that Φ is non negative for all $m_1,m_2>0$ when $W\left(m\right)+mW'\left(m\right)$ is weakly decreasing, i.e. when $mW\left(m\right)$ is concave. In this case, from what precedes, $S\left(\sigma\right)\geq S\left(\sigma'\right)$.

If $W\left(x\right)+xW'\left(x\right)$ is K-Lipschitz continuous for some $K>0$, then $\left|\Phi\left(x,y\right)\right|\leq\frac{K}{2}\left|y-x\right|^2$ so

$$\left|\sum_{c\in A(\sigma)}\Phi\left(m\left(c,\sigma'\right),m\left(c,\sigma\right)\right)\right|\leq\frac{K}{2}\sum_{c\in A(\sigma)}\left|m\left(c,\sigma'\right)-m\left(c,\sigma\right)\right|^2\leq\frac{K}{2}\left(d\left(\sigma',\sigma\right)\right)^2.$$

The following example shows that the concavity of $mW\left(m\right)$ is necessary for ε-optimality to imply locally optimality.

Example 2 *Suppose* $C=[-1,1]$, *preferences are given by*

$$U_i\left(c,J\right)=-\left|c-\theta_i\right|-W\left(\mu\left(J\right)\right),$$

and individual types θ_i *are distributed uniformly on* $[0,1]$ *with density* Δ*. Then consider the strategy profile* σ^n *such that* $A(\sigma) = \{c_k^n : k = 1, .., n\}$ *where* $c_k^n = \frac{k-1/2}{n}$ *and* $M(c_k^n, \sigma^n) = \{i \in I : \theta_i \in [(k-1)/n, k/n[\}$*. One can easily check for all* n*,* σ^n *is* ε*-optimal.*[16] *By moving the individuals in* $[1/n - \varepsilon, 1/n]$ *from community* c_k^1 *to* c_k^2*, total welfare changes by*

$$\begin{aligned} & -\Delta\frac{\varepsilon^2}{2} + \begin{bmatrix} \left(\frac{\Delta}{n} + \Delta\varepsilon\right) W\left(\frac{\Delta}{n} + \Delta\varepsilon\right) + \\ \left(\frac{\Delta}{n} - \Delta\varepsilon\right) W\left(\frac{\Delta}{n} - \Delta\varepsilon\right) - 2\frac{\Delta}{n} W\left(\frac{\Delta}{n}\right) \end{bmatrix} \\ = & -\Delta\frac{\varepsilon^2}{2} + (\Delta\varepsilon)^2 \frac{\partial^2 [mW(m)]}{\partial m^2}\left(m = \frac{\Delta}{n}\right) + o\left(\varepsilon^2\right) \end{aligned}$$

Hence, if $\frac{\partial^2[mW(m)]}{\partial m^2} > 0$ *for some* m_o*, by letting* $n \to \infty$ *and* $\Delta = nm_o$*, the above quantity will be positive for* ε *sufficiently small.*

9.4 Proof of Proposition 1

Suppose first that $\lim_{m\to 0} W(m)$ is finite and is a lower bound of W. We denote $c_1^*, .., c_n^*$ the elements of C^*. Let $c_1, .., c_l \in C \backslash C^*$. Consider a partition of society in which for $k = 1..l$, community c_k has a mass m_k of members, for $p = 1..n$, community c_p^* has a mass m_p^* of members and all other communities are not active. Facing such a partition, let I_k be the set of individuals i who weakly prefer community c_k to all other communities described above and to any $c \in C^*(i) \setminus A(\sigma)$. Let I_p^* be the set of individuals i who weakly

[16] Consider an individual in a community c whose type is closer to c than to a neighoring community c' by a factor d. If this individual, carrying a social weight ε, moves to the neighboring community c', he will exert an immigration externality on the members of c' which exactly offset the emigration externality he exerts on the members of c. Moreover, he will not benefit from this migration whenever ε is sufficiently small compared to d. Hence, total welfare will decrease.

prefer community c_p^* to all other communities described above and to any $c \in C^*(i) \setminus A(\sigma)$. Finally, let I^* the set of individuals i who weakly prefer some $c \in C^*(i) \setminus A(\sigma)$ to all others communities described above. Since $\lim_{m \to 0} W(m)$ is a lower bound of W,

$$I^* \cup \left(\cup_{k=1}^{l} I_k\right) \cup \left(\cup_{p=1}^{n} I_p^*\right) = I.$$

From assumption 1, $A(\sigma)$ is countable. Together with assumption 2, this implies that the set of individuals who belong to more than one of the sets I^*, I_1,..,I_l,I_1^*,..,I_n^* is of measure 0.

Consider the map $M : [0, \mu(I)]^l \times [0, \mu(I)]^n \to [0, \mu(I)]^l \times [0, \mu(I)]^n$ such that for $k = 1, .., l$ and $p = 1, .., n$,

$$\begin{aligned} M_l(m_1, .., m_n, m_1^*, .., m_n^*) &= \mu(I_k), \\ M_p(m_1, .., m_n, m_1^*, .., m_n^*) &= \mu\left(I_p^*\right), \end{aligned}$$

where I_k and I_p^* are defined above. From what precedes, M is well defined and a fixed point of M is a Nash equilibrium. From assumption 2 and 1, M is continuous. Brower's fixed point theorem concludes the proof.

If $\lim_{m \to 0} W(m) = -\infty$, then I^* is empty and the preceding argument holds unchanged. If $\lim_{m \to 0} W(m)$ is not a lower bound of W, then from assumption 2, C is finite and one can prove existence by applying Brower's fixed point theorem to the function M from $[0, \mu(I)]^{|C|}$ to itself as defined in the first case.

9.5 Proof of proposition 3

Suppose first that C is finite. From lemma 1, we can restrict attention to strategies which are consistent with a potential. Let $\sigma(p)$ be consistent with a potential p. Let $v = \sup_{i \in I, c \in C} |V_i(c)|$, $q = \min\{p(c) : c \in A(\sigma)\}$ and $r = \max\{p(c) : c \in A(\sigma)\}$. From assumption 2, $v < \infty$. By definition of q and r, necessarily $0 \le r - q \le 2v$, and one can easily check that $\sigma(p)$ is consistent with the potential p' defined as follows: if $c \in A(\sigma)$, $p'(c) = p(c) - q$ and if $c \notin A(\sigma)$, $p'(c) = -2v$. Hence, one can restrict attention to potentials in $[-2v, 2v]^C$. From remark 1, a potential p defines a unique strategy. Moreover, from assumption 1, the size of each community is a continuous function of the potential, so $S(\sigma(p))$ is continuous in p. The extreme value theorem completes the proof.

Let us now assume that C is compact and V_i is continuous in c for all i. From what precedes, $S(c_1, .., c_n) = \max_{\sigma \in \{c_1, .., c_n\}^I} S(\sigma)$ has a solution. For any measurable partition $P = (P_1, .., P_n)$ of I, let $\sigma^P(c_1, .., c_n)$ be the strategy such that $i \in P_k \Rightarrow \sigma_i = c_i$. Since $V_i(.)$ is continuous in c and bounded over $I \times C$, the dominated convergence theorem implies that $S\left(\sigma^P(c_1, .., c_n)\right)$ is continuous in $(c_1, .., c_n)$. Therefore, $S(c_1, .., c_n)$ is upper semi-continuous. Since C is compact, $S_n = \max_{c_1, .., c_n \in C} S(c_1, .., c_n)$ exists. To complete the proof, it suffices to show that S_n is constant after some threshold.

If it is not the case, then there exists a sequence of strategies $(\sigma^n)_n$ such that for all n, σ^n is optimal among strategies such that $|A(\sigma^n)| \le n$, and the size $\underline{m}^n$ of the smallest community $\underline{c}^n$ under σ^n goes to 0 as $n \to \infty$. Since $S(\sigma^n)$ is bounded below and $\lim_{x \to 0} W(x) = -\infty$, the size $\overline{m}^n$ of the largest community $\overline{c}^n$ under σ^n is bounded away form 0, say by $b > 0$. Since W is continuous on $[b, \mu(I)]$, it is bounded on $[b, \mu(I)]$ and since W' is bounded below, $\frac{\partial[mW(m)]}{\partial m}$ is bounded below on $[b, \mu(I)]$ by some $K \in \mathbb{R}$. Let $v = \sup_{i \in I, c \in B} |V_i(c)|$. If all individuals in $\underline{c}^n$ migrate to $\overline{c}^n$, the effect of social welfare is

$$\begin{aligned}\Delta S &= \left(\overline{m}^n + \underline{m}^n\right) W\left(\overline{m}^n + \underline{m}^n\right) - \overline{m}^n W\left(\overline{m}^n\right) \\ &\quad -\underline{m}^n W\left(\underline{m}^n\right) - \int_{i \in M(\underline{c}^n, \sigma)} V_i(c) - V_i(c) \\ &\geq \underline{m}^n \left(K - W\left(\underline{m}^n\right) - 2v\right),\end{aligned}$$

which shows that $\Delta S > 0$ for n sufficiently large, which is impossible by construction of σ^n.

References

[1] Alesina, A.. and E. Spolaore, 1997, "On the Number and Size of Nations", The Quarterly Journal of Economics, Vol. 112, No. 4, 1027-1056

[2] Arnold, B., 1987, "Majorization and the Lorenz Order: A Brief Introduction". Springer-Verlag Lecture Notes in Statistics, vol. 43

[3] Bolton P. and G. Roland, 1997, "The Break-up of Nations: a Political-economy Analysis", Quarterly Journal of Economics 112: 1057-1090

[4] Brown, C. and W. Oates, 1987, "Assistance to the poor in a federal system", Journal of Public Economics, Elsevier, vol. 32(3), pages 307-330

[5] Buchanan, J. M., 1965, "An Economic Theory of Clubs", Economica, vol. 32, pp. 1-14

[6] Demange, G., 1994. "Intermediate preferences and stable coalition structures", Journal of Mathematical Economics, Elsevier, vol. 23(1), pp. 45-58

[7] Epple, D. and T. Romer, 1991, "Mobility and Redistribution", Journal of Political Economy, 99, 828-858

[8] Greenberg, J. and S. Weber, 1986, "Strong Tiebout equilibrium under restricted. preference domain", Journal of Economic Theory 38, 101-117

[9] Greenberg, J. and Weber, S., 1993, "Stable coalition structures with an unidimensional set of alternatives", Journal of Economic Theory 60: 62-82

[10] Haimanko, O., Le Breton, M. and S. Weber, 2004, "Voluntary formation of communities for the provision of public projects", Journal of Economic Theory, vol. 115

[11] Hardy, G., J. Littlewood and G. Polya, 1988, "Inequalities", Cambridge University Press

[12] Jehiel, P. and S. Scotchmer, 2001 "Constitutional Rules of Exclusion in Jurisdiction Formation", Review of Economic Studies 68, pp. 393-411

[13] Konishi, H., M. Le Breton and S. Weber, 1997a, "Pure Strategy Nash Equilibrium in a Group Formation Game with Positive Externalities", Games and Economic Behavior, 21(1-2):161–182

[14] Konishi, H., M. Le Breton and S. Weber, 1997b, "Equilibrium in a Model with Partial Rivalry", Journal of Economic Theory, vol. 72, 1997, p. 225-237.

[15] Konishi, H., Le Breton, M. and Weber, S. (1998), "Equilibrium in a Finite Local Public Goods Economy", Journal of Economic Theory, vol. 79, 224-244.

[16] Le Breton, M. and S. Weber, 2003, "The art of making everybody happy: how to prevent a secession", IMF. Staff Papers 50 (3), 403-435

[17] Luenberger, D., 1969, "Optimization by Vector Space Methods", John Wiley and Sons, Inc., New York.

[18] McGuire, M., "Group segregation and optimal jurisdictions", Journal of Political Economy, vol. 82, No. I, pp. 112-132

[19] Milchtaich, I., 1996, "Congestion games with player-specific payoff functions", Games and Economic Behavior 13 (1), pp. 111–124

[20] Rockafellar, R., 1970, "Convex Analysis", Princeton University Press. ISBN 0-691-01586-4.

[21] Wildasin, D, 1991, "Income Redistribution in a Common Labor Market Income Redistribution in a Common Labor Market", The American Economic Review, Vol. 81, No. 4, pp. 757-774