Peer Group Effects, Sorting, and Fiscal Federalism

Sam Bucovetsky Amihai Glazer

January 25, 2012

Abstract

Suppose both the rich and the poor prefer to live in communities where many of the residents are rich. This preference may lead to excessive public spending in rich communities, which want to deter immigration by the poor. If the poor prefer pooling, they may benefit from national policies restricting local taxes, or imposing centralized provision of local public goods. In a separating equilibrium, both the rich and the poor may benefit from a policy of subsidies to communities inhabited by the poor. Some voters may therefore favor central government regulation of local governments, or favor uniform provision of a local public good.

Keywords peer group effects, sorting, decentralization

JEL Classification R23 · R38 · H73

1 Introduction

Many individuals may prefer living in a community where many of the residents are rich. The motive can arise from enjoying the high taxes paid by the rich, from peer-group effects in education, or from the higher status of living in a rich community. Such preferences can lead a community to adopt policies which appeal more to the rich than to the poor. In particular, if services provided by a local government are normal goods, then a community which spends a lot on local public services, and consequently imposes high taxes, can attract the rich but not the poor.

Though it is well–known that an equilibrium can have stratification by income (even if explicit exclusion based on income is infeasible), insufficiently explored is the influence of central government policy on the possibility of such equilibria. For example, if income stratification arises in a decentralized equilibrium, a low–income majority may favor a policy of centralization which prevents such stratification. We thus offer a complementary view of federalism, which considers both policies local government may adopt to deter immigration, and policies voters may want a central government to adopt to constrain such local policies. In particular, central government constraints on local governments may determine whether in equilibrium regions are stratified by income.

We shall see that there are two types of equilibria, with a central government able to determine which arises by restricting local governments. Spending ceilings or floors imposed on lower–level governments may change a separating equilibrium into a pooling equilibrium (see section 6). Lower–income residents will always prefer a pooling equilibrium to a separating equilibrium, so these results may explain some limits in jurisdictions in which lower–income residents form a majority. But more direct intervention is possible; if decentralization generates a separating equilibrium, then a poor majority may prefer to abolish local government autonomy, as shown in section 7. Moreover, the distortions in local government policy which are needed to support a separating equilibrium are costly to the rich. In section 8 we show that equalizing intergovernmental transfers, by reducing these distortions, may benefit the rich.

2 Literature

Discussions of federalism often start with the Oates decentralization theorem (Oates 1972), which shows that when a central government is constrained to offer uniform services across jurisdictions, the choice between local and central provision depends on the size of spillovers from local public goods, and on differences in preferences for (or costs of providing) public goods between regions. An additional incentive for decentralization appears when jurisdictions differ in their distributions of income. A region with low income inequality may favor separation that reduces the high taxes and redistribution that occur under centralization. A region with high income inequality may want to separate so that the majority in the region could impose taxes on the minority. Mobility of residents among jurisdictions, however, leads to complete homogeneity across jurisdictions, even without political integration of any kind (Bolton and Roland 1997).

The desire to live near nice neighbors is mentioned by Tiebout (1956), who, however, did not pursue the implications of such preferences.[1] People who care about what types of people live in their communities may enjoy "gains from grouping" and an "associational surplus" (Fennell 2009).

Some work considers clubs where members of each type gain direct utility from the presence of the other type (Brueckner and Lee 1989). Also modeled are workers who prefer to work with persons of the same group (Becker 1957 and Arrow 1972), and constituents who prefer to be served by clerks in a government agency of the same race (Borjas 1982). The characteristics of a competitive equilibrium when firms hire workers with different skills (Berglas 1976, and McGuire 1991).

The consequences of the poor (potentially) chasing the rich have been explored in the literature. Local governments may use restrictive zoning, housing codes, and high spending on schools that raise property taxes; such policies would appeal more to the rich or other groups the municipalities want to attract (Becker and Murphy 2000). In models with a land market, minimum lot–size zoning has been explained as an exclusionary device. Hamilton (1975, 1976), Wheaton (1993), and Fernandez and Rogerson (1997) are examples. The incentives of rich jurisdictions to over–provide local public goods as an exclusionary device are analyzed by Wilson (1998). His assump-

[1] Strahilevitz (2006) pursues the idea in examining the amenities some developments provide.

tions ensure that efficiency requires homogeneous communities, and that the equilibrium is a separating one. In an important extension to Wilson (1998), Hoyt and Lee (2003) show that a rich jurisdiction may also subsidize private goods, as an exclusionary device. In contrast, our main contributions are to compare welfare of the poor and of the rich under homogeneous and heterogeneous communities, and to explore policies of a central government that the rich or the poor may favor.

A static model with voters who choose local policies, and policies which affect the inter-community migration equilibrium, is considered by Epple and Romer (1991), though they focus on intra-community redistribution rather than on public-good provision. A peer-group effect in schools, with residents voting on how much to spend on education, is studied by Epple and Romano (2003). But in their model the vote on spending is made after people choose where to live, and therefore voters cannot use spending decisions to affect the composition of the schools. We, instead, allow people to move in response to spending decisions, and so spending can be strategic.

The number of jurisdictions is taken as fixed in most of this literature. In contrast, we consider explicitly free entry of new jurisdictions. The free-entry assumption lets us more clearly relate the equilibrium policy choices with those obtaining in models of sorting under perfect competition with asymmetric information.[2]

A large literature examines entrepreneurial behavior in the local public sector. Important examples are Berglas (1976), Berglas and Pines (1981), Scotchmer and Wooders (1987), Brueckner and Lee (1989), Scotchmer (1997), and Conley and Wooders (1998). Much of that literature assumes that entrepreneurs are "small," taking the utility attained by different types of people as given. In contrast, we explicitly consider the effect of policy on people's location decisions.

The interaction of voting and migration is an important phenomenon which has been well analyzed in an extensive literature, including Ellickson (1971), Westhoff (1977), Rose–Ackerman (1979), Epple, Filimon and Romer (1984, 1993), and de Bartolome (1990). A local official may bias services with the aim of attracting people who would likely vote for the incumbent, and encourage the out-migration of political opponents. That strategy was adopted by Mayor Curley of Boston, who used wasteful redistribution to his

[2]The classic paper is Rothschild and Stiglitz (1976). Dionne, Doherty and Fombaron (2000) survey this literature.

poor Irish constituents, accompanied with incendiary rhetoric, to encourage richer citizens to emigrate from the city, thereby shaping the electorate in his favor (see Glaeser and Shleifer 2005). A model which resembles Glaeser and Schleifer's in considering how current policy affects migration and thus future policy is Brueckner and Glazer (2008). Whereas Glaeser and Shleifer (2005) focus on the incentives of vote-maximizing incumbent officials, Brueckner and Glazer (2008) consider the preferences of residents, and allow for a broader range of policies than redistribution.

Lastly, voters at the federal level may favor a minimum federal standard for local provision of a public good when the median voter in the federal electorate has preferences that can differ from the median voter in the district in which he resides. See Cremer and Palfrey (2000).

3 Assumptions

3.1 City managers

It is easiest to establish the existence of equilibria by considering not voters, but city managers. We later give conditions under which the equilibria so derived also apply under voting.

Though we assume perfect information, the analysis resembles that of competitive markets under asymmetric information. In the canonical model of competitive insurance under asymmetric information (Rothschild and Stiglitz 1976), firms try to earn positive profits from the introduction of new insurance contracts, taking as given the contracts offered by other firms. An equilibrium is a set of contracts for which no such profitable entry by new firms is possible. The assumptions here have jurisdictions' policies under private control, with free entry of new jurisdictions, and with each jurisdiction's city manager taking as given the policies offered by competing jurisdictions.

Each city manager aims is to attract residents, subject to a zero-profit condition in each jurisdiction.[3]

We assume throughout that the cost per capita of a given level of local public good provision does not vary with the number of people. This as-

[3]As Proposition 4 in the appendix proves, the equilibrium menus will be the same, whether jurisdictions are run by city managers, or by profit–maximizing entrepreneurs who can select admission fees. The menus will also be the same if public output levels are chosen by voters.

sumption of constant returns to scale in population is consistent with most of the literature (starting with Tiebout 1956).[4]

3.2 Population

The population consists of $L^1 > 0$ poor people, each with income y_1, and $L^2 > 0$ rich people, each with income $y_2 > y_1$.[5]

The average income in the population is $\bar{\lambda}$:

$$\bar{\lambda} \equiv \frac{y_1 L^1 + y_2 L^2}{L^1 + L^2}. \tag{1}$$

The preference of a type–i ($i \in \{1, 2\}$) person can be represented by a utility function $U_i(g, \lambda)$, with g the level of the public good in the jurisdiction in which the person lives, and with λ the average income in the community. As argued below, several phenomena can be represented by this reduced–form utility measure.

Exogenously set rules determine how much each resident pays in taxes to finance the provision of the public good. For example, suppose the cost of the local public sector is shared equally by all residents in a jurisdiction. Suppose also that residents care directly about the income composition of their jurisdiction, perhaps arising from concerns about status, or from peer-group effects. Thus, the utility of a resident of type i could be represented by some function $V_i(x, g, \lambda)$, where x is consumption of a numéraire private good. Then if the cost per person per unit of the local public good is c, then $U_i(g, \lambda) \equiv V_i(y_i - cg, g, \lambda)$.

Alternatively, if the cost of the public good is financed by a proportional income tax, then the tax rate t in a jurisdiction satisfies

$$t(g, \lambda) = \frac{cg}{\lambda}, \tag{2}$$

so that $U_i(g, \lambda) = V_i((1 - t(g, \lambda))y_i, g, \lambda)$, which would depend on the population composition even if the direct utility measure $V_i(\cdot, \cdot, \cdot)$ were independent of the community income level λ.

[4]Data indeed show that in much of the U.S.,housing prices equal construction costs (Glaeser and Gyourko 2003), with the supply of physical houses very elastic. And where prices are higher, the cause is zoning and regulation, consistent with our story of the poor making it difficult for the rich to segregate.

[5]Many of our results survive generalization to more than two income classes, as the technical appendix shows. However, the uniqueness of the equilibrium solution may not hold when the number of income classes exceeds two.

Since the cost for the public sector is subsumed in the utility measure $U_i(g, \lambda)$, increases in g may increase or decrease a resident's utility. We make the following "standard" assumptions about this utility measure:

1. For each income class $i \in \{1, 2\}$ and each possible community income level $\lambda \in [y_1, y_2]$ there is a unique preferred level of public output $g_i^*(\lambda)$ such that

$$\frac{\partial U_i(g, \lambda)}{\partial g} > 0 \quad \text{if} \quad \text{and} \quad \text{only} \quad \text{if} \quad g < g_i^*(\lambda).$$

2. $\frac{\partial U_i(g,\lambda)}{\partial \lambda} > 0$ for all $i \in \{1, 2\}$, $g > 0$ and $\lambda \in [y_1, y_2]$

3. For some finite $g_1^E > g_1^*(y_1)$,

$$U_1(g_1^E, y_2) = U_1(g_1^*(y_1), y_1).$$

4. If $g' > g$, and if $U_1(g, \lambda) = U_1(g', \lambda')$, then $U_2(g', \lambda') > U_2(g, \lambda)$

Assumption 3 makes sense if the poor must pay some share of the cost of the public sector, no matter how small. Then a sufficiently high level of public output would drive their private consumption to zero, and it seems reasonable to assume that no peer–group or status effect can compensate for starvation.

Assumption 4 is the usual single–crossing property. It ensures that the indifference curve of a rich person through any (p, λ) combination is flatter than the indifference curve through the same point of a poor person — when g is depicted on the horizontal axis and λ on the vertical axis. (Since indifference curves can slope down or up, "less than" here means greater in absolute value if both groups' indifference curves sloped down.) Figure 1 illustrates indifference curves for preferences which are consistent with the above assumptions. Curve IP an indifference curve of a poor person; curve IR is an indifference curve of a rich person.

Each person chooses from a menu of jurisdictions, ignoring his own influence on the income distribution in the different jurisdictions. So he chooses his most preferred jurisdiction from those available, taking as given the average income λ in each jurisdiction

Of course this average income λ must be consistent with the location choices people make. And we assume that a city manager cannot directly exclude people from a jurisdiction. Either a person's income is private information when he chooses where to live, or a legal proscription prohibits any discrimination based on income.

Since beliefs about the income distribution must be based on the actual income distribution, we define a distribution of population as **consistent** with the available menu of local public output levels, and with the perceived income distribution, if the perceived proportion of rich people in a jurisdiction equals the actual proportion, after people make their optimal location choices.

3.3 Definition of equilibrium

Our equilibrium concept resembles a competitive equilibrium in sorting models. In these models, an equilibrium is a menu of contracts offered by entrpreneurs such that no competing entrepreneur can profitably introduce a new contract, given this existing menu. Here contracts would consist of a jurisdiction providing some level of public output, and an admission fee to that jurisdiction.

We modify the above definition in two ways. First, we replace each entrepreneur with a city manager, who is required to break even on (each of) her contracts, rather than making positive profits. This zero-profit condition lets us analyze the existence of equilibrium using two-dimensional graphs. But an outcome will be an equilibrium under this modified definition if and only if it is an equilibrium when entrepreneurs are free to choose whatever admission fee they wish.[6] As in competitive insurance models, even without this restriction, competition reduces profits to zero. Second, we allow city managers to introduce (different) multiple jurisdictions. However, no cross–subsidization is allowed; a city manager cannot lose money on any single jurisdiction.

The equilibrium concept used here differs in an important respect from that in Rothschild and Stiglitz (1976). Consider the effect of entry by a new firm in the model by Rothschild and Stiglitz. Cream–skimming by a new entrant may alter the mix of consumers served by existing firms, and this cream–skimming may cause some of the contracts offered by existing firms to become unprofitable. As is usual in models of competitive behavior, new entrants are assumed not to anticipate the effect of their entry on other firms'

[6]Provided that they charge the same admission fee to all.

profits, and on the potential exit of no–longer–profitable firms.

In the model presented here, cream–skimming affects not the profits of existing firms, but the average income λ in other jurisdictions. This change will affect the consistent allocations. In other words, in the Rothschild–Stiglitz (1976) model, each firm plays Nash with regard to its strategic variables, the price–quantity pair it offers, and assumes that altering its own behavior will not change the other firms' contracts. In the current model, each city manager plays Nash in output g. The λ's are not strategic variables, but consequences of the city managers' output choices. A new entrant assumes its own behavior will not change other city managers' output choices, but will change the distribution of residents, since the residents are not strategic players but instead respond passively to city managers' choices.

So the model here is formally closer to the modification of the Rothschild–Stiglitz model made by Wilson (1977), ensuring the existence of equilibrium. Not surprisingly, the equilibrium outcomes here are more similar to those in this modified insurance model.

The timing of the model is as follows. First a central government decides which restrictions, if any, to impose on prospective city managers who set local policies. Then each city manager sets provision of the local public good in her jurisdictions, anticipating the migration responses of residents. Lastly, residents sort themselves, so that each lives in the jurisdiction he prefers (which choice depends, of course, on the location decisions of others).

4 Equilibria

Consider the usual subgame perfect Nash equilibrium. City managers anticipate the locational choices people will make. A city manager chooses to enter if and only if each of the jurisdictions she operates can attract a positive population, given the choices made by other city managers, and given the subsequent location decisions of individuals. Since the population distributions consistent with a given menu of local public output levels need not be unique, we must further restrict the population distributions which result from city managers' decisions. These formalities are discussed in section 10.

4.1 Separating equilibrium

The results are presented graphically here. Formal analyses and proofs are in the Appendix. In Figure 1, curve IP is an indifference curve of a poor person. It is constructed to be tangent to the horizontal axis; that is, it represents the highest utility a poor person can obtain if his community has no rich people. A poor person who lives in a community with only the poor has highest utility when the level of g is at point “all poor” (this is the tangency point with the horizontal axis). Anywhere above this curve represents higher utility for a poor person than points on curve IP.

The point “all rich” is constructed to satisfy the selection constraint, that no poor person prefers to move into rich jurisdictions. It lies on the right-hand side of indifference curve IP, where the average income reaches its maximum level y_2. Curve IR is the indifference curve of a rich person through point “all rich”. The horizontal dashed line in the middle of the graph is the average income of the whole population, an analogue to the “pooling line” in models of insurance.

In this figure, the indifference curve IR of the rich lies above this pooling line. Under these conditions, a separating equilibrium exists. All rich people live in communities consisting only of rich people. Each consumes at point “all rich”, on indifference curve IR. All poor people live in communities consisting only of poor people. Each consumes at point “all poor”, on indifference curve IP. Note that if these points represent the allocations, then no rich person would want to move to any newly established community which gives a consumption point below curve IR, so no community which lies between curves IR and IP is feasible. And no poor person would want to live in a community which lies below curve IP.

The poor and the rich would both prefer to live in a community which gives a consumption point in the area above curve IR. But no such community could be established—if it were, all the poor and all the rich would want to live there; so the proportion of the rich in such a community would necessarily be the same as in the population as a whole so that $\lambda = \bar{\lambda}$, the level indicated by the horizontal dashed line. Since this horizontal line lies below curve IR, entry of such a new community is impossible.

The high level of g at point “all rich” can be interpreted to mean that the rich want to provide high g (with associated high taxes) to deter the poor from moving into their communities; in the absence of such a consideration, the rich would prefer to consume somewhere to the left of point “all rich”, with less g.

4.2 Pooling equilibrium

A different equilibrium, with pooling, appears if the rich are a sufficiently large fraction of the population. Indifference curve IP in Figure 2 is the same as in Figure 1; indifference curve $IR1$ in Figure 2 is the same as indifference curve IR in Figure 1. At point "Pooling" Curve $IR2$ is tangent to the dashed horizontal pooling line. This average income level is higher than in Figure 1.

Point "Pooling" is an equilibrium. Any point above $IR2$ is infeasible: it would attract all the rich and the poor in the population, implying that the proportion of rich in each community is greater than their proportion of the nation's population — an impossibility. Any community with an allocation below indifference curve $IR2$ would be less attractive to the rich than is the community at point "Pooling". Therefore, any such community would attract no rich people, leading the poor to live in a community at point "all poor", which each finds inferior to living in a community at point "Pooling".

Notice that the rich prefer point "Pooling" to point "all rich". As long as the pooling line lies above indifference curve $IR1$ through the point "all rich", the rich will benefit from an increase in the proportion of the rich in the population.[7] And since the poor prefer point "Pooling" to point "all poor", they too benefit from an increase in the proportion of the rich in the population.

This upsetting of the separating equilibrium with a pooling equilibrium operates exactly as in adverse selection models. Unlike the result in Rothschild and Stiglitz, however, an equilibrium can exist in this situation. Pooling outcomes cannot be upset so easily by new entrants who cream–skim. A new entrant who attempts to attract only the rich residents of a mixed jurisdiction would lower the average income in the existing mixed jurisdiction. In insurance markets, such selection reduces the profits of an existing firm below zero, but it does not change the attractiveness of the existing contract to high–risk (or low–risk) customers. Here the fall in λ harms directly the poor (and rich) residents of an existing jurisdiction. In doing so, it may induce the poor, as well as the rich, to leave the existing jurisdiction. This of course would undo the new entrant's intention of attracting only the rich.

[7] Once the pooling line is below this indifference curve, as in Figure 1, further decreases in the proportion of rich people do not affect the equilibrium utility of either group.

4.3 Characterization of equilibrium

Suppose that some central planner decided on distinct public output levels g_j for some communities $1, 2, \ldots, J$, and then allowed people to sort themselves among these communities. The equilibrium just described (separating or pooling) is the unique solution to the planner's maximization under the "lexmax" criterion:[8] maximize the utility of the poor over all plans which maximize the utility of the rich.

If there are two income classes, then this lexmax solution is unique, and is the unique equilibrium possible under free entry by city managers.

If the number of income classes exceeds two, this lexmax solution is still unique. The solution still corresponds to an equilibrium when local public output levels are set by competing city managers (or entrepreneurs). With more than two income classes, however, additional equilibria may appear.

All these claims are proved in the technical appendix.

5 Voting

The model presented above assumes that city managers set local public policy, and limits individual residents' responses to voting with their feet, rather than having any say in choosing policies. But the equilibria derived here still will be equilibria, even if local residents have the final say over local spending.

Suppose then that residents of each community can vote on local spending. When voting, residents of each community take the policies as given in every other community. They also take as given the number of other communities; residents vote after city managers created communities.

Voters, however, do not take the population distribution as given. They anticipate the migration response to any policy changes they might choose. That is, voters in community i choose their local public output level g_i, taking as given the public output levels $(g_1, g_2, \ldots, g_{i-1}, g_{i+1}, \ldots, g_N)$ in other communities, and anticipating the average income levels $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ resulting from these output levels.[9]

So consider an equilibrium described in the previous section. The equilibrium will be said to be **stable** with respect to voting if no unilateral public

[8] See Bossert, Pattanaik, and Xu (1994).

[9] Formally, using a term defined in the appendix, they take the other public output levels as given, and anticipate average income levels $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ which *correspond* to the public output levels $(g_1, g_2, \ldots, g_N)$.

output change in any community makes the majority of residents of that community better off than in the original equilibrium, after migration has occurred.

The one assumption required to ensure stability is that there be several communities of each type. As mentioned above, the number of communities here is indeterminate. If some $\{(g_1, \lambda_1), (g_2, \lambda_2), \ldots, (g_N, \lambda_N)\}$ constitutes an equilibrium, then there will be an equilibrium in which each community is cloned, replacing each community with two communities, each with the same public output level and average income as the original community.

An equilibrium which is stable with respect to voting always exists (see Proposition 5, proved in the appendix). In particular, if there are at least two communities of each type in an equilibrium, then that equilibrium will be stable with respect to voting. With a pooling equilibrium, that means at least two identical communities; with a separating equilibrium that means at least two identical all–rich communities and at least two identical all–poor communities.

The key to this stability is voters' foresight. In equilibrium, no city manager can introduce a new community, with a distinct level of public output provision, which attracts some residents by offering them greater utility than they had before. If city managers cannot make any residents better off by creating communities — after people re–sort themselves — then the residents will be unable to make themselves better off by changing existing communities. As long as voters foresee the consequences of their policy changes, voters will not want to change the public output levels chosen by city managers.

6 Spending limits

The separating equilibrium has the poor potentially worse off than they would be in an equilibrium where the poor live in the same communities as the rich. In the equilibrium the poor prefer each community has a proportion of the rich and poor identical to that in the population as a whole.

Suppose then that the poor were a majority in the population, controlling a central government which could restrict local governments. In a separating equilibrium, the poor are dissuaded from living among the rich by the high level of g and associated taxes. The poor may therefore benefit from a spending limit on g. Figure 4 shows how. In the absence of a spending constraint, Figure 4, like Figure 1, shows a separating equilibrium with the rich consuming at "all rich" and the poor consuming at "all poor". In Figure 4, indifference curve IP is the same as in Figure 1—it shows the highest utility the poor can get in a community consisting only of the poor. Indifference curve $IR2$ is tangent to the horizontal line, at point T.

Now suppose a constraint is imposed (say by the central government) setting the maximum level of g at the level associated with point "Mixed". If the constraint were binding, one outcome would have all the rich live in mixed jurisdictions, and some of the poor continue to live in the homogeneous jurisdiction "all poor". The proportion of poor in the mixed community would have to be so high that the poor were indifferent between the homogeneous "all poor" and the mixed community (labeled "Mixed" in Figure 4).

This outcome, with all the rich in the community "Mixed", and the poor distributed between communities "Mixed" and "all poor" is Pareto dominated by a pooling outcome, at point T in Figure 4. That is, if the spending constraint is stringent enough, the rich would rather let all the poor in, and not have to distort the consumption of the rich.

Thus, if the spending limit is sufficiently stringent, then the equilibrium has pooling, with all communities represented by point T. The poor are better off at point T, than at the separating equilibrium in which the poor consume at point "all poor". The rich would prefer to consume above indifference curve $IR2$. But the constraint on g restricts the rich to consume to the left of the point "Mixed"; and if a rich community is to exclude the poor, the mobility of the poor restricts the rich to consume on indifference curve IP. The rich are better off in a pooling equilibrium at T. Note that if the constraint is to induce a move to pooling, then the limit on spending must be no more than the level at the point "Mixed". But the spending constraint may not actually bind on the mixed community (at T).

Though a spending limit can benefit the poor, as just seen, a small spending limit may hurt the rich without helping the poor. This is shown in Figure 3. It differs from Figure 4 in having a less stringent spending limit. If the original equilibrium (with no spending limit) were a separating equilibrium, then the outcome of a spending limit which is just binding must be the one depicted in this figure. The equilibrium in this situation would have all the rich and some of the poor live in a community with allocation at point "Mixed", and some of the poor consuming at point "all poor". The limit on spending on g hurts the rich, but does not help the poor.[10]

[10]In Lee (1993), spending limits harm bureaucrats, but only benefit voters if they are

We so far discussed caps on spending by local governments. But note that an effective cap can be imposed by setting some (nonlinear) spending requirements. Suppose the central government requires that any city which provides some service must provide at least $g_H + \Delta$. An example would be building codes for theaters in schools. The constraint would induce the rich to go below g_H, thereby effectively imposing a cap.

And floors on spending could also benefit the poor. Suppose the central government imposes a binding minimum level on g, slightly greater than that provided in "all poor" in the separating equilibrium. That minimum forces "all poor" to raise g to the new legal minimum, making "all poor" less attractive to poor people. So it shifts them to a lower indifference curve in Figure 1, indicated by the dotted curve. The city manager of the rich jurisdiction, who wants to keep out the poor, must make "all rich" less attractive, which he does by increasing g there. So the new equilibrium has both "all poor" and "all rich" provide a higher g. The new outcome has the poor indifferent between "all poor" and "all rich" — but on a lower (further right) indifference curve than before. The rich are also worse off, since the excessive spending in "all rich" has been exacerbated. The small increase in the minimal level of g has harmed all residents.

But as this binding minimum increases, the utility of the rich in "all rich" falls to the level they would get under pooling. (That is, the point "all rich" moves right, until it is on the rich people's indifference curve which is tangent to the pooling line.) If the floor on spending increases a tiny bit further, then the rich are better off giving up their attempts to exclude the poor through over–provision. The new equilibrium has pooling. So — for the same reasons that a sufficiently tight cap on spending can induce a switch to pooling — a sufficiently tight floor can induce a switch to pooling. That move from a separating to a pooling equilibrium can (sometimes) benefit the poor.[11]

sufficiently large. However, unlike the case considered here, in Lee (1993) small spending limits make everyone (voters and bureaucrats) strictly worse off .

[11] As with the maximum, we would not see the constraint actually binding, if it is effective. If the minimum works at switching the nature of the equilibrium, then all we see in equilibrium is pooling. The minimum spending constraint would bind only on "all poor", and in equilibrium no city manager wants to provide "all poor".

7 Uniform service

The second generation of fiscal federalism literature emphasizes that uniformity of public good provision differs from centralization. But uniform provision is feasible if the national government sets the policy of the local public sector. Consider then the choice between uniform provision by the central government, and the equilibrium prevailing under competitive behavior by city managers (with none of the central government regulations described in the previous section).

Consider a first, “constitutional” stage, which has residents choose whether to have (*i*) free–entry, competitive provision by many city managers, with no central government regulation, or (*ii*) uniform provision, with the level chosen by majority rule at the national level. If the poor are in the majority nationally, then the poor necessarily do better under uniformity: they set g to maximize their utility, and also enjoy the benefits of living in a community with rich people. So the poor are off than under decentralization: in a separating equilibrium they do not enjoy the benefits of living in a community with rich people, and in a pooling equilibrium they do not enjoy their preferred level of g. Whereas uniform central provision is better for the poor in this situation, it is necessarily strictly worse for the rich.

Conversely, if the rich are in the majority, they necessarily do at least as well under decentralization as under uniformity. But this preference may be weak; with a rich majority, uniform provision yields the same outcome as a pooling equilibrium under decentralization. What benefits the rich may not necessarily harm the poor here. With a rich majority, the poor may be better off, or worse off, in a separating equilibrium under decentralization than under uniform central provision. Figure 6 illustrates how the poor benefit under decentralization; Figure 5 illustrates how they suffer under decentralization.

8 Transfers among jurisdictions

The model here may also explain the popularity of transfers from higher levels of government to low–income communities. If local government policies are decentralized, and if residents are mobile, these transfers may benefit both the rich and the poor.[12]

That is, consider again a prior constitutional stage, at which a national legislature decides whether to provide transfers to communities with a large proportion of poor residents. These transfers may be financed directly by payments from richer communities, or indirectly through a national income tax.[13] In either case, the rich will be making transfers to poor people who choose to live in homogeneous communities. But the rich may still benefit from such a transfer policy, because it affects the location decisions of poor people. Transferring resources to "all poor" makes that poor–only community more attractive. Doing so means that "all rich" does not have to distort its public expenditure decision as much, in order to keep out the poor. Introducing these transfers may induce a Pareto–preferred separating equilibrium under decentralization.

In particular, suppose that peer-group benefits are a normal good, with the poor valuing them little, and the rich valuing them much. Suppose also that under decentralization the equilibrium is a separating one. A central government policy of transfers to any community for which λ is below some (low) threshold would make a poor community more attractive, thereby allowing the rich to reduce spending on g in rich jurisdictions, without inducing the poor to migrate. As an extreme example, suppose a poor person pays almost no taxes. An increase in g in a rich community will not prevent migration. But a subsidy to a poor community can deter migration.

Notice that subsidies here are directed not at poor people, but at communities inhabited exclusively by poor people. Indeed, because an increase in a poor person's income increases his willingness to pay for g and his willingness to pay for peer-group benefits, a direct subsidy hurts the rich. Instead, the rich favor a subsidy or transfer that is contingent on the individual residing in a jurisdiction populated by the poor. A subsidy to the jurisdiction itself accomplishes the goal.

Transfers to jurisdictions are common. Among members of the OECD, in the following countries more than 5 percent of government spending is for fiscal equalization: Austria, Italy, Spain, Switzerland, and Finland, with Japan the highest at 11 percent.[14]

The European Union spends about a third of its budget on transfers to

[12]Wilson (1998) discusses transfers by the central government to the jurisdictions as solving the problem of the non-existence of an equilibrium.

[13]Transfers through the income tax seem more common in practice.

[14]See Blogchliger et al. 2007).

poor regions (see *http : //europa.eu/pol/reg/index_en.htm*), with projected spending on such transfers over the period from 2007 to 2013 amounting to about 350 billion euros. Some of these transfers are to the poor countries in eastern Europe, but some go to regions in southern Italy, East Germany, Greece, Portugal, Spain, and the United Kingdom. In Canada, the federal government makes payments to less wealthy Canadian provinces; in 2010-2011, six provinces got $14.4 billion in equalization payments from the federal government. This policy is enshrined in the Constitution: "Parliament and the government of Canada are committed to the principle of making equalization payments to ensure that provincial governments have sufficient revenues to provide reasonably comparable levels of public services at reasonably comparable levels of taxation" (Subsection 36(2) of the Constitution Act, 1982). Policies in both the European Union and in Canada are consistent with our model of rich areas wishing to limit immigration from poorer regions.

In the United States, states constrain local governments in several ways. Under Dillon's Rule, in effect in 31 states, cities can exercise only those powers specifically delegated by the state.[15] Tax and expenditure limitations, such as California's well known Proposition 13, greatly constrain even rich communities from providing local public goods; such limitations appear in 46 states.[16]

Many states, sometimes but not always under court order, adopted policies to equalize school finances, so that districts with rich residents do not spend much more than districts with poor residents. It appears that some of the policies aimed not only to increase spending in poor districts, but to reduce spending in rich districts. Hoxby (2001) notes that equalization schemes have generated both leveling-up and leveling-down, with states having the most dramatic equalizing showing reduced total spending on schools: California (15 percent lower), New Mexico (13 percent lower), and Utah (10 percent lower). In terms of our model, that pattern would suggest support for the policies by the poor.

Foreign aid, aiming to limit immigration, offers another example of the transfer to the poor we are considering. Empirical studies indeed find that foreign aid reduces migration into donor countries, and that donors actively use aid to reduce immigration (Azam and Berlinschi 2008, and Bermeo and

[15]Richardson, Gough, and Puentes 2003.
[16]Mullins and Cox 1995.

Leglang 2010).

9 Conclusion

Residents, or city managers, who care about the composition of their communities may favor policies which affect the composition. They may adopt policies which are more highly favored, or less strongly opposed, by people it wants to attract. Voters at the national level may favor limiting the city's ability to adopt such discriminatory policy. And voters may favor fiscal transfers, not to individuals, but to communities. An examination of all three such approaches finds that some policies can benefit all residents (including the rich who pay higher taxes). But some policies, which can be efficient when stringent, when lax can hurt the rich without helping the poor.

10 Appendix : Analytical solution

10.1 Summary

The main goal of this appendix is to prove that a competitive equilibrium exists when entrepreneurs can set up townsites, under free entry by entrepreneurs, when people are perfectly mobile among townsites, and cannot be excluded (directly) on the basis of their type.

The familiar equilibrium to the Rothschild–Stiglitz insurance model, as modified by Wilson, must be one of these competitive equilibria. It may not be the only competitive equilibrium when there are more than two income classes. But it always is a competitive equilibrium. And every competitive equilibrium must be Pareto optimal, among the set of feasible allocations (when people are free to move among communities).

Entrepreneurs here are assumed to commit to the provision of particular public output levels in their townsites, and to let prospective residents then sort themselves.

We first define and characterize a particular set of output levels, and assignment of people to communities providing those outputs, which we call the "lexmax solution". This lexmax solution is what a planner would provide, if the planner had the exclusive right to administer townsites, but could not control (directly) the movement of people among them — if the planner's welfare criterion is to maximize the well–being of the richest group of residents[17]. Ultimately, we show that this lexmax solution must be a competitive equilibrium when there is free entry by competing entrepreneurs.

We first show that this lexmax solution is unique.

Then we show that the lexmax solution persists, even if a bunch of other townsites are available, providing different levels of output than those proposed by the planner. That is, if an arbitrary number of different new communities is created, they will not be able to attract any residents away from the planner's lexmax solution.

This result is the crucial one for characterizing competitive equilibrium, since it shows that the lexmax solution will not be upset by entry of new competitors.

Since the planner does not charge admission, we construct another artificial intermediate step, between profit–maximizing entrepreneurs and plan-

[17] extended lexicographically

ners. We call the agents in this intermediate step "city managers" : they compete to attract residents, just like private entrepreneurs would, but are constrained to make zero profits. We then show that the lexmax solution will be an outcome in this intermediate step, and then finally we show that the free entry ensures that the outcome with profit–maximizing entrepreneurs is the same as that with city managers.

10.2 crucial assumptions

Two crucial sets of assumptions are needed in the proofs below.

One is our treatment of the allocation of population among an arbitrary set of communities. Since the utility–maximizing choices by residents creates a sort of coordination game, there may be many equilibrium location patterns, for the given set of communities. We narrow down this set. First we rule out patterns which are Pareto dominated by other feasible location patterns. Second, in many location patterns, some communities would be uninhabited. In such a situation, residents must form conjectures about what would be the income composition of these uninhabited communities. People might avoid moving to an otherwise–attractive community if they feared it would be inhabited only by lower–income people. We restrict these conjectures : we assume that residents understand the single–crossing property of their own preferences, and use that understanding to infer what type of people might choose a townsite with a given output level.

The other set of assumptions concerns entry by entrepreneurs. We allow new entrepreneurs to enter with multiple new townsites : if we restricted each entrepreneur to a single townsite, then more outcomes (in addition to the lexmax solution) might survive competitive entry. But we do not permit cross–subsidization by individual entrepreneurs : they must charge non–negative admission fees for each townsite.

10.3 The Model

There are N types of individual, differing only in their income y_i. By convention, the groups are ordered in increasing order of their income :

Convention : $y_1 < y_2 < \cdots < y_N$

Each person's preferences can be represented by a utility function $V_i(g, \lambda, f)$, where g is the public output in the community in which the person lives, λ is the average income there, and f is a fee that the person may have to pay. The main assumptions made are that preferences obey the single–crossing property, and that each income class has a (finite) most–preferred public output level $g_i^*(\lambda)$, which may depend on the average income in the community :

1. For each income class $i \in \{1, 2, \cdots, N\}$ and each possible community average income $\lambda \in [y_1, y_N]$ there is a unique preferred level of public output $g_i^*(\lambda)$ such that
$$\frac{\partial V_i(g, \lambda, f)}{\partial g} > 0 \quad \text{if} \quad \text{and} \quad \text{only} \quad \text{if} \quad g < g_i^*(\lambda).$$
2. $\frac{\partial V_i(g,\lambda,f)}{\partial \lambda} > 0$ for all $i \in \{1, 2, \cdots, N\}$, $g > 0$ and $\lambda \in [y_1, y_N]$
3. $\frac{\partial V_i(g,\lambda,f)}{\partial f} < 0$
4. For each $i < N$, there is some finite $g_i^E > g_i^*(y_i)$ such that
$$V_i(g_i^E, y_{i+1}, f) = V_i(g_i^*(y_i), y_i, f).$$
5. If $g' > g$, and if $V_i(g, \lambda, f) = V_i(g', \lambda', f)$, then $V_j(g', \lambda', f) > V_j(g, \lambda, f)$ for all $j > i$.

Except when (briefly) discussing the behavior of profit–maximizing entrepreneurs, we will not need to consider the effects of the fees on people's behavior. So we can, for the most part, restrict attention to the case $f = 0$ and define

Definition : $U_i(g, \lambda) \equiv V_i(g, \lambda, 0)$

10.4 Feasible Outcomes for a Planner

If a central planner were to create communities, and decide on an output level g^i for each, but could not directly exclude any person from any community, an allocation L_i^j of people of income i to community j, and an output menu $(g^1, g^2, \cdots, g^J)$ for communities is feasible if

1. For each community j, $L_i^j > 0$ for some $i \in \{1, 2 \cdots, N\}$.

2. $\lambda^j = \frac{\sum_{i=1}^N L_i^j y_i}{\sum_{i=1}^N L_i^j}$

3. If $L_i^j > 0$, then $U_i(g^j, \lambda^j) \geq U_i(g^m, \lambda^m) \quad \forall \quad 1 \leq m \leq J$

The first restriction rules out completely empty communities : the possibility of empty communities, and the conjectures people make about the average incomes there, will have to be considered when examining competitive outcomes, but is not needed in considering the planner's problem. The third restriction incorporates the assumption that the number of people is sufficiently large that no person considers the impact of her own migration decision on average incomes in either the origin or destination community.

A plan is Pareto optimal if it is feasible, and if no feasible plan Pareto dominates it. Among the Pareto optima, we consider a particular plan, since it will be our candidate for the result of competition among developers.

10.5 separating and pooling

Proposition 1 *Suppose a plan is feasible, and arrange the communities in increasing order of their public output provision, so that $g^1 < g^2 < \cdots < g^J$. Then*

(i) If $L_i^j > 0$ and $L_n^j > 0$, with $n > i$, then $L_k^j = L_k$ for all $k \in \{i+1, i+2, \cdots, n-1\}$.

(ii) $L_i^j > 0$ and $L_i^m = 0$ for $m > j$ imply that $L_n^h = 0$ for all $h \geq m$, $n < i$.

That is, in a feasible outcome, each community gets a slice of the income distribution : types $1, 2, \ldots, I_1$ in community #1, types $i_2, i_2+1, \ldots, I_2$ in community 2, and so on, where i_j is either I_{j-1} or $I_{j-1}+1$.

10.6 The Lexmax Solution

Among all Pareto optimal plans, define the lexmax solution recursively as follows :

let C^N be the set of all plans which maximize U^N, among the set of Pareto optimal plans ; then for any set of plans C^j, define C^{i-1} as the set of all plans in C^i which maximize U^{i-1}

Definition The set of lexmax solutions is C^1.

That is, the lexmax solution is the plan which would be chosen by a planner with a **lexmax** social welfare function.[18]

10.7 The truncated economy

Take an economy with N income types, and a lexmax solution C^1 for that economy. The **truncated economy** is constructed by removing all the communities in which any type–N people locate in the candidate equilibrium, and removing all the people in those communities. So, for a given lexmax solution in the original economy, let

$$\mathcal{J}(N) \equiv \{j | L_N^j > 0\}$$

The truncated economy is a new economy, with L_i' people of income y_i, where

$$L_i' = L_i - \sum_{j \in \mathcal{J}(N)} L_i^j \quad i = 1, 2, \ldots, N$$

By construction, $L_N' = 0$, so that the truncated economy will serve as a mechanism for proving results by induction, since the truncated economy has $N - 1$ or fewer distinct income levels.

Lemma 1 *Suppose* $\{L_i^j, g^j | i = 1, 2, \ldots, N; j = 1, 2 \ldots J\}$ *is a lexmax solution, and that* $\#\mathcal{J}(N) = k$ *for this lexmax solution. Then the plan in which there are* $J - k$ *communities, with community* j *providing output* g^j*, and attracting a population* L_i^j*, for each* $i < N$*, each* $j < J - k$*, will be a lexmax solution for the truncated economy.*

Proof Since the original plan was feasible for the original economy, the "truncated plan" will be feasible for the truncated economy. Suppose, contrary to the lemma, that some other plan for the truncated economy Pareto–dominated this truncated plan.

We will show that, if this were the case, some plan for the original "un–truncated" economy Pareto dominates the original candidate equilibrium, implying a contradiction (since a lexmax solution must be Pareto optimal).

First of all, consider the utility of the highest income level in the truncated economy. Let $M < N$ denote this income level. Let K be the number of communities in this new plan for the truncated economy.

[18] as defined in Bossert, Pattanaik, and Xu (1994)

If the new plan for the truncated economy yielded exactly the same level of utility to income class M (as they got in the original candidate equilibrium), then we are done. If we add back the missing $L_i - L_i^j \quad (i \geq M, j > J-k)$ people to the truncated economy, and add k new jurisdictions to the new plan for the truncated economy, providing the output levels g^j of the communities in $\mathcal{J}(N)$ in the original un–truncated economy, then the people who have been added back in will locate in the same communities, and get the same utility levels, as in the original lexmax solution. So this adjusted plan (use the new plan for the truncated economy, and restore the communities in $\mathcal{J}(N)$) is feasible, and Pareto dominates the candidate equilibrium for the original un–truncated economy, a contradiction.

So next, assume that the new plan for the truncated economy actually yields higher utility to income level M than the original lexmax solution in the original un–truncated economy.

Now add back the missing people to the truncated economy, and add back the missing communities in $\mathcal{J}(N)$ (with their original public output levels g^j) as above. Call these "restored" communities from the original economy : $K+1, K+2, \ldots K+J-k$. (The renumbering may be necessary, since in the new plan for the truncated economy, the number of populated communities K may differ from the number of populated communities $J-k$ in the original plan for the truncated economy.) But this time, the fact that U^M has increased in the truncated economy means that any type–M people assigned to community $K+1$ will want to move, since their utility from living in community K has increased. Move all the L_M^{K+1} people of this type from community $K+1$ to community K. This move will actually increase the attraction of community K, since it raises the average income level there. So at the same time, increase g^K, so as to keep the utility of type i_K people (the poorest people now living in community K) exactly the same as it was in the new plan for the truncated community. Since U^i eventually decreases with g, it is always possible to do so.

So the overall effect of the change is to move richer people (of income y_M) into community K, and to raise g^K to keep the utility of the poorest residents of community K unchanged. By construction, this change means that no–one will want to move to or from communities $1, 2, \cdots, K-1$. And single–crossing implies that all residents of community K are at least as well off from this combination (immigration and increased g) as they were in the truncated economy, since higher income people value increases in public output relatively more.

This change may have made community K so attractive that people of income level $M+1$ (assigned to communities in $\{K+1, K+2, \ldots, K+J-k\}$) will also want to move to community K. If this is the case, move them as well to community K, again increasing g^K so as to keep constant the utility of people of income class i_K. These changes have no effect on the utility of people of income $1, 2, \ldots, i_K$, and increase strictly the utility of people of income $i_K + 1, i_K + 2, \ldots, M + 1$.

If it now turns out that people of income type $M + 2$ want to move to community K, do the same thing : move them, and increase g^K so as to keep constant utility of type i_K. Keep moving people, and increasing g^K until either : everyone of type $M+1, M+2, \ldots, N$ has been moved to community K, or people of income type i ($i > M$) don't want to move to community K.

In either case, after a finite number of moves (and accompanying public output changes), a feasible plan for the original economy has been produced, which Pareto dominates the original one.

Therefore, no new plan for the truncated economy can Pareto dominate the truncated version of the lexmax solution for the original economy. •

10.8 no splitting

Lemma 2 *In any Pareto optimum (among the set of feasible outcomes), for any income group i, $L_i^j > 0$ for exactly one community j.*[19]

Proof Start with the case $N = 2$. The single–crossing condition means that there can be at most three distinct communities inhabited in any feasible outcome.

If we have 3 communities, with community #2 mixed with both income groups, then move all the rich people from community 3 to community 2, at the same time increasing g^2 so as to keep constant $U^1(g^2, \lambda^2)$. The new plan is feasible, leaves the utility of type–1 people unchanged, and (because of single–crossing) increases the utility of type–2 people.

If we have 2 communities, with community #2 mixed (and community #1 populated only by type–1 people), then move all the poor people in community #2 to community #1, increasing g^2 so as to keep unchanged $U^1(g^2, \lambda^2)$. This change results in a feasible plan in which U^1 is unchanged and U^2 has increased.

[19] modulo cloning

Finally (for the case of $N = 2$), if we have 2 communities and community #1 is mixed, move all the rich people in community 2 to community 1, which results in a feasible plan which Pareto dominates the original one.

The proof proceeds by induction. Suppose the lemma holds for economies with $N - 1$ income classes. Now take an economy with N income classes, and consider a Pareto optimum for that economy, with J communities.

Suppose first that the richest people in this economy, those of income class N, were split between communities J and $J - 1$. Then moving all the rich people in community J to community $J - 1$, at the same time raising g^{J-1} so as to keep constant the utility of the poorest people in community $J-1$ will produce a Pareto–improvement, raising the utility of income classes $i_{J-1}, i_{J-1}+1, \ldots, N$ while keeping unchanged the utility of all poorer people.

Suppose next that community J had more than one income class residing in it, but that people of income class i_J were split between communities $J-1$ and J. Now move all the type–i_J people from community J to community $J - 1$, raising g^{J-1} to keep constant the utility of the poorest people in community $J - 1$. This change is a Pareto–improvement. It results in a feasible outcome as long as people of income $i_J + 1$ still prefer community J after the change. If not, then move those people to community $J - 1$, increasing g^{J-1} further so as to keep constant the utility of the poorest person in the community. Continue this until either everyone left in community J prefers it to community $J-1$, or until community J is empty. In either case, a feasible plan has been reached which Pareto dominates the original plan.

So it must be the case that community J contains all of the people in income classes $i_J, i_{J+1}, \ldots, N$. Lemma 1 shows that the remaining $J - 1$ communities form a Pareto optimum for the truncated economy, when all of the people in income classes $i_J, i_{J+1}, \ldots, N$ are removed from the economy. The induction hypothesis says that there is no splitting between communities by any income class in a Pareto optimum with $M < N$ income classes. Therefore, no community $j < J$ can contain only a (positive) proper fraction of income classes $i \quad (1 \leq i < N)$, completing the proof of the lemma. •

10.9 public output levels

Lemma 3 *In any lexmax solution, either the output level in community j is that preferred by the richest residents, so that $g^j = g^*_{I_j}(\lambda^j)$, or the output level is "excessive" and serves to keep the poorest people in the next–poorest community out : $g^j \geq g^*_{I_j}(\lambda^j)$ and $U^{I_{j-1}}(g^{j-1}, \lambda^{j-1}) = U^{I_{j-1}}(g^j, \lambda^j)$.*

Proof As with most of the lemmata, the proof proceeds by induction.

$N = 2$

Start with the lexmax solutions when $N = 2$. Lemma 2 implies that the only possible lexmax solutions involve either complete pooling, an outcome in which everyone resides in the same community, or complete separation.

In the first case, suppose that, in the one community, $g < g_2^*(\bar{y})$, where $\bar{y}$ is the average income for the whole economy. Suppose a new community is created, in addition to the existing community, with g^2 in the new community in the interval $(g, g_2^*(\bar{y})$. By construction, $U_2(g^2, \bar{y}) > U_2(g, \bar{y})$. So if it is also true that $U_1(g^2, \bar{y}) \geq U_1(g, \bar{y})$, a feasible Pareto–preferred outcome results, in which everyone moves to the new community, and is better off. If $U_1(g^2, \lambda) = U_1(g, y_1)$ for some $\lambda < \bar{y}) < y_2$, then a new feasible outcome results in which some of the poor people, and all of the rich people, move to the new community, resulting in the rich people getting a utility of $u_2 > U_2(g^2, \bar{y}) > U_2(g, \bar{y})$, so that the original outcome could not have been the lexmax solution (since another feasible outcome offers higher utility to the richest group). If $U_1(g^2, y_2) \leq U_1(g, y_1)$, then a new feasible outcome results in which all the poor people stay in community #1, and all the rich people move to community #2, attaining utility of $U_2(g^2, y_2) > U_2(g^2, \bar{y}) > U_2(g, \bar{y})$ so that again the original situation could not have been a lexmax solution.

Suppose then there is separation. If the poor strictly prefer community #1 to community #2, then small changes in g^2 will not affect the mobility constraints. So it must be the case that $g^2 = g_2^*(y_2)$, since otherwise a small change in g^2 would make better off the rich without violating the constraint.

On the other hand, if $U_1(g^1, y_1) = U_1(g^2, y_2)$ and $g^2 < g^*(y_2)$, then there exists some $\hat{g} > g_2^*(y_2)$ such that $U_2(g^2, y_2) = U_2(\hat{g}, y_2)$. The single–crossing property ensures that $U_1(\hat{g}, y_2) < U_1(g^2, y_2) = U_1(g^1, y_1)$. Therefore, increasing public output in community 2 from g^2 to something slightly less than $\hat{g}$ would make the rich strictly better off, without causing any poor to want to move in to community #2, again violating the assumption that the original plan was a lexmax solution.

Therefore, the hypothesis of the lemma must hold when $N = 2$

induction hypothesis

Assume then, that the lemma holds for all economies with $N - 1$ or fewer income classes. Consider now an economy with N income classes, and a lexmax solution for that economy with J communities populated.

Lemma 2 ensures that all the people in income class N live in community J. If we truncate the economy by getting rid of community J,and all its inhabitants, Lemma 1 implies that the resulting truncated plan for the remaining $J-1$ communities is a lexmax solution for the truncated economy with $N-n < N$ income classes, where n is the number of income classes residing in community J in the original "un–truncated" economy. Therefore, the induction hypothesis ensures that $g^j \geq g^*_{I_j}(\lambda^j)$ in communities $1, 2, \ldots, J-1$, with equality only if people of income class I_{j-1} are indifferent between communities $j-1$ and j.

What remains to be shown is that this property holds for community J as well.

community J

Suppose first that $g^J < g^*_N(\lambda_J)$. If g^J is also below the most preferred output level (given an average income of λ^J) of the richest people in community $J-1$, $g^*_{I_{J-1}}(\lambda^J)$, then there must be some output level $\tilde{g} > g^J$, such that $U_{I_{J-1}}(\tilde{g}, \lambda^J) = U_{I_{J-1}}(g^J, \lambda^J)$. Increasing output in community N from g^N to $\tilde{g}$ would therefore not induce any immigration from people of income levels of I_{J-1} or less. From the single–crossing property, the change would also increase the utility of everyone currently in community J, contradicting the assumption that the plan was a lexmax solution.

If g^J were greater than $g^*_{I_{J-1}}(\lambda^J)$, but still below $g^*_N(\lambda^J)$, then a slight increase in g^J would ; (i) make better off the richest people in community J ; (ii) not induce any immigration by people of income levels I_{J-1} or less, and (iii) not induce any emigration (if the change is small enough) by people of income $i_J, i_J+1, \ldots, N-1$. So the change would result in a feasible outcome which is better for income class N, again contradicting the assumption that the original plan was a lexmax solution.

So the public output level in community J must be at least as high as the richest people's most–preferred level. If it strictly exceeded $g^*_N(\lambda^J)$, and the selection constraint $U_{I_{J-1}}(g^{J-1}, \lambda^{J-1}) \geq U_{I_{J-1}}(g^J, \lambda^J)$ were slack, then a small decrease in g^J would make the richest people better off without inducing any migration, establishing the last part of the lemma. •

10.10 Uniqueness of the Lexmax Solution

Lemma 4 *If attention is restricted to plans in which each community's output level is different, so that $g^1 < g^2 < \cdots < g^J$, then there is at most one* *candidate equilibrium for a given population distribution.*

Proof Again the proof proceeds by induction. If $N = 2$, the only lexmax solutions have either complete separation or complete pooling. The only separating equilibrium is the one in which community 1 provides $g_1^*(y_1)$ and in which community 2 provides the minimum level of g for which $g \geq g_2^*(y_2)$ and $U_1(g_1^*(y_1), y_1) = U_1(g, y_2)$. This is unique. The only possible pooling equilibrium has $g = g_2^*(\bar{y})$. If the pooling equilibrium and separating equilibrium were both lexmax solutions, then they would both provide the same level of utility to type–2 people. Then single crossing implies that the pooling outcome would provide a strictly higher utility level to type–1 people, so that the separating outcome would not be a lexmax solution.

So assume that the lexmax solution is unique when the economy contains $N - 1$ or fewer income classes. We must show that there can be at most one lexmax solution when there are N income classes.

Suppose that there were two distinct lexmax solutions, $g^1, g^2, \ldots, g^J$ and $h_1, h_2, \ldots, h^K$. Several cases must be considered

(i) In both lexmax solutions, only type–N people live in the highest–output communities. Then the induction hypothesis says that $K = J$ and $h^j = g^j, \quad j = 1, 2, \cdots, J - 1$. If type–$N - 1$ people are indifferent between communities $J - 1$ and J in both situations, then we must have $g^J = h^J$. If the selection constraint is slack for type $N - 1$ in either situation, then $g^J = h^J = g_N^*(y_N)$ in each situation, since the definition of lexmax solution requires that U_i be the same for each income class i in each situation.

(ii) If type $N - 1$ people lived in community J in the first situation, then since we must have $U_N(h^K, \lambda^K) = U_N(g^J, \lambda^J)$, and since $U_N(h^K, y_K) > U_N(g^J, \lambda^J)$ if $h^K = g_N^*(y_N)$, it must be true that $U_{N-1} = U_{N_1}(h^K, \lambda^K)$ in the second situation : either the selection constraint binds or group $N - 1$ lives in community K in the second situation.

Since both situations give the same utility to both groups $N - 1$ and N from living in the highest–output situation, single crossing then implies that $g^J = h^K$.

(iii) Since $g^J = h^K$ in both situations, therefore $\lambda^J = \lambda^K$, so that the exact same people live in the highest–output community in both situations. Now truncate the economy by removing the residents of community J. The induction hypothesis says that then $J = K$, and that $g^j = h^j$ for each $1 \leq j \leq J - 1$, completing the proof of the lemma. •

10.11 Feasibility without a Planner

Suppose that, instead of city planners, we simply have a set of "townsites". A townsite is a community, with an associated public output level.

That is, the proprietors of the townsites commit to a level of public output g^j for their communities, but the average income in the community will be determined by people's locational choice.

As mentioned above, the number of individuals is assumed large enough that no individual migrant takes into account the effect of her own migration decision on the average income, in either the origin or destination community. That is why the feasibility condition in section 10.4 above was that L_i^j be positive if and only if community j offered income class i at least as high a level of income as any other community, taking as given the average income levels in all communities.

Without a planner, the possibility must be addressed of some community not attracting any migrants at all.

If a community is unpopulated, inferring the average income in that community poses a problem. Here we place some restrictions on people's conjectures about the average income in a community, should the community be uninhabited. Without such restrictions, some fairly "implausible" outcomes might be sustainable. Community j might offer a very attractive public output level to people of income i, but if all those people conjecture that the average income in that community is very low, no–one would be willing to live there.

As well, some inefficient outcomes are consistent with individual migration decisions, given the menu of public output levels available. For example, consider a two–type economy, in which the lexmax solution is the pooling solution, in which all people of both income classes locate in a community which provides a public output level of $g_2^*(\bar{y})$. Now suppose someone introduces a new townsite with a slightly higher level of public output provision g^2, one for which $U_2(g^2, y_2) < U_2(g_2^*(\bar{y}), \bar{y})$. If a few rich people move to the new community, the average income in the original pooled community falls. If there is some λ in $(y_1, \bar{y})$ for which $U_2(g^2, y_2) = U_2(g_2^*(\bar{y}), \lambda)$, then this emigration can be self–fulfilling : if enough rich people move to the new community, the average income in the old pooled community will fall sufficiently to make rich people in different between the communities.

Again, this outcome seems implausible. Simple communication among prospective migrants should be sufficient to eliminate it, since the outcome just described is strictly Pareto–dominated by the original pooling outcome. For this reason, we will rule out migration patterns which are Pareto–dominated by some other feasible pattern.

These modifications are incorporated in the concept of a population distribution **corresponding** to a set of townsites, defined immediately below.

DEFINITION : A distribution of population L_i^j ($i \in \{1, 2, \ldots, N\}$, $j \in \{1, 2, \ldots, J\}$) **corresponds** to a set of townsites $\{g^1, g^2, \ldots, g^J\}$ if

1. $\sum_{j=1}^{J} L_i^j = L_i \quad i = 1, 2, \cdots N$

2. $L_i^j \geq 0 \quad i = 1, 2, \cdots, N \quad j = 1, 2, \cdots, J$

3. If $L_i^j > 0$, then $U_i(g^j, \lambda^j) \geq U_i(g^k, \lambda^k) \quad \forall k \in \{1, 2, \ldots, J\} \qquad i = 1, 2, \cdots, I$

4. A set of population weights ω_i^j for each community j exists, such that $\lambda^j = \frac{\sum_{i=1}^{N} \omega_i^j y_i}{\sum_{i=1}^{N} \omega_i^j}$

5. If $\sum_{i=1}^{N} L_i^j > 0$, then $\omega_i^j = L_i^j$.

6. If $g^k > g^j$ then $\lambda^k \geq \lambda^j$.

7. No other distribution of population $\{\hat{L}_i^j\}$ satisfies all the previous properties, and Pareto dominates $\{L_i^j\}$.

Part 7 of the above definition is the refinement which rules out Pareto–dominated outcomes, as mentioned above.

Restriction 6 must be satisfied whenever townsites j and k are populated, thanks to the single–crossing property. It is motivated by the notion that people understand, in conjecturing about the population composition of unpopulated townsites, that higher–output townsites tend to attract higher–income residents.

10.12 viability

A feasible outcome will be said to be **viable** if no residents can be attracted to any additional townsites which are introduced. More formally, *Definition* Consider a feasible outcome, in which L_i^j people are assigned to community j, from a set $\mathcal{G}$ of townsites. The outcome is **viable** if this feasible outcome is a population distribution which corresponds to the set of townsites $\mathcal{H}$, for any set $\mathcal{H} \supset \mathcal{G}$.

Virtually by definition[20], a viable outcome must be Pareto optimal.

We hope that the concept of a viable outcome can be seen as closely related to competitive equilibrium : viability says that competitors cannot enter and attract anyone to their new townsites.

Viability is a fairly weak concept, in that we allow for arbitrary conjectures about the population composition of empty new townsites, subject only to the relatively weak restriction 6 of the definition of a population distribution. If people believe that new townsites have an unattractive population composition (a low level of λ), then it is fairly easy to keep people from moving to them.

Given this flexibility of conjectures, the following result holds :

Proposition 2 *The lexmax solution is viable.*

Proof Start with the lexmax solution, and now introduce new townsites. Let $g^1, g^2, \ldots, g^J$ be the output levels provided by the communities in the lexmax solution. Suppose that some new townsite has an output level h in (g^j, g^{j+1}). Then an acceptable conjecture about the average income in that townsite, if it is empty, is that the average income there is λ^j, the average income in the lexmax solution for the community in the original $\mathcal{G}$ with the next–lowest output level.

If income class i is assigned to community j in the lexmax solution, then lemma 3 shows that $U_i(h, \lambda^j) \leq U_i(g^j, \lambda^j)$ for any $h > g^j$. So the richest person in the community with output g^j will not be willing to move to the community providing $h \in (g^j, g^{j+1})$, if she conjectures that the average income there is λ^j: given an average income level of λ^j, she does not want higher public output. Single–crossing then implies that no–one of lower income, $i < I_j$, will want to move to the community with public output level h.

Suppose next that the selection constraint binds for people of income I_j, so that they are indifferent between (g^j, λ^j) and (g^{j+1}, λ^{j+1}). Then

[20] part 7 of the definition of a corresponding population distribution

$U_{I_j}(g^{j+1}, \lambda^{j+1}) = U_{I_j}(g^j, \lambda^j) \geq U_{I_j}(h, \lambda^j)$. Single crossing then implies that $U_i(g^{j+1}, \lambda^{j+1}) > U_i(h, \lambda^j)$ for all income classes $i > I_j$, so that no–one wants to move to the empty townsite promising output level h.

Finally, suppose that the selection constraint does not bind for people of income I_j, so that $U_{I_j}(g^{j+1}, \lambda^{j+1}) < U_{I_j}(g^j, \lambda^j)$. Suppose now that the poorest person in community $j+1$ did want to move to the empty community h, so that $U_{i_{j+1}}(h, \lambda^j) > U_{i_{j+1}}(g^{j+1}, \lambda^{j+1})$. For this to be the case, it must be true that more than one income class occupies community $j+1$. (If only one type lived there, and if the selection constraint did not bind from below, then in the lexmax solution $g^{j+1} = g^*_{i_{j+1}}(\lambda^{j+1})$ so that $U_{i_{j+1}}(g^{j+1}, \lambda^{j+1}) > U_{i_{j+1}}(g^*_{i_{j+1}}(\lambda^j), \lambda^j) \geq U_{i_{j+1}}(h, \lambda^j)$.)

Now consider a point $(h', y_{i_{j+1}})$ on the indifference curve of income class i_{j+1}, going through (h, λ^j). Now create a community with public output level $h' - \epsilon$, with ϵ small enough that people of income class i_j still prefer (g^j, λ^j) to $(h' - \epsilon, y_{i_{j+1}})$. Since the indifference curve of the type–i_{j+1} people through $(h', y_{i_{j+1}})$ slopes up, these people will prefer strictly $(h' - \epsilon, y_{i_{j+1}})$ to $(h', y_{i_{j+1}})$, if ϵ is small. Move a few people of income class i_{j+1} to this new community. By construction, the utility of these people goes up. But the move also increases utility in community $j+1$, since it increases the average income there. So move people of type i_{j+1} from community $j+1$ to the new community (with output $h' - \epsilon$) up until the point they are indifferent between the two communities. If ϵ is small, the number of people of i_{j+1} who will be moved must be less than $L_{i_{j+1}}$.

The move increases strictly the utility of people of type i_{j+1}. It increases strictly the utility of those remaining in community $j+1$. By construction, no–one of income less than i_{j+1} wants to move to the new community. Single crossing ensures that no–one of income greater than i_{j+1} wants to move to the new community.[21] Therefore, the assumption that $U_{i_{j+1}}(h, \lambda^j) > U_{i_{j+1}}(g^{j+1}, \lambda^{j+1})$ made it possible to create a feasible allocation which Pareto–dominated the original allocation, contradicting the assumption that the original allocation was the lexmax solution. •

Of course, there may be **other** population distributions corresponding to the augmented set $\mathcal{H}$ of townsites. Proposition 2 says only that new entry

[21] Since the poorest person in community $j+2$ must prefer that community strictly to community $j+1$ in the original lexmax solution, a sufficiently small population change in community $j+1$ will not change that preference, so that no–one in communities $j+2, j+3, \ldots, J$ will want to move.

will fail if residents are pessimistic in their assessment of the new townsites.

But, even with this flexibility, not every Pareto optimal outcome is viable. Suppose that $N = 2$. Then an allocation in which everyone is assigned to one community will be Pareto optimal, provided only that the output g in the community is between $g_1^*(\bar{y})$ and $g_2^*(\bar{y})$, and that $U_1(g, \bar{y}) > U_1(g_1^*(y_1), y_1)$, where $\bar{y}$ is the overall average income between the two groups, $\frac{L_1 y_1 + L_2 y_2}{L_1 + L_2}$. However, if $g < g_2^*(\bar{y})$, the Pareto optimum is not viable, since addition of a new townsite with output level in $(g, g_2^*(\bar{y})]$ will attract away the higher–income residents.

That means that, when $N = 2$, there are at most two viable outcomes. These are the two candidates for the lexmax solution : the pooling outcome in which everyone is assigned to a community in which $g = g_2^*(\bar{y})$ and the separating outcome in which $g^1 = g_1^*(y_1)$ and g^2 is the lowest output level in $[g_2^*(y_2), \infty)$ for which $U_1(g^2, y_2) \leq U_1(g_1^*(y_1), y_1)$. If these two outcomes cannot be Pareto–ranked, they both will be viable.

A further refinement can actually eliminate this multiplicity of equilibria, at least when there are two income classes. It seems implausible that an income group would move to a townsite, if there is no income distribution for which they would find it attractive.

Modified Definition : A population distribution corresponds to some set of townsites if it obeys the previous definition, and, in addition, the average income assigned to a townsite h^k must be some

$$\lambda^k = \frac{\sum_{i=1}^{N} \omega_i^k y_i}{\sum_{i=1}^{N} \omega_i^k}$$

where $\omega_i^k = 0$ if $U_i^j(g^j, y_1) \geq U_i(h^k, y_N)$ for some other townsite j.

Proposition 3 *If $N = 2$, the only viable outcome is the lexmax solution, if the modified definition of a corresponding population distribution is used.*

Proof If the lexmax solution is a pooling outcome, then it Pareto dominates the other possible viable outcome (the pooling outcome in which $g = g_2^*(\bar{y})$).

So it must be shown that this pooling outcome is not viable if the lexmax solution involves separation. To show this, augment the pooling outcome with 2 townsites, one with $g_1 = g_1^*(y_1)$ and one with g^2 as the smallest $g^2 \geq g_2^*(y_2)$ for which $U_1(g^2, y_2) \leq U_1(g^1, y_1)$. The new refinement in the modified definition requires that an average income level of y_2 be assigned to the townsite with output of g^2 (even if it is unpopulated). That means that the rich people will leave the original pooled community, if the lexmax solution involves separation. •

The refinement in the modified definition is somewhat plausible. Why would a poor person, for example, ever contemplate moving to a high–output community, when it is dominated by some other community, for every possible average income level in the two communities? And then why should anyone ever expect any of these poor people ever to move to this community?

When $N > 2$, the set of viable outcomes is not necessarily limited to the lexmax solutions, even with the refinement added to the definition of a corresponding population distribution.

Imagine, for example, that $N = 3$, and that the lexmax solution pooled income classes 2 and 3, and left income class 1 alone in a separate, lower–output community. Another Pareto optimum might involve pooling income groups 1 and 2, and leaving group 3 alone with a very high level of output. If this second optimum was better for income class 1, and worse for the other two groups, then it might be viable.

In this example, let g^1 and g^2 be the public output levels associated with the lexmax solution, and h^1 and h^2 the output levels associated with the other Pareto optimum mentioned, with $g^1 < h^1 < g^2 < h^2$. The associated average income levels are $\lambda^1 = y^1$ and $\lambda^2 = \frac{L_2 y_2 + L_3 y_3}{L_2 + L_3}$ corresponding to g^1 and g^2, and $\mu^1 = \frac{y_1 L_1 + L_2 y_2}{L_1 + L_2}$ and $\mu^2 = y_3$ associated with h^1 and h^2. The second Pareto optimum would be viable here, even with the modified definition, if we had

$$U_1(h^1, \mu^1) > U_1(g^1, \lambda^1) = U_1(g^2, \lambda^2)$$

$$U_2(g^2, \lambda^2) > U_2(h^1, \mu^1) = U_2(h^2, \mu^2) > U_2(g^2, y_2)$$

$$U_3(g^2, \lambda^2) > U_3(h^2, \mu^2)$$

So here, if we started with the "other" Pareto optimum, a rival planner who introduced the two townsites corresponding to the lexmax solution would not be guaranteed to attract residents away from the existing townsites. The townsite with output level g^2 is attractive only if its average income is relatively high. If the type–2 residents of the townsite with (h^1, μ^1) were guaranteed that all the highest–income people would join them in a move to the g^2 townsite, they would be willing to move. They can be assured (through the refinement in the definition) that none of the lowest–income people would go to this town. But even with that assurance, if they (and the high–income people) believed the average income in the g^2 town was close to y_2, no–one would be willing to move there.

10.13 entrepreneurs and managers

Since the specification of people's preferences includes the budget constraint of their community, some generalization is needed to consider profit–maximizing entrepreneurs. Entrepreneurs can charge an entry fee f to residents of townsites they set up, but that this fee must be the same for all residents.

Entrepreneurs set up townsites, and commit to provide specific output levels in each of their townsites.

The outcome when several entrepreneurs set up different townsites is the population distribution corresponding to that townsite. The profit earned by an entrepreneur a who operates a set $\mathcal{T}_a$ of townsites is

$$\pi_a \equiv \sum_{j \in \mathcal{T}_a} f^j \sum_{i=1}^{N} L_i^j$$

if the entrepreneur charges the fee f^j to each resident of townsite j. We require that all fees be non–negative.

Definition : A set of entrepreneurs, and their associated menus $\{f^j, g^j\}$ of output levels and fees constitutes a competitive equilibrium if no new entrepreneur can enter, and earn strictly positive profits from a population distribution corresponding to the new menu resulting from the new entrant's menu being added to the existing list.

Naturally, competition will drive entrepreneurs' fees to zero. For that reason, we can restrict attention to **city managers** who are just entrepreneurs who are constrained not to make profits. City managers simply create townsites, and commit to output levels there, trying to attract residents.

So if some outcome is viable, then no city manager can enter, and attract any residents to a new townsite. Conversely, if no city manager could enter and attract residents to a new townsite, the original outcome is viable.

If city managers can't attract any residents to new townsites, then profit–maximizing entrepreneurs can't make any profits from new townsites. That is,

Proposition 4 *The lexmax solution is always a competitive equilibrium.*

Proof Since the lexmax solution is viable, no city manager could open a set of new townsites and attract any new residents. That means no entrepreneur can make profits by entering with new townsites. •

This result also ensures that a competitive equilibrium must exist. When there are only 2 income classes, the modified definition of a population distribution implies that the competitive equilibrium is unique.

Theorem 1 *When $N = 2$, the competitive equilibrium is unique (up to cloning of identical jurisdictions).*

Let $g_2 = \min(g_1^E, g_2^*(y_2))$. If $U_2(g_2, y_2) > U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$ then the competitive equilibrium is a separating equilibrium, in which all poor people live in jurisdictions providing a level $g_1^*(y_1)$ of the public output, and all rich people live in jurisdictions providing g_2. If $U_2(g_2, y_2) \leq U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$, then the competitive equilibrium is a pooling equilibrium, in which all people, rich or poor, live in a jurisdiction providing $g_2^*(\bar{\lambda})$ of the public output.

10.14 Stability under Voting ($N = 2$)

Proposition 5 *If there are at least two communities of each type, then the equilibrium is stable with respect to voting.*

Proof There are three cases to consider: (*i*) a separating equilibrium; (*ii*) a pooling equilibrium with the rich in the majority; (*iii*) a pooling equilibrium with the poor in the majority.

We will consider each case in turn.

1. Separating Equilibrium

Consider first voting in "all rich". Residents of a rich community will favor a new public output level g_2' only if it will yield higher utility to rich people than the utility $u_2(g_1^E, y_2)$ they attain currently. But higher utility for the rich means that the output and population in the community after the change, (g_2', λ') must be on a higher indifference curve for the rich than the original (g_1^E, y_2). This pattern is only possible if $g_1^E > g_2^*(y_2)$, and therefore can only happen if $g_2' < g_1^E$. But then single crossing implies that $U_1(g_2', \lambda') > U_1(g_1^E, y_2)$. Since $U_1(g_1^E, y_2) = U_1(g_1^*(y_1), y_1)$, this last inequality implies that all the poor residents will also move to the rich community which changed its public output unilaterally to g_2'. Because the community attracts all the poor people, it must have an average income λ' which is less than or equal to the overall average $\bar{\lambda}$. But this implies a contradiction: if the equilibrium is separating, then $U_2(g_1^E, y_2)$ is no less than the utility the rich can attain from any outcome on or below the pooling line.

Now consider a change in public output in one of the poor communities, to some $g_1' \neq g_1^*(y_1)$, which increases the utility of the poor (original) residents of the community. If the change benefits the residents of this community, then it benefits all poor people. Therefore, it will attract all the poor people. Therefore, the average income λ' in this community, after migration has occurred, will be less than or equal to the overall average $\bar{\lambda}$. But that means that none of the rich people will choose to live in the new community, since it offers them no higher utility than they can get in "all rich". (The fact that the equilibrium is separating implies that "all rich" is at least as attractive to the rich as any outcome on or below the pooling line.) So the proposed change in public output in the poor community can attract only poor people, which means that the new population composition $\lambda' = y_1$. But by definition, $g_1^*(y_1)$ maximizes $U_1(\lambda, y_1)$, so that no change in public output can benefit the poor residents.

2. Pooling Equilibrium : Rich Majority

In this case the rich determine policy in each of the identical communities.

In this equilibrium, the public output in each mixed community is $g_2^*(\bar{\lambda})$. Now consider the indifference curve for the rich, which is tangent to the horizontal line $\lambda = \bar{\lambda}$. This curve is marked $IR2$ in Figure 2. Let $\tilde{g} > g_2^*(\bar{\lambda})$ be the larger of the two public output levels for which this indifference curve hits the "all–rich" composition line $\lambda = y_2$. (That is $\tilde{g}$ is the unique public output level for which $\tilde{g} > g_2^*(\bar{\lambda})$ and for which $U_2(\tilde{g}, y_2) = U^2(g_2^*(\bar{\lambda}), \bar{\lambda})$.)

The fact that $\tilde{g} > g_2^*(\bar{\lambda})$, and the single–crossing property, imply that the poor must prefer strictly $(g_2^*(\bar{\lambda}), \bar{\lambda})$ to $(\tilde{g}, y_2)$. Therefore, any new outcome (g', λ') which is preferred by the rich to the original outcome $(g_2^*(\bar{\lambda}), \bar{\lambda})$ must also be preferred strictly by the poor.

So any policy change in any of the identical mixed communities which attracts any of the rich must attract all the poor residents. [If it attracts any of the rich, it must have an average income λ' which is greater than $\bar{\lambda}$; if it raises λ in this community, it must lower λ in each of the others; and since the poor prefer (g', λ') to $(g_2^*(\bar{\lambda}), \bar{\lambda})$, they must prefer (g', λ') to any $(g_2^*(\bar{\lambda}), \lambda)$ with $\lambda < \bar{\lambda}$.]

Therefore, when population adjusts, no public output change in any of the mixed communities can benefit the rich majority.

3. Pooling Equilibrium : Poor Majority

In this case, Lemma 5 proved below will be useful.

Suppose that the poor majority in some mixed community can get higher utility $U_1(g', \lambda')$ from changing their community's public output level from $g_2^*(\bar{\lambda})$ to g', after the population response.

If the change makes the poor majority of this community strictly better off, then it must make all the poor people better off. Therefore it must attract all the poor people, so that $\lambda' \leq \bar{\lambda}$. In that case each of the other mixed communities, which have not changed their public output levels, will have an average income of $\bar{\lambda}$ or more. So the utility of the rich in one of these other communities, $U_2(g_2^*(\bar{\lambda}), \lambda)$ with $\lambda \geq \bar{\lambda}$, will be greater than the utility they would get in the community which changed its spending, $U_2(g', \lambda') \leq U_2(g', \bar{\lambda}) < U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$.

Therefore, the proposed policy change would attract only poor people, but no rich people, resulting in an outcome (g', y_1). Since, by definition, $U_1(g_1^*(y_1), y_1) \geq U_1(g', y_1)$, Lemma 5 shows that the poor will not benefit from this policy change.

The last two sections of the proof actually show that neither the poor, nor the rich, can benefit from any unilateral public output change in a pooling equilibrium. Therefore, the razor's edge case in which $L^1 = L^2$ is also accounted for, neither group wants to change local public output unilaterally. •

Lemma 5 *In any pooling equilibrium, the poor must get higher utility* $U_1(g_2^*(\bar{\lambda}), \bar{\lambda})$ *than they would get in any all–poor community,* $U_1(g_1^*(y_1), y_1)$.

Proof Take the U–shaped indifference curve of the rich, through $(g_2^*(\bar{\lambda}), \bar{\lambda})$, labeled $IR2$ in Figure 2. This curve is tangent to the line $\lambda = \bar{\lambda}$. If there is a pooling equilibrium, spending g_1^E must lie to the right of the rightmost intersection of this curve with the horizontal line $\lambda = y_2$: the fact that the equilibrium is pooling means that (g_1^E, y_2) is on a lower indifference curve for the rich than $IR2$; the fact that $g_1^E > g_2^*(y_2)$ means that g_1^E must lie the right of the rightmost intersection of $IR2$ with the line $\lambda = y_2$.

So $U^2(g_2^*(\bar{\lambda}), \bar{\lambda}) > U_2(g_1^E, y_2)$ and $g_1^E > g_2^*(\bar{\lambda})$. Therefore, single crossing implies that $U_1(g_2^*(\bar{\lambda}), \bar{\lambda}) > U_1(g_1^E, y_2)$. By definition, $U_1(g_1^E, y_2) = U_1(g_P^*(y_1), y_1)$, so that $U^1(g_2^*(\bar{\lambda}), \bar{\lambda}) > U_1(g_P^*(y_1), y_1)$. •

10.15 Uniformity versus Decentralization

Proposition 6 *If the poor constitute the national majority, then the poor are strictly better off under uniform centralization, and the rich are strictly better off under complete decentralization.*

Proof Under uniform centralization, with policy determined by the poor, the outcome is $(g_1^*(\bar{\lambda}), \bar{\lambda})$. The assumption that each group's preferred public output level is unique, and the assumption that the rich prefer more public output than the poor, imply that $U_1(g_1^*(\bar{\lambda}), \bar{\lambda}) > U_1(g_2^*(\bar{\lambda}, \bar{\lambda})$. The assumption that utility increases with λ implies that $U_1(g_1^*(\bar{\lambda}), \bar{\lambda}) > U_1(g_1^*(y_1), y_1)$, so that the poor do strictly better under uniform centralization than they do in a separating equilibrium or in a pooling equilibrium.

Under complete decentralization, the utility of the rich is either $U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$ (in a pooling equilibrium), or else greater than or equal to $U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$ (in a separating equilibrium). Since $U_2(g_2^*(\bar{\lambda}), \bar{\lambda}) > U_2(g_1^*(\bar{\lambda}), \bar{\lambda})$, the rich must be strictly better off under complete decentralization. •

If the rich are in a majority nationally, then they weakly prefer decentralization. With the rich in a majority, uniform centralization and the pooling equilibrium under complete decentralization give the same outcome. If there is a separating equilibrium, then the rich must prefer their outcome (g_1^E, y_2) strictly to the outcome $(g_2^*(\bar{\lambda}), \bar{\lambda})$ of uniform centralization.

However, the conflict between rich and poor over the form of organization may be less clear if the rich are in a majority. If the outcome under complete separation is a pooling equilibrium, and if the rich are in a majority, then uniform centralization leads to the same outcomes as complete decentralization. But if the outcome under complete decentralization is a separating equilibrium, then the poor may be better or worse off than under uniform centralization.

Proposition 7 *If the rich are in the majority, they will prefer weakly complete decentralization to uniform centralization. The poor may be better, or worse off, in a separating equilibrium than under complete centralization with the rich in a majority.*

Proof What remains to be shown is the existence of separating equilibria in which $U_1(g_1^*(y_1), y_1) < U_1(g_2^*(\bar{\lambda}), \bar{\lambda})$ and separating equilibria in which $U_1(g_1^*(y_1), y_1) > U_1(g_2^*(\bar{\lambda}), \bar{\lambda})$. Figure 5 illustrates the first case, Figure 6 the second.

Underlying the figures are the following preferences. In each case, $U_2(g, \lambda) \equiv (48 - g)g + 500\lambda$. In each case, $y_1 = 0$ and $y_2 = 1$. In Figure 5, $U_1(g, \lambda) \equiv (36 - g)g + 144\lambda$, whereas in Figure 6 $U_1(g, \lambda) \equiv (36 - g)g + 40\lambda$. Therefore $g_2^*(\lambda) = 24$ and $g_P^*(\lambda) = 18$ for any λ in either figure, $g_1^E = 30$ in Figure 5, $g_1^E = 24.3246$ in Figure 6, $\bar{\lambda} = 0.54$ in each case, and in each case $U_2(g_1^E, y_2) > U_2(g_2^*(\bar{\lambda}), \bar{\lambda})$ so that the equilibrium under complete decentralization is separating. •

References

[1] Arrow, Kenneth (1972) "Some mathematical models of race discrimination in the labor market." In A.H. Pascal, ed. *Racial Discrimination in Economic Life*, Lexington MA.: Lexington Books, pp. 187-203.

[2] Azam, Jean Paul and Ruxanda Berlinschi (2008) "The aid-migration trade-off." Presented at the Annual Bank Conference on Development Economics, Cape Town, South Africa

[3] Becker, Gary S. (1957) *The Economics of Discrimination*. Chicago: University of Chicago Press.

[4] Becker, Gary S. and Kevin M. Murphy (2000) *Social Economics: Market Behavior in a Social Environment*. Cambridge: Harvard University Press.

[5] Berglas, Eitan (1976) "On the theory of clubs." *American Economic Review*, 66(2): 116–121.

[6] Berglas, Eitan, and David Pines (1981) "Club, local public goods, and transportation models: A synthesis." *Journal of Public Economics*, 15(2): 141–162.

[7] Bermeo, Sarah Blodgett and David Leglang (2010) "Foreign interests: Immigration and the political economy of foreign aid." Working paper, Duke University.

[8] Blogchliger, Hansjorg et al. (2007) "Fiscal equalisation in OECD countries." OECD Network on Fiscal Relations Across Levels of Government, Working Paper No. 4.

[9] Bolton, Patrick and Gerard Roland (1997) "The breakup of nations: A political economy analysis." *Quarterly Journal of Economics*, 112(4): 1057–1090.

[10] Borjas, George J. "The politics of employment administration in the federal bureaucracy." *Journal of Law and Economics*, 25(2): 271–300.

[11] Bossert, W., Pattanaik, P., and Y. Xu (1994) "Ranking Opportunity Sets : An Axiomatic Approach." *Journal of Economic Theory*, 63(2): 326–345

[12] Brueckner, Jan K. and Kangoh Lee (1989) "Club theory with a peer-group effect." *Regional Science and Urban Economics*, 19(3): 399–420.

[13] Brueckner, Jan and Amihai Glazer (2008) "Urban extremism." *Journal of Law, Economics, and Organization*, 24(2): 307–318.

[14] Conley, John and Myrna Wooders (1998) "Anonymous Lindahl pricing in a Tiebout economy with crowding types." *Canadian Journal of Economics*, 31(4): 952–974.

[15] Cremer,Jacques and Thomas R. Palfrey(2000) "Federal Mandates by Popular Demand." *Journal of Political Economy*, 108 (5): 905–927.

[16] de Bartolome, Charles A.M. (1990) "Equilibrium and inefficiency in a community model with peer group effects." *Journal of Political Economy*, 98(1): 110–133.

[17] Dionne, Georges, Neil Doherty, and Nathali Fombaron (2000) "Adverse selection in insurance markets." In G. Dionne ed. *Handbook of Insurance*, Dordrecht: Kluwer Academic Publishers.

[18] Ellickson, Bryan (1971) "Jurisdictional fragmentation and locational choice." *American Economic Review*, 61(2): 334–339.

[19] Epple, Dennis, Radu Filimon and Thomas Romer (1984) "Equilibrium among local jurisdictions: Toward an integrated treatment of voting and residential choice." *Journal of Public Economics*, 24(3): 281–308.

[20] Epple, Dennis, Radu Filimon and Thomas Romer (1993) "Existence of voting and housing equilibrium in a system of communities with property taxes." *Regional Science and Urban Economics*, 23(5): 585–610.

[21] Epple, Dennis N. and Richard Romano(2003) "Neighborhood schools, choice, and the distribution of educational benefits." In Caroline M. Hoxby, editor, *The Economics of School Choice*, Chicago: University of Chicago Press.

[22] Epple, Dennis and Thomas Romer (1991) "Mobility and redistribution." *Journal of Political Economy*, 99(4): 828–858.

[23] Fennell, Lee Anne (2009) *The Unbounded Home: Property Values Beyond Property Lines*. New Haven: Yale University Press.

[24] Fernandez, Raquel, and Richard Rogerson (1997) "Keeping people out: Income distribution, zoning, and the quality of public education." *International Economic Review*, 38(1): 23–42.

[25] Glaeser, Edward L. and Joseph Gyourko (2003) "The impact of building restrictions on housing affordability." *Federal Reserve Bank of New York Economic Policy Review*, 21–39.

[26] Glaeser, Edward L. and Andrei Shleifer (2005) "The Curley effect: The economics of shaping the electorate," *Journal of Law, Economics, and Organization*, 21(1): 1–19.

[27] Hamilton, Bruce W. (1975) "Zoning and property taxation in a system of local governments." *Urban Studies*, 12: 205 – 211

[28] Hamilton, Bruce W. (1976) "Capitalization of intrajurisdictional differences in local tax prices." *American Economic Review*, 66(5): 743–753.

[29] Hoxby, Caroline M. (2001) "All school finance equalizations are not created equal." *Quarterly Journal of Economics*, 116(4): 1189–1231.

[30] Hoyt, William H. and Kangoh Lee (2003) "Subsidies as sorting devices." *Journal of Urban Economics*, 53(3): 436–457.

[31] Lee, Kangoh (1993) "Bureaucrats and tax limitation." *Journal of Urban Economics*, 34(1): 24–42.

[32] McGuire Martin. C. (1991) "Group composition, collective consumption, and collaborative production." *American Economic Review*, 81(5): 1391–1407.

[33] Mullins, Daniel R. And Kimberley A. Cox (1995) "Tax and expenditure limits on local governments." Center for Urban Policy and the Environment, Indiana University.

[34] Oates, Wallace (1972) *Fiscal Federalism*. New York: Harcourt-Brace.

[35] Richardson, Jesse J. Jr., Meghan Zimmerman Gough, and Robert Puentes (2003) "Is home rule the answer? Clarifying the influence of Dillon's Rule on growth management." Brookings Institution Center on Urban and Metropolitan Policy.

[36] Rose–Ackerman, Susan (1979) "Market models of local government: Exit, voting and the land market." *Journal of Urban Economics*, 6(3): 319–337.

[37] Rothschild, Michael, and Joseph Stiglitz (1977) "Equilibrium in competitive insurance markets." *Quarterly Journal of Economics*, 90(4): 629–649.

[38] Scotchmer, Suzanne (1997) "On price–taking equilibria in club economies with nonanonymous crowding." *Journal of Public Economics*, 65(1): 75–87.

[39] Scotchmer, Suzanne and Myrna Wooders (1987) "Competitive equilibrium and the core in economies with anonymous crowding." *Journal of Public Economics*, 34(2): 159–173

[40] Strahilevitz, Lior Jacob (2006) "Exclusionary amenities in residential communities." *Virginia Law Review*, 92: 437–499.

[41] Tiebout, Charles (1956) "A pure theory of local expenditures." *Journal of Political Economy*, 64(5): 416–424.

[42] Westhoff, Frank (1977) "Existence of equilibrium in economies with a local public good." *Journal of Economic Theory*, 14(1): 84–112.

[43] Wheaton, William (1993) "Land capitalization, Tiebout mobility, and the role of zoning regulations." *Journal of Urban Economics*, 34(2): 102–117.

[44] Wilson, Charles (1977) "A model of insurance with incomplete information." *Journal of Economic Theory*, 16(2): 167–207.

[45] Wilson, John D. (1998) "Imperfect solutions to the musical suburbs problem." In D. Pines, E. Sadka, and I. Zilcha, eds., *Topics in Public Economics: Theoretical and Applied Analysis.* Cambridge, UK: Cambridge University Press.

11 Figures

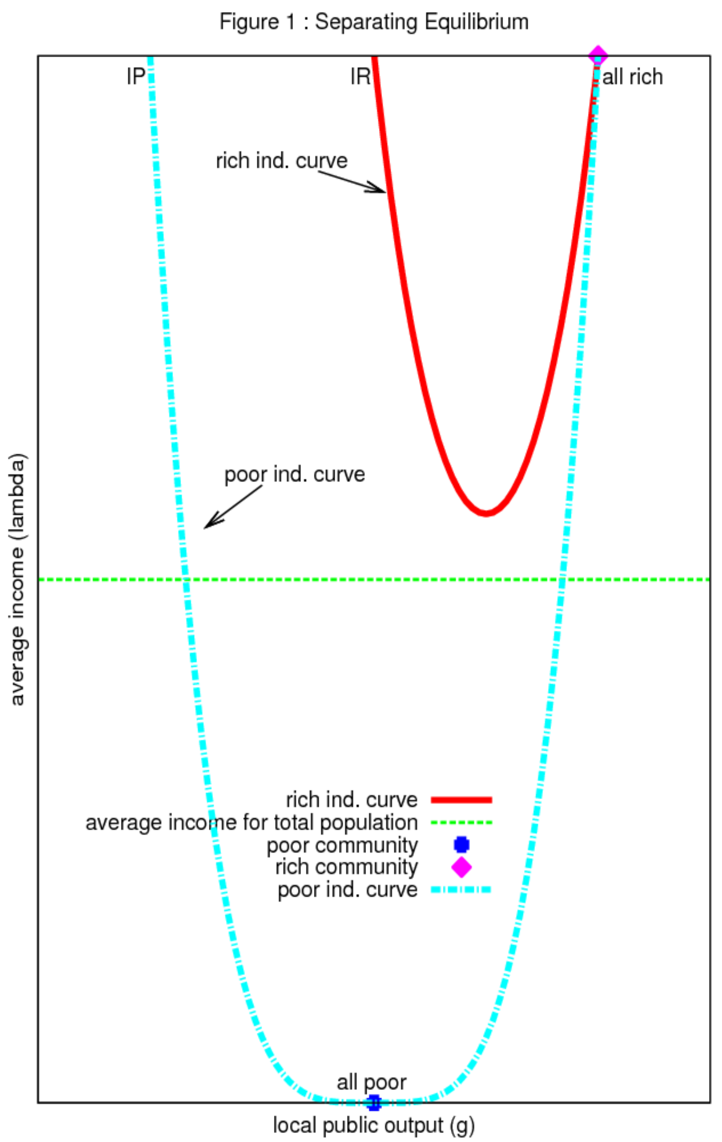


Figure 1 : Separating Equilibrium
IP
IR
all rich
rich ind. curve
poor ind. curve
average income (lambda)
rich ind. curve
average income for total population
poor community
rich community
poor ind. curve
all poor
local public output (g)

Figure 2 : Pooling Equilibrium

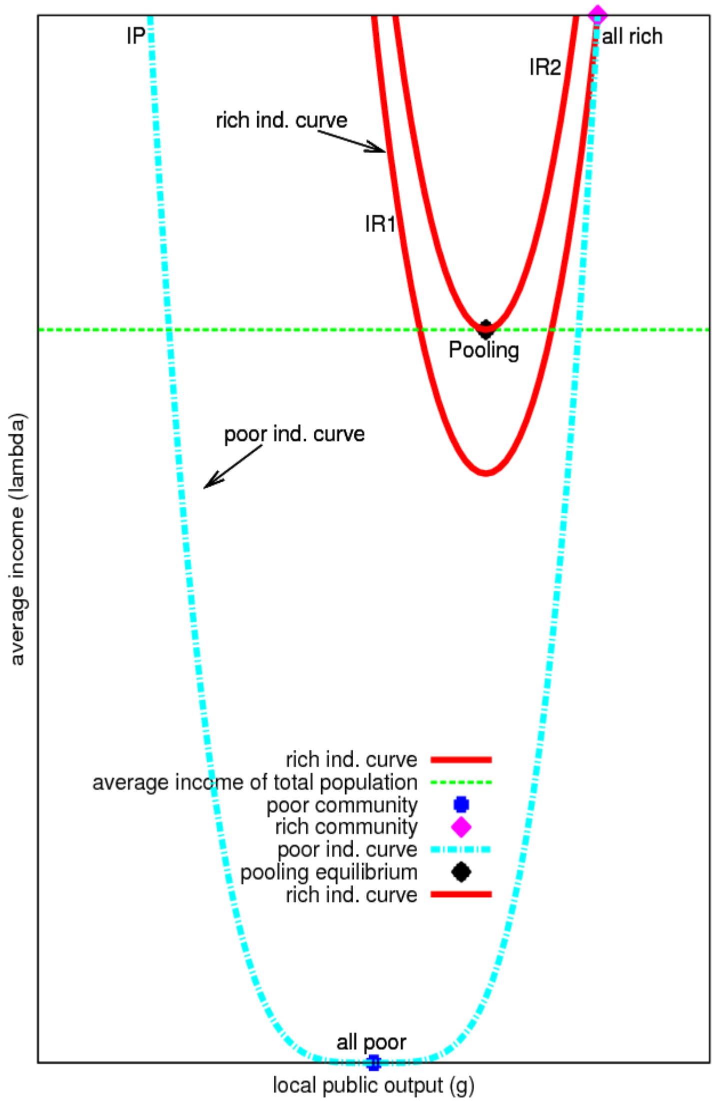
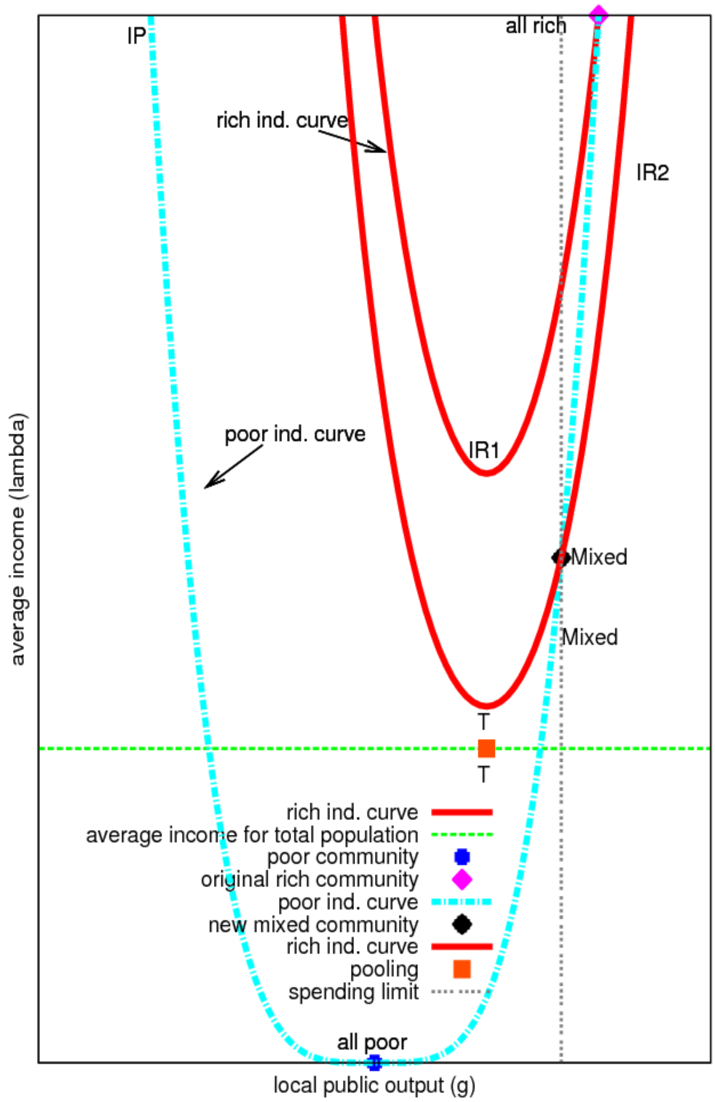


Figure 3 : Spending Limits Help No-one



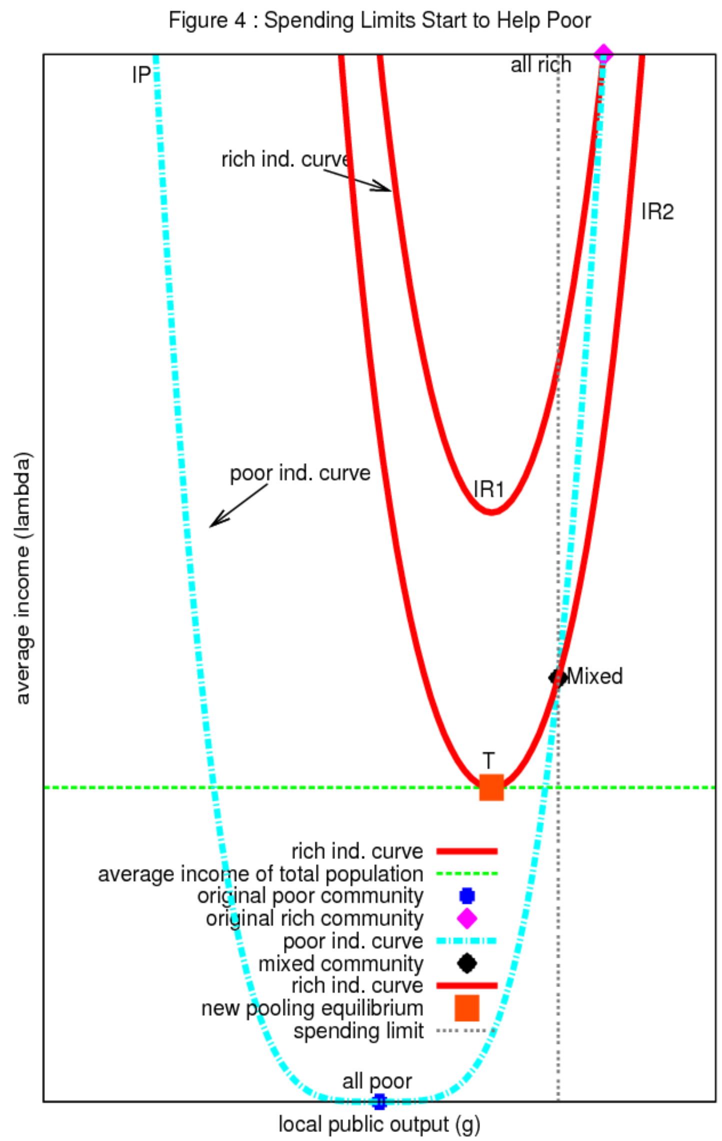


Figure 4 : Spending Limits Start to Help Poor
IP
all rich
rich ind. curve
IR2
poor ind. curve
IR1
Mixed
T
rich ind. curve
average income of total population
original poor community
original rich community
poor ind. curve
mixed community
rich ind. curve
new pooling equilibrium
spending limit
all poor
average income (lambda)
local public output (g)

Figure 5 : Poor Prefer Uniform Provision

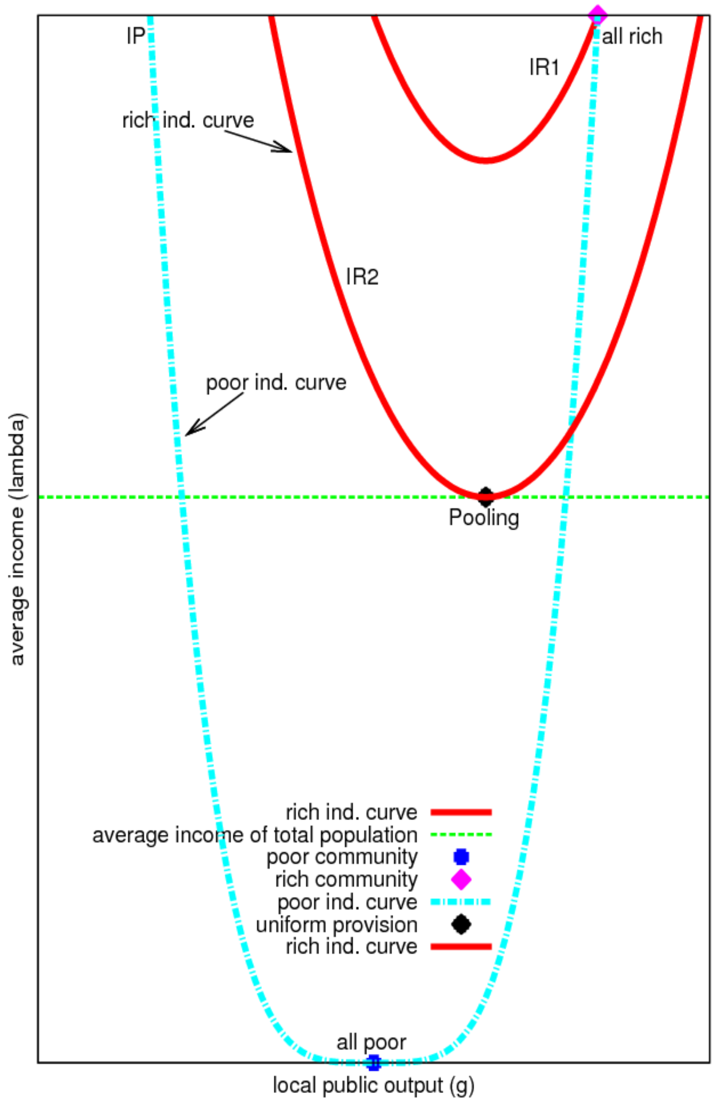


Figure 6 : Poor Prefer Decentralization