Competition among Fallible Governments

by

James R. Hines Jr.
University of Michigan and NBER

October 2007

I thank Timothy Besley, Andrew Haughwout, Erzo Luttmer, and numerous seminar participants for helpful comments on an earlier draft of this paper.

Competition among Fallible Governments

ABSTRACT

Should government policies be harmonized or instead permitted to vary among competing jurisdictions? This paper considers the welfare effects of policy heterogeneity with freely mobile populations. Jurisdictions adopting better policies experience population inflows. If the costs of population concentration are small, then competition, in affording greater choice, promotes higher welfare levels. With significant convex congestion costs, however, this welfare ordering is reversed: competition induces so great a concentration of population in jurisdictions adopting preferred policies that average consumer welfare is lower than with harmonization. To the degree that interjurisdictional mobility is limited by income redistribution policies, intergovernmental transfers, congestion-related taxes or rising land prices, competition among decentralized governments may again produce higher welfare than harmonization. Hence the welfare impact of policy harmonization depends critically on the nature of crowding costs and the extent to which market forces or government policies reduce population mobility.

JEL Classification: H73, R13, D71.

Keywords: fiscal federalism, tax competition, policy coordination.

James R. Hines Jr.
Department of Economics
University of Michigan
343 Lorch Hall
611 Tappan Street
Ann Arbor, MI 48109-1220

jrhines@umich.edu

1. Introduction.

Governments everywhere compete over the extent to which they offer worthwhile educational opportunities, effective transportation links, safe streets, moderate taxes, and other attractive policies. Under what circumstances would a federation benefit from limiting this kind of competition among its constituent jurisdictions? The case in favor of limiting competition typically starts by identifying circumstances in which competing jurisdictions have incentives to enact policies designed to advance their own interests at the expense of others. For example, a small town might adopt lax environmental regulations if much of the cost of local pollution is borne by nonresidents who live downwind or downriver, and there are no mechanisms in place that capture the benefits of improving environmental quality elsewhere. In settings with significant externalities of this type, an optimizing centralized authority can improve on the decentralized outcome by setting policies that incorporate spillover effects.

Competition among decentralized governments offers what are potentially offsetting benefits, notably including the greater range of choice that competition typically affords individuals and firms selecting among locations. Competing jurisdictions have incentives to tailor their policies to local needs and the demands of the marketplace. In competing with each other governments become more accountable for their actions than they would be otherwise, thereby quite possibly stimulating the formation and adoption of better public policies than would be forthcoming from a centralized authority. With sufficient information, policy instruments, and presence of mind a centralized government authority could replicate the outcomes produced by competing jurisdictions, and conceivably improve upon them by incorporating the costs and benefits of spillovers between jurisdictions. In practice, however, policies chosen by central governments exhibit uniformity across locations that obscures the nuances of local needs and betrays the kind of diversity that might be expected to accompany local innovation. Hence the choice between local discretion and central government fiat typically entails a choice between policies that vary across locations and those that are effectively harmonized by the center.

The purpose of this paper is to compare the welfare properties of policies chosen by competing governments with the properties of policies that are harmonized by central authorities. While there are many aspects of such a comparison, the analysis focuses on the impact of population mobility, doing so because the ability of individuals and firms to discipline governments with exit carries profound implications for government competition. Jurisdictions adopting tax, education, criminal justice, environmental, social welfare, and other policies that produce high satisfaction levels for their residents become attractive places, drawing population and mobile economic resources that might otherwise have located elsewhere. While it is customary to analyze the relative merits of competition and harmonization in settings in which governments respond shrewdly to the incentives created by their strategic situations, a focus on mobility makes it useful to consider settings in which governments choose policies nonstrategically – indeed, randomly.

Random policy choice is a formalization of the commonplace observation that governments are fallible. In a competitive setting with numerous jurisdictions, it can be expected that some governments will adopt unfortunate policies, not because they want to do so, but because they are incapable of doing otherwise. Governments often lack the information necessary to enact sound policies, and they frequently operate under severe financial and bureaucratic constraints that impede their abilities to adopt even those measures that are generally agreed to be worthwhile. The political nature of government decisionmaking notoriously diverts attention from the public welfare to other considerations in policy development and reform, and even granting the natural tendency to overstate these critiques of government, it is clear that the search for an ideal public sector is illusory.

In a competitive setting, consumers can avoid the consequences of undesired government policies by locating elsewhere. If central governments are no better or worse than local governments in selecting desirable policies, and in the absence of spillovers, so that a jurisdiction's policies affect only its own residents, then the welfare comparison of competition and harmonization turns on the impact of the induced population distribution. The analysis of random policy choices implies that if crowding costs are unimportant, then competition supports higher welfare levels than does harmonization, since competition affords residents the opportunity to select jurisdictions adopting the best policies. As crowding costs become more pronounced, the desirability of harmonization increases; if crowding costs are sufficient to ensure that all jurisdictions are occupied, then policy harmonization produces greater average welfare than does competition. To the degree that crowding costs are captured in land prices, tax charges, or other forms of interregional redistribution that are correlated with population density, the relative benefits of harmonization are mitigated, so with sufficient congestion-related pricing, competition again becomes the preferred alternative.

2. Fiscal federalism.

Federalist governmental structures, such as those in the United States, Canada, and Germany, institutionalize jurisdictional competition, consciously limiting the harmonizing influence of the central government in an effort to promote local decisionmaking, diversity, and competition among subnational governments. The desirability of such competition has been debated for ages.[1] It is useful to identify the costs and benefits of decentralization as identified in much of this literature.

Competition provides the opportunity for jurisdictions to experiment with policies that are new and possibly superior to what a central authority might adopt; competition also affords citizens a greater range of choice than they would face under uniform policies prescribed centrally. Against this must be weighed the costs of any resource misallocation that accompanies the competitive process, including the costs of population reallocation and the costs of any substandard policies adopted by competing governments.

Tiebout's (1956) classic treatment of local public good provision with population mobility considers the welfare properties of competition to provide goods and services to consumers with heterogenous tastes. The Tiebout model is one in which consumer/voters express their preferences with exit rather than voice; a sizable modern literature considers the impact of voice that takes the form of voting. Governments are perfect optimizers in the Tiebout model. The setting evaluated in this paper is one of pure exit, but in the absence of heterogeneous tastes. Hence it considers the role of exit in mitigating the effect of poor public policy choices on consumer welfare, in a setting in which there is universal agreement about what would constitute good and bad public policies.

3. Fallibility and exit.

This section considers the implications of fallibility in settings in which government behavior is disciplined by population mobility. Government policies can be summarized by a

[1] Oates (1999) and Wellisch (2000) offer surveys of this literature.

vector $\boldsymbol{\theta} \equiv (\theta_1,...,\theta_n)$, the elements of which consist of the complete set of policy parameters. Thus, θ_1 might correspond to the top personal marginal income tax rate, θ_2 to average per-pupil elementary school expenditures, θ_3 to the maximum permitted size of commercial office buildings, and so on. For most analytic purposes it is sufficient, as well as convenient, to treat the $\boldsymbol{\theta}$ vector as though it has a scalar-valued single element, denoted θ. The consequences of relaxing this assumption, and treating government expenditures instead as a vector, are considered in section 5.

In order to isolate the impact of government policies, jurisdictions are taken to be identical in all respects (as suburban towns so often are), except for their chosen policies; and individual consumers are likewise taken to be identical. Individual utility is a function of individual incomes, *y*, the quality of government policies, θ, and the size of the local population, denoted *n*, and can therefore be represented by a utility function $u(y,\theta,n)$. In this formulation, policies do not have direct spillover effects: an individual's utility is a function of the policies adopted by the jurisdiction in which he resides, and not the values of θ adopted by other jurisdictions. The variable θ is defined so that higher values are always better: $u(y,\theta',n) > u(y,\theta'',n) \forall \theta' > \theta''$. The policies of other jurisdictions have at most an indirect effect on utility, by influencing the allocation of the population. Jurisdictions are subject to crowding, with associated average costs that may rise in a convex fashion with each additional resident. Thus, $\frac{\partial u(y,\theta,n)}{\partial n} \leq 0$, and if congestion costs are convex, then $\frac{\partial^2 u(y,\theta,n)}{\partial n^2} \leq 0$. The national population fixed, but free population mobility within a country implies that crowding will be most severe in jurisdictions choosing highly desirable public policies.

3.1 An example: two jurisdictions, binary policy choices.

The implications of the model become apparent in a simple example with just two jurisdictions and an aggregate population of six individuals who can move freely between the two locations. Governments face simple a simple discrete choice between two alternatives corresponding to different values of θ (for example, whether or not to build an expensive new sports arena) with unknown payoffs. In the absence of coordination imposed by the central government, each jurisdiction chooses its own value of θ, and the population moves in response. If crowding costs are trivial (very close to zero) and policies differ, then the entire population chooses to locate in the jurisdiction selecting the more valuable policy option; whereas if the jurisdictions choose identical policies, then the population divides itself evenly between them. If, instead, the central government were to coordinate policies among jurisdictions, requiring that both select the same θ, then even very small crowding costs imply that the population will divide itself evenly between them.

What are the welfare implications of policy harmonization imposed by the central government? The answer turns on a comparison of the distribution of the central government's choice of θ and the distribution of choices made by the subnational governments. If both select from the same distribution, that is, if a subnational government is just as likely to select the more desirable policy option as is the central government, then it follows that the expected welfare of the citizenry is greater with uncoordinated policies. The reason is simply that uncoordinated policies offer more options, and in this case, greater range of choice improves the likelihood that the set of choices available to individuals includes the best possible public policy. If governments choose the superior policy half the time, and choices are independent, then under a harmonized system citizens will enjoy the superior policy half the time, whereas under independent decisionmaking by two subnational jurisdictions, citizens will enjoy the superior policy three-quarters of the time.

The greater expected welfare produced by decentralized decisionmaking follows from the assumption that crowding costs are tiny, and carries the counterfactual implication that small differences in the quality of public policies will produce enormous swings in the distribution of the population. This is just a formalization of the notion that federalist structures enjoy the benefits of experimentation by competing subnational governments. In this case, experimentation takes the form of repeated draws from a single distribution of potential policies.

In order to evaluate the impact of crowding costs, it is useful to consider the case in which utility is additively separable in income, the value of the public policy, and a quadratic cost of population crowding, and therefore given by:

(1) $$u(y,\theta,n) = y + \theta - n^2 .$$

Suppose that governments half the time choose a policy with value $\theta = 40$, and half the time choose a policy with value $\theta = 16$. With full knowledge of the consequences of their actions, governments would prefer to maximize utility by selecting the better policy, but they do not have either the information or the ability to make such a choice. Free mobility implies that population levels adjust endogenously to the choices of θ.

A central government can impose policy harmonization (i.e., identical choices of θ by both jurisdictions), in which case the total population population of six people divides itself evenly between the jurisdictions, with three residents in each location. Suppose that decisionmaking by the central government is similar to that in the two jurisdictions, half the time selecting a policy with value $\theta = 16$, and half the time selecting a policy with value $\theta = 40$.

Setting $y = 0$ for simplicity, consumer utility will therefore be 40 – 9 = 31 half of the time, and 16 – 9 = 7 the other half of the time, for an expected value of 19.

Free policy choice among governments is an alternative to the centralized, and therefore coordinated, outcome just described. With jurisdictions choosing their policies independently, they will choose identical policies half the time (one quarter of the time they both choose the policy for which $\theta = 40$, and one quarter of the time they both choose the policy for which $\theta = 16$), and choose different policies half the time. The expected value of consumer utility is again 19 for the half of the time that independent selection yields identical policy choices, so the interesting comparison lies in the cases in which policies differ.

When one jurisdiction chooses a policy with value $\theta = 40$ while the other chooses a policy with value $\theta = 16$, the population will be drawn into the jurisdiction with the superior public policy, up to the point that greater crowding eliminates the welfare advantage of locating there. Hence five residents will occupy the jurisdiction with the superior public policy, and one resident will occupy the other jurisdiction, all citizens achieving utility levels of 15 (40 – 25 = 16 – 1 = 15). Since an average utility level of 19 is available with centralized policies, decentralization entails an expected reduction in welfare. Table 1 presents these calculations in a schematic fashion.

Why does decentralization reduce expected welfare? The answer is to be found in convex crowding costs. Decentralization produces an inefficient allocation of the population, with excessive crowding in the jurisdiction offering better policies. Population movement offers the potential benefit of allocating more of the population to places with better policies, but this benefit is more than erased by free mobility, which produces excessive crowding costs. The equilibrium condition for population mobility is that utility is the same in every occupied jurisdiction, so population movement may not actually expose citizens to better average policies, since, if every jurisdiction is occupied, then everyone's utility level equals the utility of a resident of a jurisdiction with below-average policies.

It is noteworthy that neither centralization nor decentralization maximizes expected consumer welfare in this example. Consider, for example, a policy in which the jurisdiction adopting the public policy valued at 40 has four residents, and the jurisdiction adopting the policy worth 16 has two residents. Individuals in the jurisdiction with better policies and more crowded streets have utilities of 24, while those occupying the other jurisdiction have utilities of 12. Expected utility is therefore 20 *[(4/6)(24) + (2/6)(12) = 20]*, which exceeds average utility with centralization (19) and average utility with decentralization (15). Expected utility is maximized by policies that entail differences between the utilities of residents of the two locations whenever policies differ, a set of outcomes that is impossible to support with free population mobility.

Figure 1 depicts average utility as a function of *m*, the population of the jurisdiction with the superior public policy. Since the total population is 6, average utility (*E(u)*) is the sum of (*m/6*) times the utility enjoyed by residents of the jurisdiction with better public policy and *(6-m)/6* times the utility available in the other jurisdiction:

(2) $$E(u) = \frac{1}{6}\left[m\left(40 - m^2\right) + \left(6 - m\right)\left(16 - \left(6 - m\right)^2\right)\right] = -20 + 22m - 3m^2 .$$

Average utility as described in equation (2) is maximized by setting the derivative with respect to *m* equal to zero, a condition that is satisfied by $m = 3\tfrac{2}{3}$, at which point average utility equals $20\tfrac{1}{3}$. Average utility is represented by a symmetric parabola with a peak at $m = 3\tfrac{2}{3}$ – which is why *m = 3*, the harmonized outcome, yields higher utility than *m = 5*, the decentralized outcome, since *3* is closer to $3\tfrac{2}{3}$ than is *5*.

3.2. Continua of policies and jurisdictions.

It is instructive to generalize from the example of binary policy choices to one in which there is a continuum of jurisdictions and consumers. Again assume that individual utility is additively separable in income, the value of the public good, and the cost associated with crowding:

(3) $$u(y,\theta,n) = y + \theta - c(n),$$

in which $c(n)$ is the crowding cost, with $c'(n) > 0$ and possibly $c''(n) > 0$. With a continuum of identical jurisdictions the population indicator *n* can be thought of as the population per square mile of land. Since jurisdictions are identical, *n* is a function of θ, and without loss of (much) generality, it is possible to restrict attention to cases in which the function $n(\theta)$ is continuous, and continuously differentiable, in θ. Subnational jurisdictions draw the policy variable θ from a distribution with a density $f(\theta)$ on the interval $[0,1]$. For ease of analysis it is convenient to assume that the distribution of choices made by local governments exactly mirrors the distribution from which they draw their θs. Hence one can take the θs chosen by local governments as being distribution from which the central government selects its θ.

The (fixed) total population of $\bar{n}$ is distributed among the jurisdictions that can be sorted according to their choices of θ, the only aspect in which they differ:

(4) $$\bar{n} = \int_0^1 n(\theta) f(\theta) d\theta .$$

Thus, the average population density must equal $\bar{n}$, since the total land area is normalized to equal one. The average value of public policies is denoted $\bar{\theta}$, defined as:

(5) $$\bar{\theta} \equiv \int_0^1 \theta f(\theta) d\theta$$

The central government has the ability to compel subnational jurisdictions to harmonize their policies on a value of θ that it chooses. In the absence of significant crowding costs, and with identical choice functions used by the central government and local jurisdictions, decentralization produces a higher average utility level than does centralization. This follows directly from the simple observation that the maximum of a distribution exceeds its mean. Of course, it is not necessarily the case that choice functions are identical: the central government might do a better or worse job at selecting policies than do decentralized authorities. Consider, for example, the case in which the distribution of θ is uniform for choices made by local governments and the central government, but that the upper support of the distribution for the central government has a smaller range: θ might vary uniformly from 0 to *z* when θ is chosen by the central government, whereas θ varies uniformly from 0 to *x* when θ is chosen by subnational governments. Then the expected value of θ when chosen by the central government is *z/2*, whereas with a decentralized system the population will select the best available policy, enjoying a level of θ equal to *x*. If $x > z/2$, then welfare is higher with decentralization.

The introduction of convex crowding costs again changes the outcome quite dramatically. Taking the distributions used by the central and local governments to be identical, it follows that expected welfare in a decentralized system (W_d) equals:

(6) $$W_d = \int_0^1 n(\theta)[\theta - c(n)] f(\theta) d\theta ,$$

whereas expected welfare in a centralized and harmonized system (W_h) equals:

(7) $$W_h = \int_0^1 \overline{n}[\theta - c(\overline{n})] f(\theta) d\theta .$$

The difference between equations (6) and (7) is that equation (6) is an integral that is evaluated across jurisdictions – which can be thought of as states, in the American context – whereas equation (7) is an integral that is evaluated across states of the world.

The comparison of W_d and W_h is greatly facilitated by noting that the free population mobility implies that $[\theta - c(n)]$ takes the same value for any realizations of θ for which $n(\theta) > 0$. If congestion costs are high enough that every jurisdiction has positive population in the decentralized equilibrium, i.e. $n(0) > 0$, then values of $[\theta - c(n)]$ can be freely substituted for each other, so:

(8) $$\int_0^1 n(\theta)[\theta - c(n)] f(\theta) d\theta = \int_0^1 \bar{n}[\theta - c(n)] f(\theta) d\theta .$$

The equality expressed in equation (8) is an essential element in facilitating the comparison of W_d and W_h. It follows from (8) that:

(9) $$W_d - W_h = \int_0^1 \bar{n}\theta f(\theta) d\theta - \int_0^1 \bar{n} c(n) f(\theta) d\theta - \int_0^1 \bar{n}\theta f(\theta) d\theta + \int_0^1 \bar{n} c(\bar{n}) f(\theta) d\theta ,$$

or:

(10) $$W_d - W_h = \bar{n}\int_0^1 [c(\bar{n}) - c(n)] f(\theta) d\theta .$$

The evaluation of whether it is better to centralize or decentralize decisionmaking therefore turns on the sign of $\int_0^1 [c(\bar{n}) - c(n)] f(\theta) d\theta$. With convex crowding costs, this term simply has to be negative, since aggregate costs are minimized by allocating the population evenly across all jurisdictions. Hence harmonized policies produce higher welfare than decentralized policies whenever congestion costs are convex, and of sufficient magnitude to ensure that there is positive population in all jurisdictions. Alternatively, if crowding costs are linear in *n*, then welfare is the same for centralized and decentralized policies; and if crowding costs rise with *n* but do so in a concave fashion, then decentralized policies produce higher average welfare.

Figure 2 depicts the population distribution for a decentralized outcome in which all jurisdictions are populated. The function $n(\theta)$ rises with θ, but does so with a declining slope that reflects the impact of convex crowding costs. As the figure indicates, welfare is lower with decentralization than it is with harmonization, since the population level exceeds $\bar{n}$ in the jurisdiction where the public good has the average value $\bar{\theta}$.

3.3 Outcomes with empty jurisdictions.

The proof of the welfare superiority of policy harmonization in settings with convex costs relies on the assumption that congestion costs and population size are together sufficient to populate every jurisdiction in equilibrium. This is not necessarily the case: as $\bar{n}$ approaches zero, $c(n) \approx c(0)$ and is therefore constant for all jurisdictions, so the population will concentrate in the jurisdiction offering the highest value of θ. The expected utility level in the harmonized outcome is $\bar{\theta} - c(\bar{n})$, so if, in the decentralized outcome, the jurisdiction whose policy induces a value $\bar{\theta}$ has a population less than $\bar{n}$, then decentralization is welfare-superior to harmonization. Figures 3a and 3b illustrate the two possibilities. In the setting depicted in Figure 3a, welfare is greater with harmonization, whereas in the setting depicted in Figure 3b, welfare is greater with decentralization.

The convexity of the crowding cost function determines the difference between welfare with decentralization and welfare with harmonization. In order to probe the impact of crowding costs it is useful to consider the effect of population size, since as the population rises, the convexity of the crowding function becomes more important.

One question that naturally arises in a setting in which decentralization produces a higher welfare level than harmonization is whether a greater national population always increases the relative merits of harmonization. It is clear that there is a tendency in this direction, since decentralization produces higher welfare if the national population is small enough to make crowding costs unimportant, whereas harmonization produces higher welfare if the population is sufficiently great that every jurisdiction is occupied. It is, however, possible to construct examples in which a small increase in population triggers a greater reduction in welfare with harmonized policies than it does with decentralization. Consider, for instance, the case in which policies are decentralized, some jurisdictions are unoccupied, and the density $f(\theta)$ is unusually high at the value of θ for which jurisdictions are just barely on the margin of positive occupation. A small increase in national population then largely sorts itself into the marginal jurisdictions, with very little impact on average welfare. Under the alternative of policy harmonization, the same increase in population would contribute to crowding and thereby reduce welfare more substantially.

The somewhat counterintuitive example in which rising national population has a greater impact on welfare with harmonized policies reflects that marginal crowding costs are largely unaffected by population growth if, with decentralization, there is a concentration of jurisdictions just at the margin of occupation. Rising population is more directly associated with increased crowding costs under decentralization if the distribution of policies $f(\theta)$ is uniform, and in this case it is possible to demonstrate that, starting from a situation in which decentralization is produces a higher welfare level than harmonization, greater population increases the relative merits of harmonization.

Consider the impact of population change on welfare enjoyed by residents of the jurisdiction for which, with decentralization, $n(\theta)=\bar{n}$. Greater population reduces welfare in this jurisdiction, and every jurisdiction, by imposing greater crowding costs. Since welfare is the same in every populated jurisdiction with decentralization, it is sufficient to ask whether the population in the jurisdiction for which $n(\theta)=\bar{n}$ rises by more or less than the national average as *n* increases. If the population of this jurisdiction rises by more than the national average, then population growth makes decentralization worse compared to harmonization, since with centralized policies every jurisdiction has a population of $\bar{n}$ that grows at the national average.

In order to evaluate the impact of population growth it is helpful to start from the equilibrium condition that, for some constant utility level *k*, $[\theta - c(n)] = k$ for all populated jurisdictions. Population growth has the effect of reducing *k*, from which it follows that:

(11) $$\frac{dn(\theta)}{dk} = \frac{-1}{c'(n)}.$$

Equation (11) implies that population accumulates primarily in the jurisdictions with little population, for which $c'(n)$ is smallest. In order to evaluate whether the jurisdiction at which $n(\theta)=\bar{n}$ is above or below this average, it is helpful to differentiate the population equilibrium condition with respect to θ, obtaining:

(12) $$\frac{dn(\theta)}{d\theta} = \frac{-1}{c'(n)}.$$

Comparing equations (11) and (12), it immediately follows that the derivative of a jurisdiction's population with respect to a rise in the national population is proportional to the local derivative of the population function with respect to the quality of public policies. Hence the welfare question – whether rising population is worse under decentralization or with centralization – can be answered by asking whether the change in population for a given change in θ at the point at which $n(\theta) = \bar{n}$ is greater than or less than the population average value of this derivative. This rather different, but nonetheless equivalent, formulation of the question can be answered by imposing the requirement that the entire population must locate *somewhere*.

Since $n(0) = 0$, the average value of $n'(\theta)$ is simply $n(1)$, a property readily verified from the fundamental theorem of the calculus: $\int_0^1 n'(\theta)d\theta = n(1) - n(0) = n(1)$. Given that the range of θ is [0, 1], it follows that $n(1)$ is also the slope of a line drawn from the origin to the $n(1)$ point in Figure 4. The impact of population growth on the relative virtue of decentralization then turns on whether the slope of this line is more steep or more shallow than the value of $n'(\theta)$ at the point at which $n(\theta) = \bar{n}$.

If crowding costs are convex, so that $c''(n) > 0$, then the $n(\theta)$ function is concave, a property that follows from differentiating (12) to obtain:

(13) $$\frac{d^2 n(\theta)}{d\theta^2} = \frac{-c''(n)}{[c'(n)]^3}.$$

The concavity of the $n(\theta)$ function in turn implies that $n(1)$ must exceed $2\bar{n}$. The reason is that if instead $n(1) < 2\bar{n}$, then some of the population would be missing. Since the average population level is $\bar{n}$, and by assumption (that the decentralized outcome produces a higher welfare level than harmonization) the $n(\theta)$ function intersects the $n = \bar{n}$ line to the right of $\bar{\theta}$, then even with a linear $n(\theta)$ function it is not possible to satisfy the $\int_0^1 n(\theta)d\theta = \bar{n}$ condition; and the concavity of the $n(\theta)$ function only makes this requirement more stringent.

The fact that $n(1) > 2\bar{n}$ implies that the slope of $n(\theta)$ at the point that $n(\theta) = \bar{n}$ must exceed $\bar{n}$, since otherwise it would not be possible for the $n(\theta)$ function ever to reach the value $n(1)$. This follows from the fact that the value of θ for which $n(\theta) = \bar{n}$ exceeds $\bar{\theta}$, whereas the ray from the origin to $n(1)$ intersects the line $n(\theta) = \bar{n}$ at a value of θ below $\bar{\theta}$, as depicted in Figure 4. Hence it follows that, if θ is uniformly distributed, and decentralization produces greater welfare than harmonization, a growing population and the accompanying congestion costs everywhere reduces the attractiveness of decentralization compared to the alternative of harmonization.

4. Optimal population allocation and corrective policies.

It is useful to characterize the optimal allocation of the population in the model analyzed in section 3.2. Given the continuity assumptions, the optimum can be characterized by the first-order condition that adding a single person to the population of any jurisdiction has the same impact on total welfare. Aggregate social welfare is given by $\int w(\theta) f(\theta) d\theta$, in which $w(\theta)$ is welfare attributable to the jurisdiction with policy θ: $w(\theta) = n[\theta - c(n)]$. It follows that the change in aggregate welfare from adding one more unit of population to this jurisdiction is:

$$\frac{dw(\theta)}{dn} = \theta - c(n) - nc'(n). \tag{14}$$

The first order condition for the optimal population allocation is that the right side of equation (14) is the same for all values of θ.

It is immediately clear from equation (14) why decentralization cannot produce an efficient allocation of population except in the extreme case of no congestion costs, in which case the population is concentrated at the jurisdiction featuring the highest value of θ (and the variational method used to establish that the constancy of the right side of (14) is required for the optimum is no longer valid). Free population mobility implies that $\theta - c(n)$ is the same for all occupied jurisdictions, so the right side of equation (14) can take the same value for all jurisdictions only if $nc'(n)$ does not change with θ. This condition is satisfied either if *n* is constant (which is impossible, since *n* rises with θ), or if $c(n) = k\ln(n)$, for some constant *k*, a function that is not convex in *n*.

The optimal population allocation is one for which the value of (14) is constant for all values of θ, which implies:

(15) $$\frac{dn}{d\theta} = \frac{1}{[2c'(\theta) + nc''(\theta)]}.$$

From equation (15) it is clear that it is not optimal to allocate the population evenly among jurisdictions, since the right side of (15) is positive. Starting from an even distribution of the population across jurisdictions, a welfare improvement is possible by allocating population toward jurisdictions with higher values of θ, and this comes at no additional congestion cost, since $c'(n)$ is the same everywhere. Since harmonization of government policies is equivalent to allocating population evenly across values of θ, it follows that harmonization is not an optimal policy.

What policies support an optimal allocation of population? Inspection of (14) reveals that optimal congestion charges do so, as long as congestion charges equal $nc'(n)$ in equilibrium. Alternatively, appropriate redistribution of resources from congested jurisdictions to less populated rural jurisdictions can be tailored to achieve an optimal allocation of population, assuming that these redistributions affect individual utilities and therefore individual location patterns. There is also a separate question of whether it is sensible to consider optimal congestion charges or interregional distribution policies in a framework in which governments are fallible.

To the extent that any pecuniary cost associated with residence in a jurisdiction rises with population concentration, population mobility is reduced and the resulting allocation of population more closely resembles the optimal pattern described by (15). Consider, for example, the case in which each jurisdiction has a single unit of land, portions of which are traded on active markets. Each resident of a jurisdiction with population *n* then consumes *1/n* units of land; an individual's exogenous income includes the value of land with which he is endowed, and let $L\left(\frac{1}{n}\right)$ be the direct utility that a consumer receives from consuming *1/n* units of land. The price of a unit of land is therefore $L'\left(\frac{1}{n}\right)$, and, if utility is additive and separable in income, a consumer enjoys satisfaction equal to:

(16) $$u(y,\theta,n) = y + \theta - c(n) + L\left(\frac{1}{n}\right) - \left(\frac{1}{n}\right)L'\left(\frac{1}{n}\right).$$

Setting the derivative of this expression with respect to θ equal to zero yields:

(17) $$\frac{dn}{d\theta} = \frac{1}{c'(\theta) - \frac{1}{n^3}L''\left(\frac{1}{n}\right)}.$$

This, in turn, can be compared to the implied optimal population allocation:

(18) $$\frac{dn}{d\theta} = \frac{1}{2c'(\theta) + nc''(\theta) - \frac{1}{n^3}L''\left(\frac{1}{n}\right)}.$$

Examination of (19) and (20) reveals that if all congestion costs are priced, so that $c' = c'' = 0$, then decentralized land markets produce efficient outcomes – a reassuring but hardly surprising implication. Comparing (17) and (18) to their counterparts, in the absence of a land market, (13) and (15), it is clear that the introduction of priced congestion externalities, even if only partially correcting the problem, improves the welfare properties of decentralized government.

5. Extensions.

The model used in this paper is extremely stylized, deliberately so, in order to focus on the consequences of endogenous location of population. In this model, individuals are perfectly aware of the impact of public policies, whereas the governments that enact the policies, and that are presumably directed by at least some of these knowledgeable individuals, for some reason are unable to correct their policy choices in light of what they know. In a different type of model that incorporates learning, presumably not all of the equilibration would need to come from population movement, but also from convergence of policies. The virtue of this type of policy experimentation and subsequent widespread adoption is often thought be a point in favor of fiscal decentralization (e.g., McLure 1986, Oates 1972), though as Strumpf (2002) notes, this conclusion need not follow since the federal authority has better incentives than subnational authorities do to engage in optimal experimentation.

The model posits that subnational and central governments draw from the same distribution of underlying policies; indeed, as presented, the model assumes that the federal government draws from the distribution represented empirically by subnational government choices. This is largely a modeling convenience; modifying the assumptions to give the central government or the local jurisdictions different mean values of their draws carries obvious implications for the desirability of decentralized decisionmaking. If local governments make independent draws from an underlying distribution of policies, then there is a continuum of possible final distributions of decentralized policies, rather significantly complicating the analysis, and increasing the potential desirability of decentralization by increasing the likelihood that decentralization entails unoccupied jurisdictions in equilibrium. Conversely, there is something of a deadening effect on the variance of outcomes if governments make independent draws of many different policy parameters, increasing the likelihood that all jurisdictions are occupied in equilibrium, and thereby improving the relative desirability of harmonization.

Risk aversion is another potential consideration that generally weighs in favor of decentralization in models such as that analyzed in this paper. In reflecting many draws from the same distribution, decentralized policy adoption in effect offers a form of insurance that may be unavailable if harmonization from the center entails single uniform policy choices. The analysis in sections 3 and 4 considers expected outcomes in a risk-neutral setting, but if individuals are risk-averse decentralization starts to look even more attractive.

The model relies on congestion disutilities to support an equilibrium population allocation in the face of differing qualities of public policies. It is entirely possible that congestion effects are negative over some ranges, that there are locally increasing returns to greater concentrations of individuals and economic activity in some locations. These increasing returns must, however, turn negative at some point, otherwise the entire world would congregate at one location offering slightly better policies than others, and the analysis in this paper can be understood to apply to the regions in which congestion is costly. In a similar vein, there may be individuals who are attracted to certain jurisdictions for reasons untouched by any of the considerations in the model, and whose locations are therefore unaffected by public choices or any policies designed to correct congestion externalities. Explicit consideration of this fraction of the population would affect the analysis very little, other than contributing to underlying congestion costs.

In practice, many policies and institutions offer what may be indirect correctives to the problem posed by excessive responsiveness of population levels to variation in the quality of public policies. While few locations impose explicit congestion-related charges, zoning and development regulations, rent control, and other policies that regulate land use may have the effect of reducing urban concentrations. Many governments redistribute resources away from heavily populated areas, some doing so explicitly, others implicitly through tax and other institutions that extract resources from those earning market rents from agglomeration or desirable public policies (Albouy, 2007). It is noteworthy that while the capitalization of local attributes in land prices does not entirely correct the problems associated with congestion, this capitalization is often sufficient to make decentralization more attractive than harmonization directed from the center.

6. Conclusion.

Competition among governments affords a range of choices that typically would be unavailable if instead centralized authorities made decisions to which all others were expected to conform. Whether choice is a good or a bad thing depends critically on how individuals react to the choices they confront, and what the consequences of their actions are for themselves and for others. Even if individuals are fully rational, their behavior in a setting with significant congestion externalities can create outcomes in which the greater choice made available by decentralization has unfortunate ramifications for everyone. The pricing of congestion based externalities in a number of forms – through the land market, intergovernmental transfers, and congestion charges of various types – reverses these effects, however, and restores the benefits of greater choice in a wide range of cases.

References

Albouy, David, The unequal geographic burden of federal taxes and its consequences: A case for tax deductions? Working paper, University of California – Berkeley, May 2007.

Bardhan, P. and Dilip Mookherjee, Capture and governance at local and national levels, ***American Economic Review***, May 2000, 90 (2), 135-139.

Besley, Timothy and Stephen Coate, Centralized versus decentralized provision of local public goods: A political economy approach, ***Journal of Public Economics***, 2003, 87, 2611-2637.

Bewley, Truman, A critique of Tiebout's theory of local public expenditures, ***Econometrica***, May 1981, 49 (3), 713-740.

Buchanan, James M. and Charles J. Goetz, Efficiency limits of fiscal mobility: An assessment of the Tiebout model, ***Journal of Public Economics***, 1972, *1* (1), 25-43.

Buchanan, James M. and Richard E. Wagner, An efficiency basis for federal fiscal equalization, in Julius Margolis ed., ***The analysis of public output*** (New York: Columbia University Press, 1970), 139-158.

Edwards, Jeremy S. and Michael Keen, Tax competition and Leviathan, ***European Economic Review***, 1996, 40, 113-135.

Gordon, Roger H., An optimal taxation approach to fiscal federalism, ***Quarterly Journal of Economics***, November 1983, 98 (4), 567-586.

Hoyt, William H., Local government inefficiency and the Tiebout hypothesis: Does competition among municipalities limit local government inefficiency, ***Southern Economic Journal***, October 1990, 57 (2), 481-496.

Inman, Robert and Daniel Rubinfeld, Rethinking federalism, ***Journal of Economic Perspectives***, 1997, 11 (4), 43-64.

Kollman, Ken, John H. Miller and Scott E. Page, Political institutions and sorting in a Tiebout model, ***American Economic Review***, December 1997, 87 (5), 977-992.

Kollman, Ken, John H. Miller and Scott E. Page, Decentralization and the search for policy solutions, ***Journal of Law, Economics and Organizations***, April 2000, 16 (1), 102-128.

Lockwood, Ben, Distributive politics and the costs of centralization, ***Review of Economic Studies***, 2002, 69, 313-337.

McLure, Charles E., Tax competition – Is what's good for the private goose also good for the public gander? ***National Tax Journal***, September 1986, 39 (3), 341-348.

Oates, Wallace, ***Fiscal federalism*** (New York: Harcourt Brace, 1972).

Oates, Wallace E., Searching for Leviathan: An empirical study, ***American Economic Review***, 1985, 75, 748-757.

Oates, Wallace E., An essay on fiscal federalism, ***Journal of Economic Literature***, September 1999, 37 (3), 1120-1149.

Strumpf, Koleman S., Does government decentralization increase policy innovation? ***Journal of Public Economic Theory***, April 2002, 4 (2), 207-241.

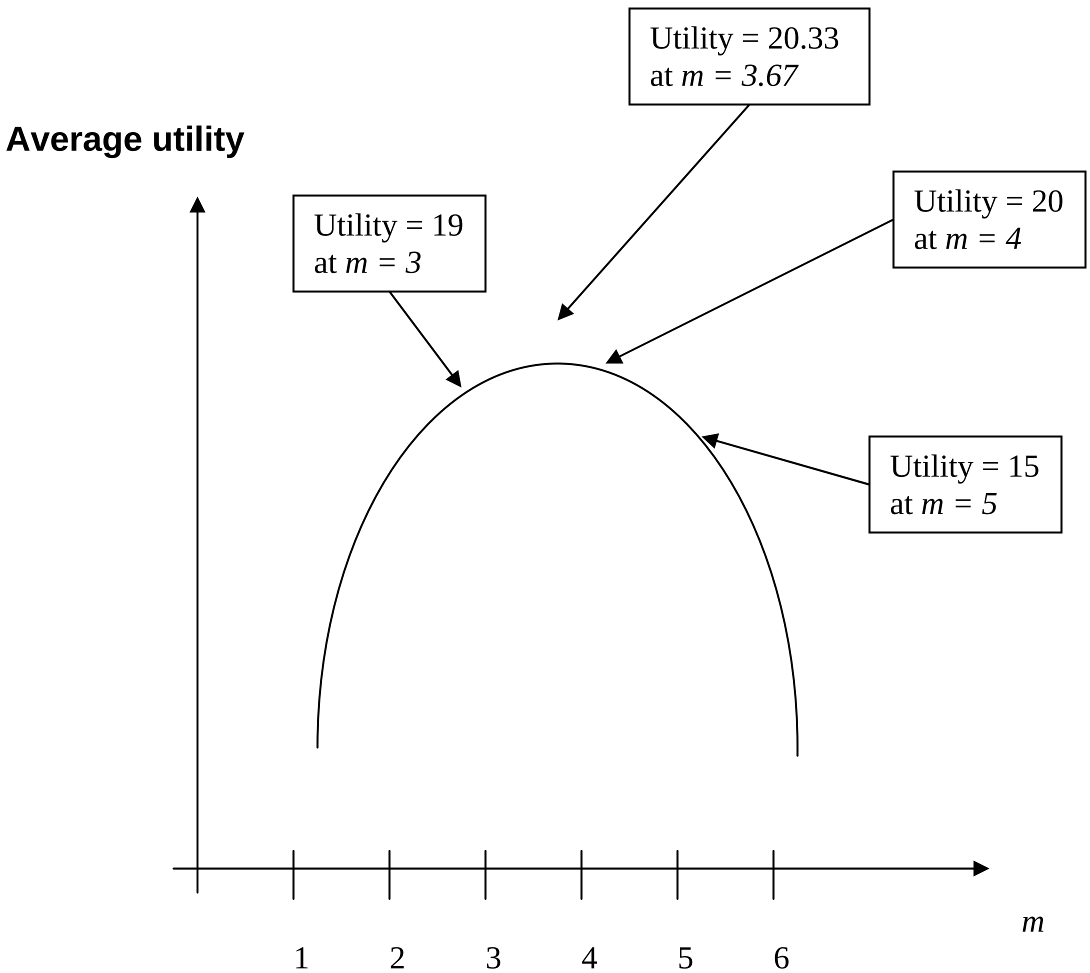
Tiebout, Charles M., A pure theory of local expenditures, ***Journal of Political Economy***, 1956, 64 (5), 416-424.

Wallis, John and Wallace E. Oates, Decentralization in the public sector: An empirical study of state and local government, in Harvey S. Rosen ed., ***Fiscal federalism: Quantitative studies*** (Chicago: University of Chicago Press, 1988).

Wellisch, Dietmar, ***Theory of public finance in a federal state*** (Cambridge, UK: Cambridge University Press, 2000).

Figure 1

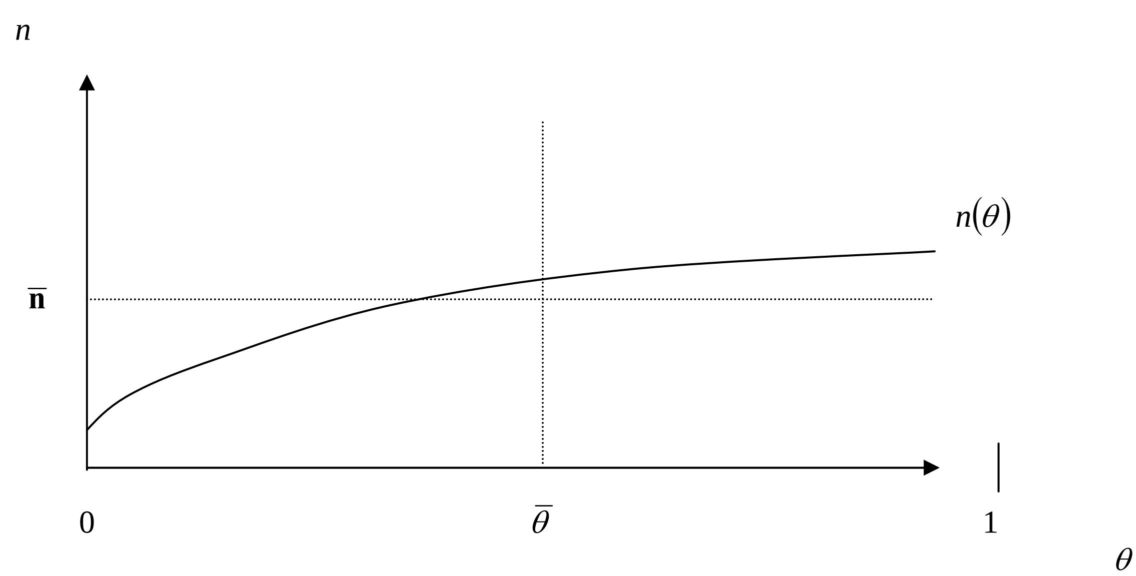
Population allocation and average welfare in the binary choice example



Note to Figure 1: The figure depicts average utility levels for consumers with utility functions given by $u(y,\theta,n) = \theta - n^2$. The total population of six is divided between two jurisdictions, with *m* consumers in the first jurisdiction, for which $\theta = 40$, and (*6 – m*) in the second jurisdiction, for which $\theta = 16$.

Figure 2

Population distribution with every jurisdiction occupied



Note to Figure 2: the figure depicts the population distribution with decentralized policies distributed over the [0, 1] interval, and crowding costs sufficient to populate every jurisdiction in equilibrium. The welfare superiority of policy harmonization is reflected in the fact that $n(\bar{\theta}) > \bar{n}$.

Figure 3a

Decentralization reduces welfare even with unoccupied jurisdictions

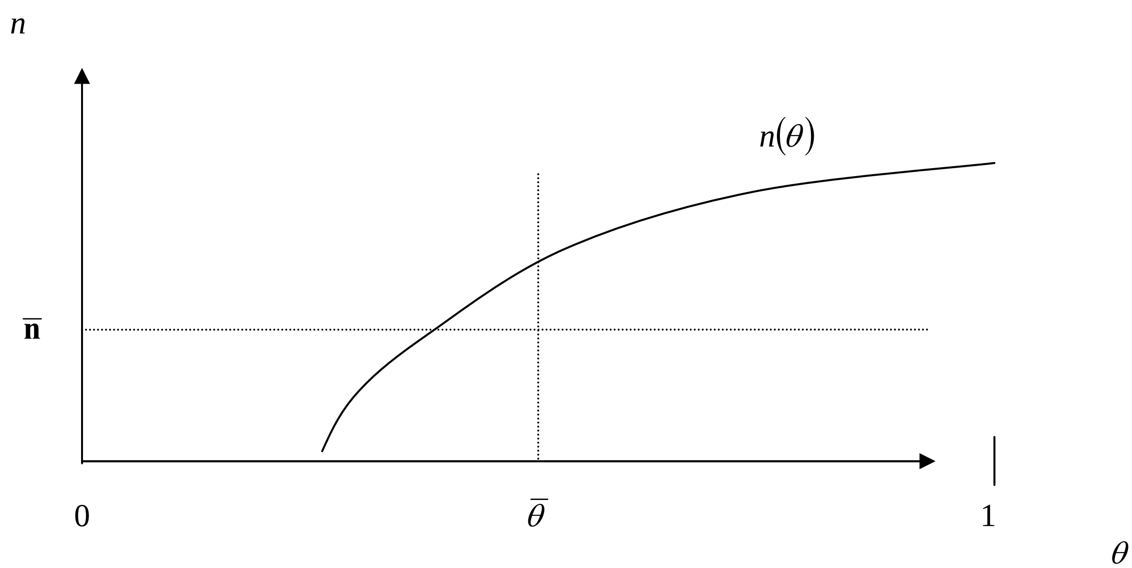


Figure 3b

Decentralization increases welfare with unoccupied jurisdictions

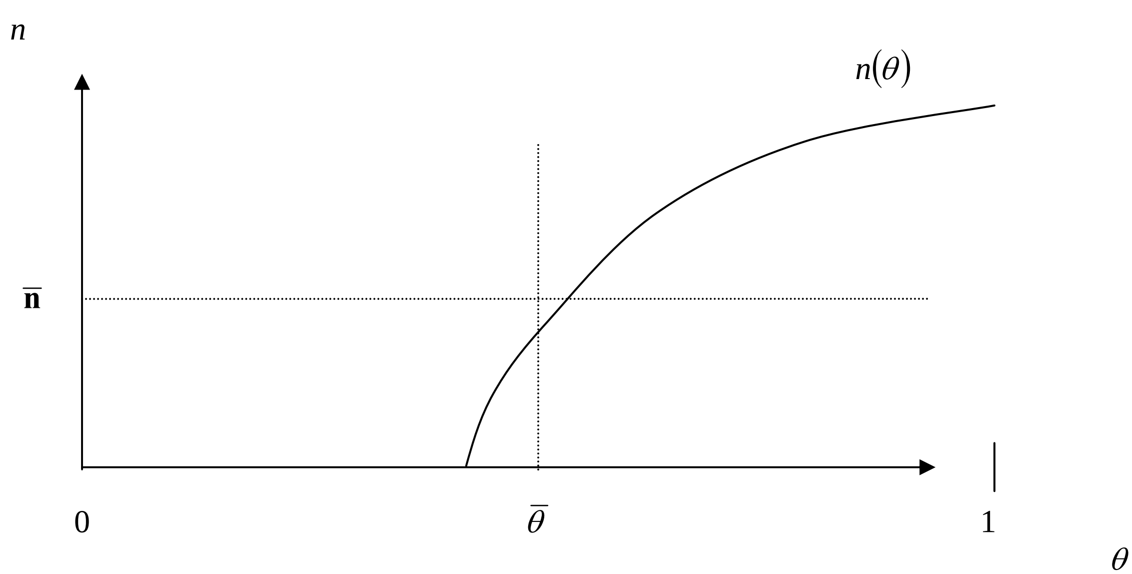
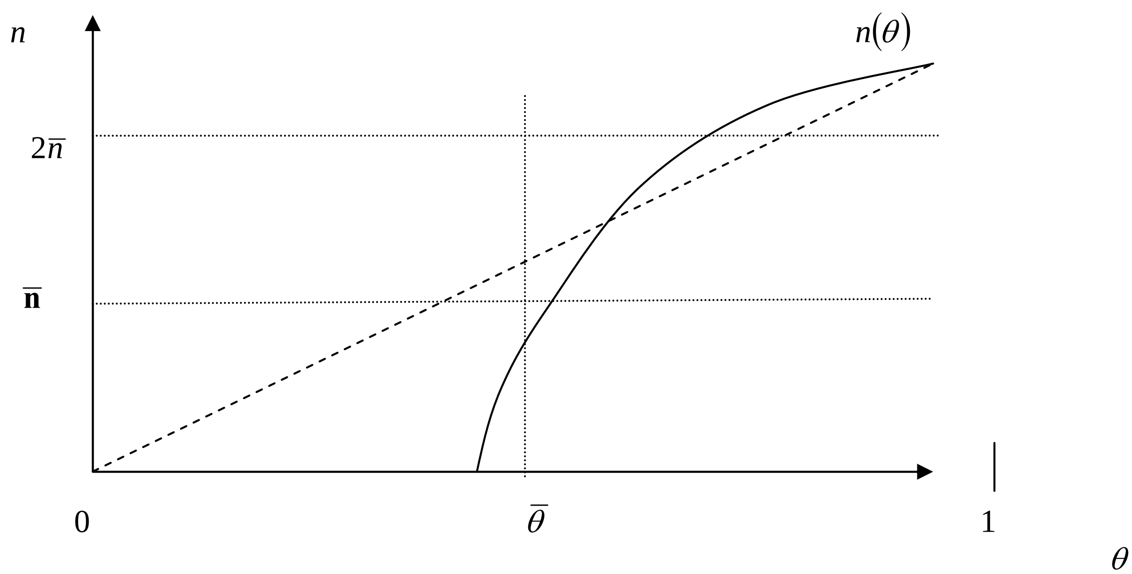


Figure 4

The effect of population growth on the difference between welfare in harmonized and decentralized systems.



Note to Figure 4: the figure depicts a decentralized outcome in which θ is uniformly distributed on the [0, 1] interval and welfare is greater under decentralization than with harmonized policies. Rising population reduces welfare with decentralization more than it reduces welfare with harmonized policies, a proposition that can be demonstrated by showing that the slope of the $n(\theta)$ function, evaluated at the value of θ for which $n(\theta)=\bar{n}$, exceeds $\bar{n}$.

Table 1

Welfare with alternative policy regimes

	Decentralization	*Harmonization*	*An alternative*
Jurisdiction 1	$\theta = 40$ m = 5 u = 40 – 25 = 15	$\theta = 40$ m = 3 u = 40 – 9 = 31	$\theta = 40$ m = 4 u = 40 – 16 = 24
Jurisdiction 2	$\theta = 16$ n = 1 u = 16 – 1 = 15	$\theta = 16$ n = 3 u = 16 – 9 = 7	$\theta = 16$ n = 2 u = 16 – 4 = 12
Average welfare	15	19	20

Note to Table 1: The table presents average welfare levels obtained with differing allocations of six individuals between two jurisdictions, one jurisdiction adopting a policy with a value of 40 for all residents, and the other adopting a policy with a value of 16. Each individual experiences crowding costs equal to the square of the jurisdiction's population. In the first column ("Decentralization"), individuals are free to move between jurisdictions, so utility levels are the same for residents of each location. In the second column ("Harmonization"), the population is evenly allocated between the two jurisdictions. And in the third column ("An alternative"), four members of the population are assigned to the jurisdiction with the policy worth 40, and two are assigned to the jurisdiction with the policy worth 16.