

Assessing the Importance of Tiebout Sorting: Local Heterogeneity from 1850 to 1990*

Paul W. Rhode**
UNC Chapel Hill and NBER

Koleman S. Strumpf
UNC Chapel Hill

December 2001
(Revised July 2003)

Abstract

This paper argues that long-run trends in geographic segregation are inconsistent with models where residential choice depends solely on local public goods (the Tiebout hypothesis). We develop an extension of the Tiebout model that predicts as mobility costs fall, the heterogeneity across communities of individual public good preferences and of public good provision must (weakly) *increase*. Given the secular decline in mobility costs, these predictions can be evaluated using historical data. We find *decreasing* heterogeneity in policies and proxies for preferences across (i) a sample of US municipalities (1870-1990); (ii) all Boston-area municipalities (1870-1990); and (iii) all US counties (1850-1990).

JEL classification: D7, H7, N3, R5

Keyword: segregation, regional convergence

*We thank David Blau, Judith Blau, Dennis Epple, Claudia Goldin, Thomas Mroz, Felix Oberholzer, Wilbert van der Klaauw, participants at the 2000 AEA, 1999 NBER Development of the American Economy Summer Institute, 2001 Public Choice, and 2001 SSHA meetings, ZEI Workshop on Federalism and Decentralization, University of Chicago, Clemson University, University of Illinois Urbana-Champaign, University of Kentucky, University of Michigan, UNC-Duke, and Wharton School seminars, and two anonymous referees for comments and suggestions. This is a substantially revised version of an earlier paper, "A Historical Test of the Tiebout Hypothesis."

**Corresponding author. UNC Chapel Hill, Department of Economics, CB# 3305, Chapel Hill NC 27599. email: prhode@email.unc.edu.

In recent years there has been renewed emphasis on decentralized governance in many countries including the United States. A key rationale for this shift is the belief that local governments provide policies better suited to citizen preferences. This wisdom is grounded in the Charles Tiebout (1956) hypothesis which states that individuals will costlessly sort themselves across local communities according to their public good preferences. This simple theory is the workhorse of the local public finance literature and has been the subject of over one thousand economics and political science articles.

Tiebout sorting remains an active current research topic, with many recent papers taking a strict interpretation of the model. As examples, Dennis Epple and Holger Sieg (1999) and Epple et al., (2001) empirically model community choice as the product of costless sorting on housing prices and public good provisions. They estimate the underlying preference parameters under the maintained hypothesis of a Tiebout equilibrium. Other recent empirical papers have used the Tiebout framework to evaluate the effects of school competition (Epple et al., forthcoming), school choice (Caroline Hoxby, 2000) or to explain the number of local jurisdictions (Alberto Alesina et al., 2000). The theoretical local public economics literature also relies heavily on the Tiebout framework and often presumes that community selection is driven exclusively by public goods and taxes. Some prominent recent examples include Raquel Fernandez and Richard Rogerson (1998), Hoxby (1999), and Thomas Nechyba (1999, 2000), who consider education quality/spending; Jan Brueckner (2000), who analyzes local tax competition; Dennis Epple and Thomas Romer (1991), who investigate redistribution; Myrna Wooders (1999), who interprets Tiebout using cooperative game theory; Fernandez and Rogerson (1997), who study the effects of zoning; and Gerhard Glomm and Roger Lagunoff (1999), Nechyba (1997), and Carlo Perroni and Kimberley Scharf (2001), who analyze generic local public goods.

Local policies clearly matter for residential choice, but are they the dominant motive? Suggestive evidence to the contrary comes from the American/Annual Housing Survey (AHS), a longitudinal, nationally-representative survey of over 50,000 homes begun in 1973. Among the AHS households who moved in the previous year, only 5 percent cited public services (including schooling) as their primary reason for moving.

Roughly 50 percent said their move was primarily due to employment or family and friends, motivations excluded in the Tiebout model and the literature cited above.¹ These results indicate that non-Tiebout incentives are important and perhaps driving forces in residential decisions. If individuals select communities in large part due to employment or social interaction opportunities, then neighbors need not have homogeneous public good preferences and one of the central implications of the Tiebout model is violated.

This paper seeks to assess more formally the relevance of Tiebout sorting. Our strategy is to derive a more realistic version of the Tiebout hypothesis and empirically to test its implications. We first generalize the Tiebout model by relaxing the assumption of perfect mobility (no moving costs). We show in a general environment that as mobility costs fall, resident preference heterogeneity *across* communities should (weakly) increase. Under some standard assumptions on individual preferences, the variation of policies across communities will also increase. The empirical section begins by documenting the dramatic reduction in mobility costs over the last two centuries. This suggests that if Tiebout incentives are of first-order importance, then heterogeneity across communities will increase in the historical record.²

To evaluate this prediction, we consider: (i) a sample of US municipalities over the 1870-1990 period; (ii) all municipalities in the Boston metropolitan area over the 1870-1990 period; and (iii) all counties in the United States over the 1850-1990 period. Almost all of our empirical results stand in opposition to the Tiebout prediction of increasing heterogeneity across communities.³ Across the US municipality sample,

¹This result is quite robust. The motives behind moves are similar if we consider only within-metropolitan area moves; if we consider future moves by households who say their current local services are “so inadequate that they want to move”; or if we consider data reports from the Current Population Survey which includes migration data in 1945-46, 1962-63, 1974-76, 1997-2000. All data sources and details on these calculations are contained in the Data Appendix, which is available at the authors’ websites.

²An alternative cross-sectional test of the Tiebout prediction is to see if increases in mobility costs lead to reduced stratification across communities. Using data for 65 large metropolitan areas in 1980 and 1990, we find higher commuting costs are associated with *greater* across municipality heterogeneity (results available upon request). This is an important complement to our main results because it does not suffer from standard criticisms of long time series (e.g. changes in the aggregate preference distribution or in the nature and type of local public services).

³These results run counter to the conventional wisdom that greater sorting has occurred in the latter part of the twentieth century. We show that most contemporary segregation occurs between neighborhoods (as measured by Census tracts) within the same municipality. Such neighborhoods receive roughly the same

heterogeneity of local policy outcomes—total local taxes per capita and school taxes per capita—has declined significantly. The coefficient of variation for school taxes fell by two-thirds between 1880 and 1992. To test the comparative static prediction regarding preference heterogeneity, we consider proxies for public good preferences: race, age, and nativity, and over the 1970-90 period, education, home ownership and income. These proxies generally exhibit diminishing heterogeneity across our municipality sample. For example, the dissimilarity index for black population share decreased from 0.72 to 0.57 between 1870 and 1990. We replicate our results using all municipalities in the Boston metropolitan area because some argue the Tiebout model should apply to small geographic regions. Even with a greater variety of preference proxies and policies including electoral outcomes and education spending, there is little evidence of increased stratification (except for racial composition, but there is no change in suburban heterogeneity and the city-suburb differences appear likely due to discrimination rather than local public goods).

We next consider county-level data. While some researchers consider counties too large to be considered communities, we show empirically that across-municipality heterogeneity is roughly equal to across-county heterogeneity plus a constant. This means that *trends* in across-county heterogeneity parallel trends in across-municipality heterogeneity. The advantage of using county-level data is that a wide range of variables is available for the full national population of counties. We assemble a vast dataset comprising all of the counties in the United States (except Alaska) over the 1850-1990 period. Our results confirm the municipal-level analysis. The dispersion of local policy outcomes across counties has declined significantly since the late nineteenth century. The coefficient of variation for local per capita education spending fell from 0.66 in 1890 to 0.25 in 1992. A similar reduction in heterogeneity occurred in per capita taxes and revenues over the 1870-1992 period. We consider numerous proxies for public good preferences including religious affiliation. Almost every preference proxy exhibits diminishing heterogeneity across counties since 1850. Two of the more graphic examples are that the proportion of blacks living in black majority counties decreased

level of local public services, and so such within-municipality stratification is difficult to explain with the

from 48 percent in 1890 to 9 percent in 1990, and that the dissimilarity index of presidential vote shares decreased from 0.27 to 0.17 between 1892 and 1988. These patterns are not solely driven by changes in the South, by rural-urban migration, or by reduced salience of our preference proxies.

In total, these results suggest that Tiebout sorting has been historically overwhelmed by forces reducing across-community heterogeneity. (A referee has noted that a closer examination of the data indicates an even greater discrepancy between the historical record and the comparative static predictions of the generalized Tiebout model. Almost every heterogeneity series declines most strongly over the late nineteenth and early twentieth centuries when the reductions in transportation and communication costs appear most rapid.) These findings do not mean that Tiebout motives are irrelevant, but rather that they have not been the primary factor in long-run location decisions. This implies that any theoretical or empirical model that adopts a pure Tiebout framework, as is common in the literature, is misspecified. In more general models where non-policy factors influence residential choice, many implications of the Tiebout theory no longer hold (for example, it is not typically possible to rank communities according to public good demands).

It is important to contrast our approach with previous empirical tests of the Tiebout hypothesis. Most papers investigate the extent of heterogeneity within communities, the motives for household mobility, and the degree to which fiscal policies are capitalized into property value (see Keith Dowding et al., 1994).⁴ These papers do not provide a basis for evaluating whether Tiebout incentives are of first-order importance. For example, when considering community composition it is unclear how large a deviation from perfect sorting is needed before concluding that non-Tiebout incentives dominate locational choices. Our comparative static approach provides a more meaningful assessment of Tiebout's importance because it implies a direction of

Tiebout model.

⁴An alternative Tiebout test considers whether greater population heterogeneity leads to increases in the number of local governments (see Ronald Fisher and Robert Wassmer, 1998 and the citations therein). But this literature is problematic because the empirical results are weak with many insignificant and wrong-signed parameters, the direction of causality is unclear, and the results are consistent with alternative sorting motives such as racism (see Jorge Martinez-Vazquez et al., 1997).

change—to greater sorting—that is empirically refutable. We recognize that no single piece of evidence presented here is convincing by itself, but the absence of historical sorting trends among the dozen or so measures we analyze constitutes a serious challenge to the view that community choice is primarily driven by Tiebout incentives. Our work suggests that non-Tiebout motives must matter and that a more general approach is needed.⁵

Our empirical results are of independent interest because they contribute to two current literatures. First, they advance the segregation literature, which explores the spatial dispersion of racial, religious, and ethnic groups. No other paper has explored segregation trends over such a long time period using such a wide variety of variables. Our analysis complements David Cutler et al., (1999) who study the segregation patterns of blacks in urban areas between 1890 and 1990. Our results are also consistent with and extend Michael Kremer (1997) who finds little change in across-tract heterogeneity in education over the 1960-1990 period. Second, our paper contributes to the growing literature on the efficiency implications of heterogeneity (e.g. Ronald Benabou, 1996). Alesina and La Ferrara (2000) show that within-community heterogeneity empirically reduces participation in various social groups while Alesina et al (1999) find ethnic diversity decreases local provision of core public services like education. Claudia Goldin and Lawrence Katz (1999) find that variation in high school graduation rates across US states at the beginning of the twentieth century was tied to the degree of religious heterogeneity. Our work provides new evidence on historical trends in several measures of community heterogeneity.

This paper is organized as follows. The next section extends the Tiebout model to include mobility costs and derives the theoretical prediction that the remainder of the paper tests: as mobility costs fall, heterogeneity across communities increases. Section II documents the long-run decline in selected measures of moving costs and Section III presents the empirical approach. Section IV investigates historical variation across

⁵Among the motives deserving more consideration are preferences for neighbors (e.g. racial discrimination or ethnic capital) and the role of employment (including commuting costs). Patrick Bayer (2000) improves upon the empirical literature by allowing residential choice to depend upon employment location and community racial composition.

municipalities in local policies and in population characteristics that proxy for public good preferences. Section V extends the analysis to the county level. The final section discusses implications of our analysis. A Data Appendix listing the sources used in our analysis is available at the authors' websites.

I. Adding Mobility Costs to the Tiebout Model

This section develops a model in which communities provide public goods and individuals, who belong to types characterized by their preferences for public goods, choose communities subject to mobility costs.⁶ As in the original Tiebout model, there is no property or employment, and communities cannot exclude individuals. We make no assumption about the initial distribution of types across communities.

Setup.

Consider a population of N individuals, indexed by i , distributed across C communities. Letting c_i be the community containing agent i , call $\mathbf{A}=(c_1, c_2, \dots, c_i, \dots, c_N)$ the *allocation* of the N individuals over the C communities. Each community c provides public goods, $\mathbf{G}_c \in \Gamma$ where Γ is a compact set. \mathbf{G}_c can be a vector of local policies, each of which may be real valued (such as taxes and spending) or unordered and categorical (such as school curriculum contents).⁷ Denote the set of community public goods as $\mathcal{G}=(\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_C)$.

We will assume that agents only care about \mathbf{G}_c in their community.⁸ Further assume that each agent belongs to a fixed type t characterized by the continuous utility function, $U_t(\mathbf{G}_c)$. Let \mathbf{G}_t be the unique ideal array $\mathbf{G}_c \in \Gamma$ for type t , and presume there are

⁶Mobility costs have been added to other locational choice models. Some examples are William Carrington et al., (1996), Zvi Hercowitz and David Pines (1997), John Kennan and James Walker (2000), and David Wildasin and John Wilson (1996). These papers contain specific assumptions (such as treating government policy as fixed or presuming agents are identical) which preclude using them to generalize the Tiebout model.

⁷We implicitly have a bound on returns to scale in provision of government services. This is necessary to preclude formation of very large and heterogeneous communities, which is also an issue in the original Tiebout model (see Truman Bewley, 1981 for examples).

⁸That is, individuals only care public good provision and not the characteristics of their neighbors. In principle, richer neighbors are more desirable because they contribute a greater tax share to the community budget constraint. The model implicitly rules out such income heterogeneity or presumes that only head taxes are possible.

T types where $T \leq N$. In some of the results derived below, we will consider special assumptions with a scalar public good. In order of increasing restrictiveness they are:

Assumptions:

A1. Single-peaked preferences: $G_c \in \mathfrak{R}$ and $U_t(G_c)$ is a twice-differentiable concave function in G_c , where $U_t''(G_c) < 0$, $U_t'(G_c) > 0$ for $G_c < G_t$, $U_t'(G_c) < 0$ for $G_c > G_t$, and $U_t'(G_t) = 0$.⁹

A2. Quadratic preferences: $G_c \in \mathfrak{R}$ and $U_{tc} = -(G_t - G_c)^2$.

Social Welfare.

Define the aggregate measure of social welfare for any allocation \mathbf{A} and set of community public goods G as the sum of all agents' utility:

$$(1) \quad W = \sum_c \sum_i U_{ti}(G_c).$$

The functional form of (1) is not essential and we discuss generalizations below.

Community decisions.

Suppose that each community c chooses its policy G_c^* to maximize the sum of utilities of its current residents:

$$(2) \quad G_c^* = \operatorname{argmax}_{G_c \in \Gamma} \sum_j U_{tj}(G_c)$$

Given our assumptions, G_c^* exists. Note that some of the communities, z , may be empty, implying $G_z^* \in \emptyset$. The functional form of (2) is not essential for our analysis; what is important is that (2) has a parallel structure to (1).¹⁰

Individual Location Decisions and the Equilibrium Concept.

Assume that the agents can move in some sequential order, i.e. one individual at a time. This ordering may be deterministic or stochastic, as long as each agent's expected order in the sequence is finite. Refer to each agent's turn to move as her *location*

⁹Single-peaked preferences can be understood as an individual maximizing a utility function containing a public good and a private good subject to a budget constraint including a tax for the public good. See Thomas Romer and Howard Rosenthal (1977).

decision event. When her decision event occurs, agent i can change communities at the cost of m_i units of utility. This “mobility cost” may be individual specific.¹¹

Assume that the mobility decisions are myopic. That is, each agent i takes the prevailing policies, \mathcal{G} , as given (thereby ignoring how her move affects the communities’ decisions or causes other individuals to move) and only considers migrating to the community currently yielding the highest utility for her type. In the migration decisions, assume each agent treats any empty community as setting policies equal to her ideal.

Definition of a myopic move: Under the *myopic movement rule*, agent i of type t_i moves from community d to community c at her location decision event if and only if:

$$(3) \quad \mathbf{G}_c^* = \operatorname{argmax}_{\mathbf{G}_e \in \mathcal{G}} U_{t_i}(\mathbf{G}_e^*) \text{ where } e \text{ is a generic community; and}$$

$$(4) \quad U_{t_i}(\mathbf{G}_c^*) > U_{t_i}(\mathbf{G}_d^*) + m_i.$$

In equilibrium, no individual will move when her decision events occur.

Definition of an equilibrium: An *equilibrium* is an allocation \mathbf{A} of individuals across communities such that no agent would choose to move at her location decision event given her mobility costs, m_i .

Results.

Tiebout’s famous claim is that if mobility costs are zero and the number of communities C is at least as large as the number of types T , then individuals of each type will sort themselves into homogeneous communities providing their ideal public good bundle. It is easy to show the following proposition, which captures the Tiebout Hypothesis.

Proposition 1. If $C \geq T$ and policies are set via (2), then W is maximized when each community contains only one type.

¹⁰(2) is used as a leading case. It is equivalent to majority rule with side payments in a world with transferable utility. See James Buchanan and Gordon Tullock (1962) pp. 190-92 for a rationalization of side-payments and intuition about why they induce efficiency.

¹¹We assume that mobility costs for individual i are constant across communities. The results below will not change if these costs vary with some measure of “distance”.

Proposition 1 makes it clear that population heterogeneity within communities reduces welfare in the Tiebout model. This point can be further illustrated by the following example. Consider a model with quadratic preferences, **A2**. To maximize resident welfare via (2), a community c (with N_{tc} residents of type t , making a total of $N_c = \sum_t N_{tc}$) should set $G_c^* = \sum_t (N_{tc}/N_c) G_t$, the population-weighted mean of the ideal policies. Welfare per capita in community c , W_c/N_c , will then equal the negative of the population-weighted variance of the ideal policies:

$$(5) \quad W_c/N_c = -\sum_t (N_{tc}/N_c) (G_t - G_c^*)^2 = -\sum_t (N_{tc}/N_c) G_t^2 + (\sum_t (N_{tc}/N_c) G_t)^2.$$

Per capita welfare in community c would be maximized if its population were homogeneous. In this example, the degree of within-community heterogeneity can naturally be measured by the population-weighted variance. In the general case, appropriately measuring the degree of heterogeneity is more difficult. The important point is that in the Tiebout model *social welfare is positively related to increased sorting* (that is, lower within-community heterogeneity and higher across-community heterogeneity).¹²

By assuming that preferences satisfy the single-peaked condition **A1** and that public good provisions are set via (2), we can show that sorting is a self-reinforcing “increasing returns” process. The movement of an individual of type s increases the attractiveness of the receiving community-- and reduces the attractiveness of the sending community-- for all type s agents. It has the opposite effects for some other types. (This is formerly developed in Appendix A.) Even if other agents’ mobility costs are fixed, the movement of a single agent can have cascading effects, inducing the movement of others. In general, the dynamics can be quite complicated. The outcome depends upon the distribution of agents (their types and individual mobility costs) across communities and upon the specification of the sequential moving order (which may be stochastic). This complexity motivates the myopic moving rule, which supposes an agent does not try to

¹²In the above example, total welfare is $W = -\sum_t N_t (G_t - \chi)^2 + \sum_c N_c (G_c^* - \chi)^2$ where $\chi \equiv (\sum_t N_t G_t)/N$ is the aggregate population-weighted mean type/equilibrium policy. The first term is constant for all allocations of a given population and the second term measures across-community heterogeneity of types (the variance of the community mean types.) A move from any allocation to perfect sorting increases W and, by the above expression, increases across-community heterogeneity.

solve through the general equilibrium implications of her move on the subsequent public good provisions or the movements of others.¹³ An additional motivation is that in a large population, a single individual has a negligible direct effect on the provision of public goods.

Proposition 2 shows that any myopic move has a positive effect on social welfare and, as a consequence, any reduction in mobility costs has a non-negative effect on social welfare. Notice that no special restrictions on individual preferences are needed for this result.

Proposition 2: When individual moves obey (3) and (4) and policies are set via (2),

- a. Any individual move strictly increases W (and does so by more than the mover's costs, m_i).
- b. If m_i falls, then individual i either stays or moves and if she moves, then W increases. The moving process yields a new equilibrium with a higher W .

Proof:

a. Suppose that individual i moves from community d to c . The utility of three groups of agents will be affected. First, the net effect among residents of community d except i (k/i) is,

$$(6) \quad \sum_{k/i} [U_{t_k}(\mathbf{G}_{d/i}^*) - U_{t_k}(\mathbf{G}_d^*)] \geq 0$$

where the inequality follows from the definition of argmax in (2). The intuition is that a community cannot be made worse off by adjusting \mathbf{G} to maximize the welfare of its current residents. The remaining residents of d (excluding i) are, by definition, at least as well off in aggregate under $\mathbf{G}_{d/i}^*$ as under \mathbf{G}_d^* . We call this the argmax argument.

Second, the net effect among the initial residents of community c (j/i) is,

$$(7) \quad \sum_{j/i} [U_{t_j}(\mathbf{G}_{c+i}^*) - U_{t_j}(\mathbf{G}_c^*)] \leq 0$$

where the inequality follows from the definition of argmax in (2). Finally, the effect on agent i has two components. From her myopic comparison of d and c ,

$$(8) \quad U_{t_i}(\mathbf{G}_c^*) - U_{t_i}(\mathbf{G}_d^*) - m_i > 0$$

where the inequality follows from (3) and (4). The other component considers how her move will change \mathbf{G} in c ,

$$(9) \quad U_{t_i}(\mathbf{G}_{c+i}^*) - U_{t_i}(\mathbf{G}_c^*).$$

But (7)+(9) equal,

¹³For a sense of the complexity of this process, see Ken Kollman et al., (1997).

$$(10) \quad \sum_{j/i} [U_{t_j}(\mathbf{G}_{c+i}^*) - U_{t_j}(\mathbf{G}_c^*)] + U_{t_i}(\mathbf{G}_{c+i}^*) - U_{t_i}(\mathbf{G}_c^*) = \sum_j [U_{t_j}(\mathbf{G}_{c+i}^*) - U_{t_j}(\mathbf{G}_c^*)] \geq 0$$

by definition of the argmax in (2). The intuition for (10) follows the argmax argument. In aggregate, the residents of c (including i) are at least as well off under \mathbf{G}_{c+i}^* as under \mathbf{G}_c^* . The change in \mathbf{G}_c may harm its initial residents, but the gain to i must more than offsets their losses; otherwise, \mathbf{G}_{c+i}^* would not be selected by a community setting policies according to (2).

Thus the total effect (6)+(7)+(8)+(9) is positive. This implies that welfare net of moving costs, $W - m_i$, increases, and therefore, W increases.

b. If agent i moves, this may induce others to migrate. By part a, no matter how many moves occur, W increases. This process must end in a finite number of moves (that is, there exists an equilibrium) because there are a finite number of possible allocations and each agent's expected order in the location decision sequence is finite. No allocation can re-occur because W is strictly increasing with each move.

Q.E.D.

Three comments are in order:

(i) Proposition 2 also holds in a world of Leviathan governments where G is fixed. Here, (6), (7) and (9) are each zero, but (8) is positive by (3) and (4).

(ii) More generally, the basic results hold if the social welfare function (1) weakly reflects individual preferences (as in $W = F(U_1, U_2, \dots, U_N)$ with $F'_i \geq 0$), and in (2) each community maximizes its residents' welfare (as measured by this W).

(iii) The myopic moving rule, which implies that (8) is positive, plays a key role in Proposition 2. Suppose instead that individuals are forward-looking and move if (8)+(9) is positive. If (8) is negative, then in principle (6)+(7)+(8)+(9) could be negative and the proposition does not hold (W may decrease).¹⁴

¹⁴In general, mobility inherently involves externalities, both positive and negative. Under the myopic movement rule, an individual moves only if the receiving community is ex ante preferable. This means that any move that benefits the individual also benefits society on net; that is, the benefits received by the mover and the other residents of the sending community exceed the costs imposed on existing residents of the receiving community. Under the non-myopic rule, an individual may find moving beneficial simply because it makes the receiving community closer to her own tastes (ex post). This can impose costs on its existing residents that are greater than the benefits that mover and the other members of the sending community enjoy. Note that under either the myopic or non-myopic rules, there may be socially beneficial moves that are not made when the private benefits fall short of the mobility costs.

A further issue of interest is how a reduction in mobility costs affects the distribution of policy outcomes. Because it is difficult to define variation in multidimensional space, we focus on cases with scalar public goods.

Observation: Under **A1** with $C=T=2$ or under **A2**, when local policies are set via (2) then individual moves which obey (3) and (4) lead to increased variation of policy outcomes across communities.

A formal demonstration of the result is contained in Appendix A.¹⁵ Consider first the case of single-peaked preferences, **A1**, when $C=T=2$ holds. In this case, a community's policy will be the weighted average of the two types' ideal policies where the weight on a type's preference depends positively on its population share. Any myopic move will widen the difference in policies between the two communities by pushing the policy in the receiving (sending) community toward (away from) the mover's ideal G . Now consider the second case which assumes quadratic preferences, **A2**. Under (2) and **A2**, the policy in a given community is the mean of residents' ideal G 's and the aggregate population-weighted mean policy is independent of the distribution of types across communities. Any move obeying (3) and (4) will widen the difference in policies between the sending and receiving communities, increasing the total population-weighted variance of policies.

To summarize, the theoretical model developed in this section extends the Tiebout framework to include mobility costs. Our results, while somewhat novel, are clearly in the spirit of Tiebout's argument. Mobility costs may prevent individuals from sorting into homogeneous communities of their own type-- the allocation that maximizes social welfare. A reduction in mobility costs has a non-negative effect on social welfare. This implies falling mobility costs increase heterogeneity across communities. We also find that sorting increases the variation of local policy outcomes under several variants of the model.¹⁶ The remainder of the paper tests these predictions empirically.

¹⁵The observation also holds for generic preferences when the population is initially completely diffuse, implying policies are identical across communities. Moving to any level of greater sorting, the variation of policies across communities will be weakly greater, and in the fully sorted equilibrium, the variation will be strictly greater.

¹⁶Note that these results depend on the Tiebout assumption that residential choice is based only on public good provision. If individual resident decisions weigh factors (employment, proximity to family,

II. Documenting Declining Mobility Costs

The conventional wisdom holds that mobility costs have fallen over time.¹⁷ Yet constructing a comprehensive measure to document this “truism” is difficult. Anyone who has relocated knows that out-of-pocket expenditures represent only a fraction of the costs of moving. As human capital theory suggests, these costs include the lost work time-- organizing before departure, traveling, and getting back up to speed at the destination. Given that real wages have generally risen, the value of this lost time would be increasing. However, several opposing forces more than offset this effect. Improvements in transportation and the increased similarity of regional cultures mean less time is now lost in the move. During the colonial period, the rigors of the transatlantic travel and the effects of exposure to a new disease environment were purportedly so severe that newly imported slaves and indentured servants required six months to two years to achieve positive levels of net output (David Galenson, 1996). Few migrants suffer such a loss today. The available evidence suggests that over the past one-and-one-half centuries, the reduction in direct travel time has more than offset the increase in the value of labor.¹⁸ In addition, as is argued below, communication improvements have reduced one of the key costs of moving, the lost contacts with one’s friends and family in the home community.

There is clear evidence that physical moving costs have fallen over the last century. (In the interest of brevity, all sources are contained in the online Data Appendix.) The most obvious change is the spread of the personal automobile. In 1900,

amenities) not perfectly correlated with preferences over G_c , then reduced mobility costs need not induce greater Tiebout sorting. Indeed, given the policy rule (2), migration for non-Tiebout reasons may reduce the variation of G and dampen Tiebout incentives to migrate.

¹⁷To perform our comparative static exercise, an exogenous reduction in mobility costs is needed. While much of the transportation infrastructure (such as highways and airports) involves endogenously determined public goods, the reduction in transportation and communications costs was in large part a product of technological improvements (in steam and internal combustion engines, communications equipment, and production techniques).

¹⁸Between 1857 and 1999, the time required to travel between New York and Chicago has fallen from 2 days to less than 2 hours 40 minutes, or by a factor of 18; that for a trip between New York and Los Angeles has fallen from about 3.5 weeks to less than 8 hours, a factor of 75. The time for short distance trips has also sharply decreased: the average automobile speed in 1970 is roughly 12 times the stage speed in 1840. Average real labor returns have increased between 5 and 7-fold over the 1860-2000 period.

there was roughly one passenger car for every 10,000 Americans; today, the ratio is nearly one car for every two. This change was due in part to sharp reductions in the costs of owning and operating automobile. The careful calculations of Hiram P. Maxim, a leading engineer, showed driving costs in 1903 equaled 143.8 cents per mile in 1998 dollars. The American Automobile Association estimates that the full cost of driving the more reliable, comfortable cars of today average 54.9 cents. As the fragmentary data on automobile costs per mile (excluding finance charges) presented in Figure 1 indicate, most of this decline occurred before the Second World War. Also facilitating the spread of the personal car were massive investments in the nation's system of public roads, nearly doubling its mileage from 2.3 million in 1900 to almost 4 million today.¹⁹ An important consequence of the spread of the automobile was to weaken the link between work and residence locations, potentially allowing greater Tiebout sorting.

Improvements in trains and airplanes have also significantly lowered mobility costs. As the series in Figure 1 reveal, the real cost of railroad service was about one-third as expensive in 1995 (13.4 cents per passenger mile) as it was in 1895 (37.4 cents). The real cost of air travel also fell sharply, with average airline revenues per passenger-mile dropping from about 108 cents in 1929 to 13.7 cents by 1995 (rough parity with railroads). In addition, the speed of air travel nearly tripled since the early-1940s.

We also know that the real cost of moving household goods has fallen substantially. Circa 1995, the real rate per ton-mile for a private COD shipment by a household goods carrier averaged 57.8 cents, which is far less than the 88.1 cents charged a decade before. Tariff schedules filed with the Interstate Commerce Commission indicate that the real rate per ton-mile for a "modern" shipment was approximately 147 cents in 1936, implying costs have fallen by over 60 percent between 1936 and 1995.²⁰ The costs of local moves have declined as well. For example, when L. S. and Anna

¹⁹This growth in mileage understates the true improvement in transportation access because most early roads were little more than dirt pathways. In 1904, for example, "surfaced" roads made up less than seven percent of total mileage. The first coast-to-coast auto trip across North America, completed in 1903, purportedly took 65 days (<http://www.nam.org/AboutMfg/timeline1901.html>).

²⁰It appears that the average weight of shipments has also risen, climbing from somewhat under 2-tons in the early period to 3-tons today, but this proportional increase is less than the fall in rates and is of course endogenous with respect to the price decrease. Over the 1994-96 period, the average billed shipment weighed just under three tons (5919 pounds) and traveled 1261 miles.

Shoen established the U-Haul Co. in 1945 at Ridgeway, WA, they charged \$2 per day for a small trailer. Today renting such a trailer at I-5 Auto Sales near Ridgeway (or at many of the firm's other 15,000 retail locations) would cost \$9.95 or about 40 percent less in real terms.

Long-run movements in communication costs reveal similar trends. This is important for several reasons. Lower communication costs improve information flows about other regions, reducing uncertainty. In addition, they allow migrants to maintain contacts with friends and family "back home." Finally, easier communication encourages more dispersed production activity, implying people are less tied to a particular community for employment reasons. Figure 2 shows the real costs of making three-minute daytime telephone calls from New York to Chicago and San Francisco have fallen almost continuously. To place a three-minute transcontinental call in January 1915 (when service first became available) cost \$20.70 in current dollars, which was almost \$314 in 1998 dollars. The real cost of such a call in 1995, even at ATT residential daytime rates, was less than *three-tenths of one percent* as high.²¹

These falling mobility costs have apparently set more Americans on the move. In 1940, about 11 percent of the American population (five years and older) had lived in a different county five years earlier. This fraction increased to 17 percent in 1970 and to 19 percent by 1990. Another useful measure of long-run mobility rates is the percentage of the native population residing in their state-of-birth. In 1870, almost 77 percent of the native population resided in their state-of-birth. Since 1900, the fraction has continuously fallen, with the most rapid rate of decline occurring during the 1940-70 period. By 1990, only about two-thirds of the native population resided in their state-of-birth. Today's migration rates appear sufficiently high to allow the American population to achieve significant sorting across local jurisdictions according to policy preferences, if they so desired.²²

²¹The reduction in postal rates, especially across country, was also notable. In 1860, during the Pony Express period, it cost \$10 to send a one-ounce letter between New York and San Francisco. By 1886, the cost fell to two cents in the currency of the day.

²²Consider a population composed of two equally sized groups that are initially evenly distributed across two regions. If 4 percent of the population moved every year in accordance with Tiebout "voting with their feet" thinking, the regions would be completely segregated within 12.5 years.

III. Empirical Implementation

Given the secular decline of mobility costs, a natural test of the prediction of the generalized Tiebout model is to examine historical trends in the dispersion of local fiscal outcomes and in the sorting of population types across localities. While it is not clear how to define Tiebout's canonical community, the most natural definition is the municipality or the Census minor civil division (MCD). Unfortunately, electronic versions of comprehensive MCD-level data are not available before 1970. Instead we created a random 10-percent sample of counties and entered data for all MCDs in these counties for the earlier years.²³ We investigate heterogeneity trends across this MCD sample. Because some argue that Tiebout sorting only applies over limited geographic areas, we next extend the analysis to the full set of municipalities in the Boston metropolitan area.²⁴ Finally, we present additional results using the full set of US counties. One potential complication is the growing number of municipalities and counties. This is mainly due to territorial division, which will lead to increases in our across-community heterogeneity measures even in the absence of individual movement (see Paul Rhode and Koleman Strumpf, 2000 for details).²⁵

The analysis investigates trends in across-community heterogeneity of policy variables and proxies for public good preferences. (Details on the availability, precise definition, and sources for all variables are presented in the online Data Appendix). Our local policy variables are per capita taxes and expenditures.²⁶ Taxes are a measure of the overall level of government activity; we consider both overall and school taxes. We also examine spending on education and protection (police and fire), the most prominent local

²³We cannot directly sample from the population of MCDs since there is no historically consistent listing of all municipalities. We entered information for years prior to 1970 where the Census tabulated MCD-level data. The years with coverage of all MCDs are: 1870 (demographics only); 1880, 1890 (government finances only); 1930, 1940, and 1960 (demographics only). We have data for all years but 1940.

²⁴We also explore more broadly whether sorting occurs over small geographic areas in the entire US. Our analysis indicates that heterogeneity across adjacent communities has tended to remain flat over time. (The results are based on county-level data and are available upon request.)

²⁵Because this is the direction of change that the Tiebout model predicts, finding empirical evidence of reduced sorting would be strong evidence against the model.

²⁶Richard Ely (1888) contains a detailed discussion of the development and historical comparability of our taxation measures.

services. For these policy variables, two heterogeneity measures are calculated, both of which increase with dispersion across governments and control for changes in mean levels. The first measure is the population-weighted coefficient of variation (CV), which is the ratio of the standard deviation to the mean,

$$(11) \quad CV \equiv G^{-1}(\sum_j P_j(G_j - G)^2)^{0.5}$$

where G_j is per capita taxes/revenues/spending for government j , G is mean per capita taxes/revenues/spending for all governments, and P_j is the share of total population in government j . The second measure (DG) calculates the proportion of total taxes/revenues/spending in each year which would have to be re-allocated across governments to yield a uniform, per capita distribution (this is related to the dissimilarity index which is discussed below),

$$(12) \quad DG \equiv \frac{1}{2}G^{-1}\sum_j P_j|G_j - G|$$

Because preferences are not directly observable, we adopt the strategy of examining numerous characteristics that proxy for individual types.²⁷ For each characteristic, we partition the population into mutually exclusive and exhaustive categories. Our proxies are admittedly imperfect measures of the true types. But as an earlier version of the paper (Rhode and Strumpf, 2000) shows, even if the observable characteristics are noisy signals or available categories are too coarse, our measures of population heterogeneity remain informative.²⁸ This version also shows that the proxies retain their salience over time (in year-by-year regressions using the proxies to explain variation in local policy or election outcomes, the R-squared's do not trend down) and that the trends in heterogeneity across communities are not exclusively the result of the shift from a rural to urban society.

²⁷An ideal test of our model would involve construction of multi-dimensional measures of individual types (i.e. using the characteristics discussed below as inputs in a hedonic model of type). However, this would require detailed, individual-specific information about *all* persons living in a given local jurisdiction. Such data simply do not exist for the modern era.

²⁸There is also evidence that several of our type measures are transmitted from parents to children. Thomas Piketty (1995) cites the extensive literature showing that political preferences have an important hereditary component even after controlling for income and social class. Frank Newport (1979) finds that in the mid-1970s over two-thirds of individuals maintain their childhood religion. Kremer (1997) shows there is a high rate of transmission of parental education to their children using the PSID (he also surveys work documenting the intergenerational transmission of many other characteristics considered here).

We consider the following proxies when available:

- *Race*: It is often observed that members of racial groups share economic interests and maintain strong common political affiliations. For example, the General Social Survey (GSS) reports 39.3 percent of blacks ($N=1,864$) identify themselves as “Strong Democrats” while only 12.5 percent of whites ($N=3,675$) do so.²⁹ There are also notable racial differences in the GSS over political ideology and attitudes towards government redistribution. We use the black population share to proxy for these beliefs.

- *Age categories*: The young population share (those between 5 and 20 years old) proxies for families with children; such households presumably prefer higher spending on local schools. The old population share (those at least 65 years old) is used, since the elderly should be less likely to favor education spending.³⁰ These variables should reflect lifecycle sorting.

- *Nativity*: The foreign-born represents another distinctive population with important ramifications for local politics (e.g. school curricula). Note that interpreting trends for the foreign-born share is complicated because immigrants may sort across communities for non-Tiebout motives, for example, to take advantage of social networks or ports of entry.

- *Party vote shares in presidential elections*: Individuals presumably vote for the party whose platform is closest to their own ideal policy, implying those voting for a particular party have similar preferences.³¹

- *Religion*: Individuals affiliated with a particular religion share their faith’s set of beliefs, values, and cultural traditions and are, therefore, likely to have relatively similar

²⁹The General Social Survey (1999) is a micro dataset of individual attitudes collected over the 1972-96 period.

³⁰There is also some support from the GSS that age groups have distinct political beliefs. For example, while 5.8 percent of those aged 18-20 ($N=1,181$) considers themselves “Strong Republicans,” 16.8 percent of those aged 75 or older ($N=2,311$) do. The GSS also indicates that similar age differences exist for the appropriate level of education spending.

³¹If all individuals in a community vote for the same party, the community is composed of like-minded residents and is sorted in the Tiebout fashion. If the residents split their votes, the community has not been sorted. It is important to notice that this measure only makes sense for elections over national office. This is because party platforms are strategically set with the objective typically being vote maximization. Even relatively homogeneous areas may split their vote on local offices because the local party platforms are likely to be quite similar. For national offices, however, parties are likely to set their platforms in a way to

policy preferences.³² In the GSS, 26.5 percent of self-identified religious fundamentalists ($N=162$) and 23.1 percent of evangelicals ($N=208$) consider themselves to be “Strong Republicans” while only 4.5 percent of religious liberals ($N=265$) do so. One of the advantages of using religious affiliation is that it allows a fine partition of the population: we can employ up to 27 denominational families.

- *Home-ownership rates*: Home-owners are typically wealthier and have greater civic involvement in the community.³³

- *Education*: Educational attainment is likely to be related to income, wealth, and attitudes toward government.³⁴ We use three groups: less than a high school degree, at least a high school degree but not a college degree, and a college degree or more.

- *Income*: This is the most natural measure of type. Unfortunately, the Census did not begin reporting data on local income *distribution* until 1949. Categorical information is available for both families and households (which include unattached individuals). The Census lists 14 income groups in 1949, 15 income groups in 1969, 17 income groups in 1979 and 25 groups in 1989.

To ensure robustness we employ several heterogeneity measures for our proxies. For variables with discrete types, the dissimilarity index and the Gini coefficient are used.³⁵ These measures, which are commonly used in the segregation literature, have

split the *national* vote. Individuals in a relatively homogeneous area are likely to have similar preferences over national parties, and so they will cast their votes for only one party.

³²Based on the 1990 National Election Study, David Legee and Lyman Kellstedt (1993) show that affiliation with many of the denominational families used in our analysis are strong predictors of individual voting behavior and ideological preference. Laurence Iannaccone (1998) suggests that the link between religion and politics is largely limited to moral and social issues such as school prayer and abortion. However, he only focuses on evangelical-fundamentalist Protestants. Mark Noll (1990) also documents the historical link between religion and politics using largely non-quantitative analysis.

³³See Robert Carroll and John Yinger (1994) and Denise DiPasquale and Edward Glaeser (1999).

³⁴In the GSS, of those with less than a high school education ($N=9,391$) 15.2 percent earn \$25,000 or more and 25.4 percent think welfare benefits are too low. For those with only a high school degree ($N=20,368$), the values are 40.5 percent and 17.6 percent; for those with a college degree or more ($N=7,632$), the values are 63.2 percent and 16.7 percent.

³⁵See Otis Duncan and Beverly Duncan (1955), Douglas Massey and Nancy Denton (1988), and Sean Reardon (1998). The dissimilarity index is the most widely used segregation measure. It shows the proportion of individuals who would have to change communities to create an evenly distributed population, expressed as a ratio of the number who would have to move if the types were completely segregated. The Gini coefficient generalizes the dissimilarity index. The main difference is that the Gini is sensitive to any change in the population distribution whereas the dissimilarity index is affected only by shifts in types between ‘surplus’ and ‘deficit’ communities.

three important properties. First, they vary between zero (when each type is equally represented in each community) and one (when the types are completely segregated). Thus a higher value indicates greater heterogeneity across communities. Second, they are normalized to control for the changing proportions of types in the aggregate population, implying they are unaffected if the groups grow at different rates nationally.³⁶ Third, the measures weight the communities by their population. In multiple (≥ 2) type comparisons, the dissimilarity index, D , and the Gini coefficient, GC , are defined as:

$$(13) \quad D \equiv \frac{1}{2} \frac{\sum_t \sum_j N_j |P_{tj} - P_t|}{(N \sum_t P_t (1 - P_t))}$$

$$(14) \quad GC \equiv \frac{1}{2} \frac{\sum_t \sum_k \sum_j N_k N_j |P_{tk} - P_{tj}|}{(N^2 \sum_t P_t (1 - P_t))}$$

where, N_j is the total population in community j , N is the total population, P_{tj} is the share of type t in the community j 's population and P_t is the share of type t in the total population.

We also employ entropy indices, which are additively separable and can be used to perform within- and between-region decompositions. Formally, let P_j be the share of the total population living in j . The community-level entropy index is,

$$(15) \quad E \equiv 1 - \sum_j P_j H_j H^{-1}$$

where $H_j \equiv -\sum_t P_{tj} \log(P_{tj})$, $H \equiv -\sum_t P_t \log(P_t)$. If a type t is absent from community j , then by convention, $P_{tj} \log(P_{tj}) = 0$. The between- and within-region decomposition is,

$$(16) \quad E \quad \equiv E_{\text{Between}} + \quad E_{\text{Within}} \\ = (1 - \sum_R P_R H_R H^{-1}) + \quad \sum_R P_R H_R H^{-1} (1 - \sum_{j \in R} P_{jR} H_j H_R^{-1})$$

where R is the region; P_R is the share of the total population living in R ; P_{jR} is the share of the region R 's population living in j ; P_{tR} is the share of type t in region R 's population; and $H_R \equiv -\sum_t P_{tR} \log(P_{tR})$.

³⁶More formally, suppose that each group reproduces at a different rate and that the offspring live in the same community as the parents. If there are two groups, then both indices are invariant to the group growth rates (proof available upon request).

The heterogeneity measures for income are richer because the data is in ordered categories. The aggregate income distribution can be decomposed into within- and between-community components using the two additively-separable Theil measures,

$$(17) \quad I_1 \equiv \mu^{-1} \sum_j P_j \sum_s P_{sj} \mu_{sj} \log(\mu_{sj}/\mu_j) + \mu^{-1} \sum_j P_j \mu_j \log(\mu_j/\mu)$$

$$(18) \quad I_2 \equiv \sum_j P_j \sum_s P_{sj} \log(\mu_j/\mu_{sj}) + \sum_j P_j \log(\mu/\mu_j)$$

where μ_j is the mean income, P_{sj} is now the proportion of individuals in income group s , μ_{sj} is the mean income of group s , all for community j , and μ is the aggregate mean income (Anthony Shorrocks, 1980). In the formulae, the first term is the within-component and the second term is the between-component. To investigate within-community heterogeneity further, we also consider the Gini income coefficient and the CV. Because the available data is grouped in income ranges, both lower- and upper-bound Gini's are computed for each community j ,

$$(19) \quad GL_j \equiv (2\mu_j)^{-1} \sum_s \sum_t P_{sj} P_{tj} |\mu_{sj} - \mu_{tj}|$$

$$(20) \quad GU_j \equiv GL_j + (\mu_j)^{-1} \sum_s P_{sj}^2 (\mu_{sj} - a_{s-1})(a_s - \mu_{sj})(a_s - a_{s-1})^{-1}$$

where a_s is the upper income boundary for income group s (see Joseph Gastwirth, 1972).

IV. Trends in Heterogeneity Across Municipalities

A. National Sample of Municipalities

We first consider local policy outcomes for our national sample of municipalities (MCDs) and calculate our two heterogeneity measures, the coefficient of variation (CV) and the reallocation index (DG). Panel (a) of Table 1 presents heterogeneity trends for per capita taxes across municipalities over the 1880-1992 period. Dispersion between MCDs markedly *decreased* over the sample. The values in 1992 are roughly half of their 1880 values. Panel (b) shows the long-run reduction in dispersion of school district taxes is even more dramatic. While there has been a mild divergence in school district taxes between 1972 and 1992, the current dispersion levels are well below those prevailing a century ago. The remaining panels (c)-(e) have current operations spending data for the

modern period, 1972-92. There are small reductions in across-MCD heterogeneity for overall spending per capita. Dispersion also falls for the two most prominent local expenditure categories-- protection (police plus fire) spending per capita, and school district education spending per pupil.³⁷

We next consider the available preference proxies: race, nativity, and age. To measure the dispersion of these proxies across municipalities we employ the dissimilarity index (D). Table 2 shows that heterogeneity across MCDs for each of the three public good proxies is lower in 1990 than in 1930, or where available in 1870. The dissimilarity index for racial composition declines by a fifth between 1870 and 1990 while the index for nativity falls by a seventh over the same period. There is also a substantial reduction between 1930 and 1990 in the heterogeneity of the young population while there is a U-shaped pattern for the elderly population. The latter is particularly noteworthy because it runs counter to the tremendous growth of retirement communities in Florida and Arizona. Again the lesson from the preference proxy data is that any small increases in heterogeneity across MCDs over the recent decades should not mislead us; in general, the long-run historical trends indicate convergence.

These results are in conflict with the Tiebout prediction that greater across-community stratification should accompany reductions in mobility costs. The data indicate that as moving became easier municipalities actually become more similar in their residential composition and local policy outcomes. Indeed, the convergence between MCDs was strongest in the early period when the decline in mobility costs appears most rapid. These findings suggest some alternative incentives, working in direct opposition to Tiebout, have been dominating residential location choices.

Such long-run mobility trends have had important effects on community composition. Two pieces of evidence reveal that there is substantial population heterogeneity *within* contemporary municipalities. First, consider the entropy measure of racial heterogeneity across all Census tracts in the US.³⁸ Given the entropy index is

³⁷Given that samples are used, it is best not to read too much into these small reductions in the contemporary data. Still, the results suggest there is no trend towards greater spending heterogeneity.

³⁸Census tracts are the smallest geographic unit with complete coverage of the US in 1980 (the smaller census block only covered MCDs with population exceeding 10,000). The Census considers tracts to be equivalent to neighborhoods. Full details on the definition and evolution of these units is presented in US

additively separable, this statistic can be decomposed into within- and between- MCD components. In 1980 the total entropy index of racial composition across all Census tracts is 0.546 (N=99,935), whereas the within-MCD component is 0.253. This means the racial heterogeneity across tracts within a municipality is about equal to the heterogeneity across municipalities.³⁹ Second, we can decompose the national family income dispersion into Theil within- and between-MCD components. Using all MCDs in 1969 (N=34,842), the total Theil₁ index is 0.265 while the within-MCD component is 0.229; in 1979 (N=34,809), the total index is 0.250 while the within- component is 0.222; in 1989 (N=35,065), the total index is 0.288 while the within- component is 0.243. This finding means that income heterogeneity within MCDs is quite high, contributing over four-fifths to the total dispersion. The Tiebout prediction of homogenous communities appears to be a poor approximation for contemporary municipalities.

B. Boston Metropolitan Area

In addition to the national sample, we replicated and extended our analysis using the 92 municipalities in the Boston SMSA (1980 definition). This area provides an attractive test case for several reasons. Boston has been intensely studied and is often put forward as the archetype of the Tiebout model. Municipalities are the only important local government and directly provide all high-profile public services such as education and protection. There is also clear evidence of a secular decline in intra-Boston SMSA transportation costs.⁴⁰ The area has been fully incorporated with only minor border changes throughout the study period, so we can compare the same set of communities. Detailed data are available for the Boston area dating back to the late-nineteenth century. It is also possible to perform an event analysis and see whether changes in the policy

Bureau of Census (1994), *Geographic Areas Reference Manual*, <http://www.census.gov>, while the data are discussed in the online Data Appendix.

³⁹While this result is partly driven by central cities, when the sample is restricted to MCDs with population less than 50,000 or to those with less than 10 tracts the within-MCD component still contributes 20% of the total. It is worth noting that these figures likely understate within-MCD heterogeneity since the average MCD has only 3.25 tracts (53% of MCDs have only one tract meaning their within-component is calculated as zero).

⁴⁰For example, see the discussion of the development of the Boston mass transit system in George Sanborn, *The Chronicle of the Boston Transit System*. <http://www.mbta.com>.

environment, such as the introduction of property tax limits in 1980 and of mandatory busing in Boston in 1974, are driving stratification patterns. Finally, as with many other metro areas there has been a dramatic shift of population from the city to the suburbs, with Boston's population share dropping from 40 to 20 percent between 1915 and 1990. This shift will itself lead to greater measured stratification.⁴¹

Figure 3 shows across-municipality heterogeneity trends for various political and demographic proxies for public good preferences over the 1855-1990 period. For one index, the black population share, there is greater sorting at the end of the period than at the beginning. This trend, which is consistent with Cutler, et al., (1999)'s findings for a large number of cities over 1890-1990, fits the standard picture of "white flight" in response to the Great Migration of African-Americans to northern metro areas. But we should be careful not to generalize from this observation into a wholesale acceptance of the Tiebout model. First, as Cutler, et al., (1999) note, the trend toward increased concentration reverses after 1970. Second, the rise was entirely due to city-suburb differences: Figure 3 also shows that segregation of blacks within the suburbs has changed little or if anything has declined, over the sample period. Moreover it is not clear that growing racial concentration within an urban area should be attributed to Tiebout sorting but rather could result from racism.⁴² Consistent with this view, in 1950 and 1960 over two-thirds of the between-tract racial heterogeneity under the entropy measure is due to within-municipality dispersion (even excluding Boston, the within-municipality heterogeneity contribution is about half).⁴³ A final reason not to leap from racial segregation trends to a wholesale acceptance of the Tiebout model is that none of the other preference proxies (vote shares, young, foreign-born) shows significant movement to greater sorting over the past century and a half. (Foreign-born

⁴¹All of our measures of heterogeneity across communities will increase under the following thought experiment: suppose the center city has a relatively heterogeneous population and is surrounded by more homogeneous, but distinctive suburbs. Then a representative set of city dwellers move to the suburbs in such a way that no municipality (including the central city) changes its population composition.

⁴²Cutler, et al (1999) uses evidence from housing markets to stress the role of racial discrimination. Such a sorting mechanism will upset a Tiebout equilibrium, since attitudes concerning one's neighbor's race are not likely to coincide perfectly with public good preferences.

⁴³Again the idea is that tracts are neighborhoods. Under a racism model individuals have preferences over who they interact with, and so neighborhoods are greatly stratified. In the Tiebout model there should be

heterogeneity did increase in the post-World War Two period, but this represented a return to the levels of the 1855-75 period.). Indeed, the party vote shares in presidential elections indicate reduced heterogeneity across Boston municipalities over the last fifty years.

The income data also reveal little sign of greater sorting. If individuals are becoming increasingly Tiebout sorted, then income inequality *within*-municipalities should fall and inequality *between*-municipalities should rise. Table 3 shows trends in household income heterogeneity over the 1949-89 period. We first create two within-MCD measures by calculating a CV and Gini index for the income distribution in *each* MCD and then average the MCD values using by population weights. The data in the left panel show that the average within-MCD income heterogeneity has stayed roughly constant. This means the income distribution within each municipality has not changed much over the post-World War Two period. The right panel examines trends in within- and between-MCD income heterogeneity using additively-separable Theil indices. These indices reveal the within- and between-components have changed little. Moreover, the between-component is quite small (less than a fifth of the within-component) which means almost all income heterogeneity is due to within-municipality diversity. Municipalities in the Boston area have strikingly similar income distributions and there has been little movement towards greater sorting over the last fifty years.

Nor is there evidence of increasing long-run heterogeneity of government policies across the MCDs of the Boston SMSA. Table 4 charts the trends in the key series over the 1906-92 period. Although there has been a small increase since the 1960s, the population-weighted CV of total government spending, measured by the per capita current operation budget, fell by a quarter between 1906 and 1992. We also investigate two major spending categories, protection and education, which typically comprise over one-half of total spending in these data. The CV for per capita protection spending fell by 30 percent over the past century. The CV for per capita education spending rose and

no stratification across neighborhoods within a municipality since each neighborhood receives the same level of public services.

then fell; the endpoint is roughly equal to the starting point.⁴⁴ These measures display considerably variability, but there is no observed long-run tendency for the policy CVs to rise as the model predicts. This finding reinforces the results for the national sample.

In conclusion, despite the urban flight from the heterogeneous central city, there is little evidence of increased sorting between municipalities in the Boston SMSA. The only measure giving clear evidence of growing concentration is racial composition between the city and suburbs, and this is likely due to racial discrimination rather than Tiebout sorting. In fact among the demographic variables we consider none ever exceeds the conventional standard of high heterogeneity (a dissimilarity index above 0.6). The population of the Boston SMSA appears to be far from the level of sorting that Tiebout would predict. Finally, changes in the dispersion of the spending variables or any of the preference proxies are not strongly linked to changes in the policy environment such as school busing in the 1970s or tax limits in the 1980s.

V. Trends in Heterogeneity Across Counties

We can gain a better understanding of trends in geographic heterogeneity by examining the more abundant county-level data. Empirically *trends* in across-county heterogeneity closely mirror trends in across-municipality heterogeneity.⁴⁵ For example, Table 2 shows that the racial dissimilarity index at the county-level is approximately the MCD-level index shifted down by a constant. Similar parallels between MCD- and county-level trends are evident for: (i) the other proxies in Table 2; (ii) the tax and spending results in Table 1;⁴⁶ (iii) all the proxies and local policy outcomes using the full

⁴⁴There is a spike in education spending heterogeneity across MCDs in the 1960s which stems from Boston's relative reduction in education spending. Seymour Sacks (1972) documents that most urban school districts reduced spending relative to their suburban counterparts during this period. The elimination of the urban-suburban spending gap by the mid-1970s can likely be linked to costs associated with forced busing in Boston (see J. Brian Sheehan, 1984) and changes in the state school aid formula (see Steven Weiss, 1970).

⁴⁵We believe that Tiebout sorting should also apply to counties. In Rhode and Strumpf (2000) we show that counties play an important role in providing local services and that reduced sorting of type proxies occurs even in states where counties have major fiscal responsibilities.

⁴⁶Table 1 also indicates there is substantial within-county heterogeneity in MCD/school district taxes or spending. However, the between-county differences are even larger and typically on the order of two-thirds the overall MCD dispersion.

national set of MCDs over 1970-90 (results omitted); (iv) all the variables from the Boston SMSA (see Rhode and Strumpf, 2000).⁴⁷

A more formal demonstration employs the MCD-level entropy index, which can be decomposed into within- and between-county components. Table 5, which contains results for a wide variety of preference proxies, shows that the decline in overall heterogeneity between MCDs is almost entirely driven by reductions in heterogeneity between counties. The within-county heterogeneity remains roughly constant and small. For example, the within-component contributes less than a fifth to the racial composition index and less than a third to the family income index. All these results suggest that computing heterogeneity trends across counties yields a reasonable approximation to heterogeneity trends between MCDs, the more typical unit for Tiebout communities. We therefore consider the far more abundant data for all US counties over the 1850-1990 period.⁴⁸

The available evidence reveals that the variation in local policy outcomes across counties fell dramatically over time. Perhaps the most prominent category is education. The top panel of Table 6 shows that the dispersion across counties of per capita spending (including all direct education expenditures within the county boundaries) steadily falls by more than 50 percent between 1890 and 1992. The second set of local policy outcomes includes real per capita taxes and revenues. Due to data availability problems, we use four different variables: Tax_1 , taxes collected by counties; Tax_2 , taxes collected by all local governments within the county; Rev_1 , revenues collected by counties; and Rev_2 , revenues collected by all local governments within the county.⁴⁹ The bottom panel of Table 6 shows a sharp drop in dispersion across counties of all these variables over the 1870-1992 period. Although there has been a slight increase in recent decades, the CV for Tax_2 fell by nearly one-half between 1870 and 1992. All of these results are robust to

⁴⁷The Boston results are of interest since the counties of Massachusetts have few fiscal responsibilities, and so there is likely to be a divergence if sub-county Tiebout sorting is the driving factor in residential choice.

⁴⁸Whenever possible the sample includes all counties in existence in a given year and the annual sample sizes are presented in the tables and figures discussed below. Alaska is omitted due to inconsistencies in its county codes.

⁴⁹The main difference between taxes and revenues is inter-governmental grants, which were typically small before 1945.

controlling for outliers, state fixed effects, and returns to scale in government services (see Rhode and Strumpf, 2000).

Heterogeneity across counties of the preference proxies remains flat or falls over the sample. Figure 4 plots the dissimilarity and Gini indices for presidential votes in elections between 1848 and 1988. (To register the importance of third parties, the figure also shows the two-party vote share.) There is a gradual downward trend, especially after 1892.⁵⁰ For example, the Gini trend line has a slope of -0.010 per decade over the entire period, and of -0.014 from 1892 on (the index decreases from 0.38 to 0.24 between 1892 and 1988). The convergence of county election results is not due to the decline in importance of third parties and appears even after accounting for the Democratic party's loss of control of the South (see Rhode and Strumpf, 2000).

The heterogeneity across counties of the black population share declines more noticeably over the sample. Figure 5 shows the dissimilarity and Gini indices as well as the fraction of blacks living in black majority counties. All series remain relatively flat from 1850 to 1890 and then begin falling. This reduction was quite dramatic: while 48.2 percent of blacks lived in black majority counties in 1890, only 9.0 percent did so in 1990. This pattern is consistent with the Great Migration of African-Americans from the South, where they were over-represented (see Carrington et al., 1996). Nonetheless, excluding the South yields indices that are lower than the national series but follow exactly the same declining pattern.⁵¹ Our investigation of county-level net migration patterns over the 1930-80 period confirms these findings.⁵²

⁵⁰The 1860 election was highly unusual because four major parties-- Republicans, Democrats, Southern Democrats, and Constitutional Unionists-- participated in the sectionally divided contest.

⁵¹Our results complement Cutler et al., (1999) who find that black urban segregation increased from 1890 to 1970 and then sharply declined. These contrasting results are likely due to the differences in the scope and level of spatial aggregation of the two analyses. They consider segregation within a city at the census-tract level whereas we are looking at all of the counties in the U.S. Their analysis captures *within city* heterogeneity while our data largely measures differences *across urban and rural areas*. So while the black rural-urban migration tended to reduce heterogeneity at the county-level, it increased heterogeneity within cities if new black migrants tended to live in disproportionately black census tracts.

⁵²In regressions both with and without controls for the South, black net migration rates have a negative and statistically significant relationship with the black population share (results omitted). That is, blacks disproportionately left counties where they were over-represented. These results run counter not only to Tiebout sorting, but also to explanations for declined heterogeneity based on purely random movement.

The data on religious affiliation, displayed in Figure 6, reveal counties have become more alike over the past century. (The convergence is more apparent if one extrapolates back using data on church seating by denominations for the 1850-90 period, see Rhode and Strumpf, 2000.) Figure 6 also shows the declining trends in heterogeneity across counties for home-ownership and education levels. The home-ownership indices each fall almost in half over the 1890-1990 period while the education indices decline slightly between 1940 and 1990.

Figure 7 presents county-level data for the age groups and the foreign born. The heterogeneity indices for the young population share have no strong trend, though dispersion clearly falls in the post-World War Two period. In this same period there is a slight growth in the heterogeneity of the old, but this is swamped by the reduction since 1850.⁵³ For the foreign-born share, there is a slight downward trend in across-county heterogeneity over the whole sample but a noticeable rise between 1960-90 (which is due to the disproportionately rapid growth of Hispanic immigrants in California, Texas, Florida, New Jersey, and New York). Finally, Table 7 presents family and household income inequality/heterogeneity measures for 1949, 1969, 1979, and 1989. Paralleling the results for the Boston SMSA, within-county heterogeneity has stayed roughly constant while the between-county component is relatively small and declined between 1949 and 1979 before increasing slightly over the 1980s. Income groups have not become more sorted.

These results reinforce and extend the MCD-level analysis. There has been measurable convergence across counties in a wide range of local policies and resident public good preferences over the last one hundred fifty years. These findings are in conflict with the prediction of the generalized Tiebout model since mobility costs have fallen over this period.

⁵³While there is no county-level data for the elderly between 1870 and 1920, we were able to compile a state-level time series over the period 1870-1970. The dissimilarity index computed from this data falls continuously, particularly during the period where we have no county data.

VI. Conclusion

This paper evaluates the empirical relevance of Tiebout sorting. Local public goods undoubtedly influence residential choice, but our evidence indicates that other factors have overwhelmed Tiebout sorting in the long run. The augmented Tiebout model predicts greater heterogeneity across communities in both resident preferences and government policies as movement becomes easier. Because of the secular decline in mobility costs, there should be a trend towards greater stratification. However, we find little evidence that the Tiebout mechanism played a dominant role in sorting over the last 150 years. In fact a wide variety of preference and policy variables indicate that communities (as measured by municipalities and counties) have become more alike one another.

These results provide an important challenge for future local public economics research. We need to determine which alternative motives empirically explain long-run residential choices and then incorporate them into our theoretical models. Such revised models will likely have implications that sharply contrast with those from the Tiebout model. This calls into question the literature that adopts a rigid Tiebout framework to explain community composition and mobility choices. To illustrate this, we conclude by briefly discussing two alternative models which are more consistent with the data.⁵⁴

⁵⁴One obvious candidate, the growing federal role in providing public services, cannot be the complete explanation for our data. Centralization limits the benefit of sorting say by providing some minimum bundle of public goods. While this is consistent with a reduction in the rate of sorting, it does not explain the unsorting we observe. Non-Tiebout incentives for residential choice are also needed: centralization increases the relative importance of these factors.

A second policy, zoning, also fails to explain the data. Bruce Hamilton (1975) argues that in the absence of zoning, poor individuals have incentives to move into high-income communities. This gives them access to the local public good at less than average cost, but also defeats the stratification that Tiebout predicts. With zoning, this poor-chasing-the-rich phenomenon could be avoided, say by imposing minimum lot sizes. Robert Nelson (1977) indicates zoning laws were almost entirely absent until the post-War period and they have grown in popularity over time. Hence the Tiebout model should be more appropriate today than in the past, and so the growth of zoning serves as an independent reason for a prediction of growing across community heterogeneity.

A final and more promising candidate is growing local government competition. If some individuals are more desirable than others (say the rich), then in equilibrium communities adopt policies which are relatively favorable to these individuals (Nechyba, 1997 uses such an argument to explain the infrequency of local income taxes). This mechanism offsets Tiebout because it limits policy heterogeneity across communities and thus reduces incentives for individual sorting.

First, suppose that individuals select communities based on employment opportunities as well as local public goods. This would hold if residents receive an exogenous, community-specific net wage (the wage differential could stem from commuting costs or community-specific labor demands which reflect the complementarity of different skill types in production). Such a model is consistent with the reduction in sorting documented in this paper if employment has dispersed or if the relative importance of local public goods has fallen over time (for example, because of growth in the central provision of public services). However, in this model communities cannot be ranked according to public good preferences which is the canonical Tiebout result invoked by a large number of papers in the local public economics literature. This critique holds more generally within the class of models where some other non-Tiebout migration motive has grown in importance (see footnote 16).

Second, suppose we adopt the prevailing view in the empirical literature that Tiebout sorting occurs over a limited geographic area such as a metropolitan area. Under this model, Tiebout sorting within metro areas should increase because mobility costs have declined. This point can be reconciled with our finding of declining national heterogeneity across communities only if the metro areas have become more similar (i.e. between-metro heterogeneity has declined). This should imply growing population diversity in the representative metro area. In fact, the average Herfindahl index of metro-area racial shares fell by over 10 percent between 1930 and 1990. But given moving costs and non-public goods motives for residential choice, this growing local diversity inhibits widespread conformity with the Tiebout model.

References

- Alesina, Alberto; Baqir, Reza and Easterly, William. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics*, November 1999, *114*(4), pp. 1243-84.
- Alesina, Alberto; Baqir, Reza and Hoxby, Caroline. "Political Jurisdictions in Heterogeneous Communities." NBER WP #7859, 2000.
- Alesina, Alberto and La Ferrara, Eliana. "Participation in Heterogeneous Communities." *Quarterly Journal of Economics*, August 2000, *115*(3), pp. 847-904.
- Bayer, Patrick. "Tiebout Sorting and Discrete Choices: A New Explanation for Socioeconomic Differences in the Consumption of School Quality." Mimeo, Yale University, 2000.
- Benabou, Roland. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review*, June 1996, *86*(3), pp. 584-609.
- Bewley, Truman. "A Critique of Tiebout's Theory of Local Expenditures." *Econometrica*, May 1981, *49*(3), pp. 713-40.
- Brueckner, Jan. "A Tiebout/Tax Competition Model." *Journal of Public Economics*, February 2000, *77*(2), pp. 285-306.
- Buchanan, James and Tullock, Gordon. *The Calculus of Consent*. Ann Arbor, MI: University of Michigan Press, 1962.
- Carrington, William; Detragiache, Enrica and Vishwanath, Tara. "Migration with Endogenous Moving Costs." *American Economic Review*, September 1996, *86*(4), pp. 909-30.
- Carroll, Robert and Yinger, John. "Is the Property Tax a Benefit Tax? The Case of Rental Housing." *National Tax Journal*, June 1994, *47*(2), pp. 295-316.
- Cutler, David; Glaeser, Edward and Vigdor, Jacob. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, June 1999, *107*(3), pp. 455-506.
- DiPasquale, Denise and Glaeser, Edward. "Incentives and Social Capital: Are Homeowners Better Citizens?" *Journal of Urban Economics*, March 1999, *45*(2), pp. 354-84.
- Duncan, Otis D. and Duncan, Beverly. "A Methodological Analysis of Segregation Indices." *American Sociological Review*, April 1955, *20*(2), pp. 210-17.

- Dowding, Keith; John, Peter and Biggs, Stephen. "Tiebout: A Survey of the Empirical Literature." *Urban Studies*, May 1994, 31(4-5), pp. 767-97.
- Ely, Richard. *Taxation in American States and Cities*. New York: Thomas Crowell, 1888.
- Epple, Dennis; Figlio, David and Romano, Richard Romano. "Competition Between Private and Public Schools: Testing Stratification and Pricing Predictions." *Journal of Public Economics*, (forthcoming).
- Epple, Dennis and Romer, Thomas. "Mobility and Redistribution." *Journal of Political Economy*, August 1991, 99(4), pp. 828-58.
- Epple, Dennis; Romer, Thomas and Sieg, Holger. "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis." *Econometrica*, August 2001, 69(6), pp. 1437-66.
- Epple, Dennis and Sieg, Holger. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy*, August 1999, 107(4), pp. 645-81.
- Fernandez, Raquel and Rogerson, Richard. "Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education." *International Economic Review*, February 1997, 38(1), pp. 23-42.
- , "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform." *American Economic Review*, September 1998, 88(4), pp. 813-33.
- Fisher, Ronald and Wassmer, Robert. "Economic Influences on the Structure of Local Government in U.S. Metropolitan Areas." *Journal of Urban Economics*, May 1998, 43(3), pp. 444-71.
- Galenson, David. Settlement and Growth of the American Colonies. In S. Engerman and R. Gallman, eds. *Cambridge Economic History of the United States*. Vo. I. Cambridge, Eng.: Cambridge University Press, 1996.
- Gastwirth, Joseph. "The Estimation of the Lorenz Curve and the Gini Index." *Review of Economics and Statistics*, August 1972, 54(3), pp. 306-16.
- General Social Survey. *GSS 1972-1996 Cumulative Datafile*. Survey Documentation & Analysis Archive. <http://csa.berkeley.edu:7502/>. 1999.
- Glomm, Gerhard and Lagunoff, Roger. "A Dynamic Tiebout Theory of Voluntary vs. Involuntary Provision of Public Goods." *Review of Economic Studies*, July 1999, 66(3), pp. 659-77.

- Goldin, Claudia and Katz, Lawrence. "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910 to 1940." *Journal of Interdisciplinary History*, Spring 1999, 29(4), pp. 683-723.
- Hamilton, Bruce. "Zoning and Property Taxation in a System of Local Government." *Urban Studies*, June 1975, 12(2), pp. 205-11.
- Hercowitz, Zvi and Pines, David. "Migration Between Home Country and Diaspora: An Economic Analysis." *Journal of Public Economics*, July 1997, 65(1), pp. 45-59.
- Hoxby, Caroline. "The Productivity of Schools and Other Local Public Goods Producers." *Journal of Public Economics*, October 1999, 74(1), pp. 1-30.
- , "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, December 2000, 90(5), pp. 1209-38.
- Iannaccone, Laurence. "Introduction to the Economics of Religion." *Journal of Economic Literature*, September 1998, 36(3), pp. 1465-95.
- Kennan, John and Walker, James. "Geographical Wage Differentials, Welfare Benefits and Migration." Mimeo, University of Wisconsin, Madison, 2000.
- Kollman, Ken; Miller, John and Page, Scott. "Political Institutions and Sorting in a Tiebout Model." *American Economic Review*, December 1997, 87(5), pp. 977-92.
- Kremer, Michael. "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics*, February 1997, 112(1), pp. 115-39.
- Leege, David and Kellstedt, Lyman, ed. *Rediscovering the Religious Factor in American Politics*. Armonk, NY: M.E. Sharpe, 1993.
- Martinez-Vazquez, Jorge; Rider, Mark and Walker, Mary. "Race and the Structure of School Districts in the United States." *Journal of Urban Economics*, March 1997, 41(2), pp. 281-300.
- Massey, Douglas and Denton, Nancy. "The Dimensions of Residential Segregation." *Social Forces*, December 1988, 67(2), pp. 281-315.
- Maxim, Hiram P. "Some Data on the Cost of Operating Automobiles for Commercial Purposes." *Scientific American* Supplement No. 1479, 7 May 1904, pp. 23694-95
- Nechyba, Thomas. "Local Property and State Income Taxes: The Role of Interjurisdictional Competition and Collusion." *Journal of Political Economy*, April 1997, 105(2), pp. 351-85.

- . “School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers.” *Journal of Public Economic Theory*, January 1999, *1(1)*, pp. 5-50.
- . “Mobility, Targeting, and Private-School Vouchers.” *American Economic Review*, March 2000, *90(1)*, pp. 130-46.
- Nelson, Robert. *Zoning and Property Rights*. Cambridge, MA: MIT Press, Cambridge, MA, 1977.
- Newport, Frank. “The Religious Switcher in the United States.” *American Sociological Review*, August 1979, *44(4)*, pp. 528-52.
- Noll, Mark, ed. *Religion and American Politics: From the Colonial Period to the 1980s*. New York: Oxford University Press, 1990.
- Perroni, Carlo and Scharf, Kimberley. “Tiebout with Politics: Capital Tax Competition and Constitutional Choices.” *Review of Economic Studies*, January 2001, *68(1)*, pp. 133-54.
- Piketty, Thomas. “Social Mobility and Redistributive Politics.” *Quarterly Journal of Economics*, August 1995, *110(3)*, pp. 551-84.
- Reardon, Sean. “Stata module to compute multiple-group diversity and segregation indices.” ideas.uqam.ca/ideas/data/Softwares/bocbocodeS375001.html, 1998.
- Rhode, Paul and Strumpf, Koleman. “A Historical Test of the Tiebout Hypothesis: Local Heterogeneity from 1850 to 1990.” NBER WP #7946, 2000.
- Romer, Thomas and Rosenthal, Howard. “Bureaucrats Versus Voters: On the Political Economy of Resource Allocation by Direct Democracy.” *Quarterly Journal of Economics*, November 1977, *93(4)*, pp. 563-87.
- Sacks, Seymour. *City Schools/Suburban Schools: A History of Fiscal Conflict*. Syracuse: Syracuse University Press, 1972.
- Sheehan, J. Brian. *The Boston School Integration Dispute: Social Change and Legal Maneuvers*. New York: Columbia University Press, 1984.
- Shorrocks, Anthony F. “The Class of Additively Decomposable Inequality Measures.” *Econometrica*, April 1980, *48(3)*, pp. 613-26.
- Tiebout, Charles. “A Pure Theory of Local Expenditures.” *Journal of Political Economy*, October 1956, *64(5)*, pp. 416-24.

Weiss, Steven. *Existing Disparities in Public School Finance and Proposals for Reform*. Federal Reserve Bank of Boston, Research Report 46. 1970.

Wildasin, David and Wilson, John. "Imperfect Mobility and Local Government Behaviour in an Overlapping-Generations Model." *Journal of Public Economics*, May 1996, 60(2), pp. 177-98.

Wooders, Myrna. "Multijurisdictional Economies, the Tiebout Hypothesis, and Sorting." *Proceedings of the National Academy of Sciences, USA*. September 1999, 96(19), pp. 10585-87.

Appendix A

This appendix contains the example of sorting with single-peaked preferences, and proofs of cases where sorting is associated with increased policy variation.

Example A: Sorting with Single-Peaked Preferences

A preliminary: define the initial level of variable X as X^0 and the level after one agent of type s moves as X^1 .

Example A: Suppose individual preferences satisfy **A1** and communities set their policies according to (2). If an agent of type s moves from community d to c , then

- (i) $|G_c^{*1} - G_s| \leq |G_c^{*0} - G_s|$ and $|G_d^{*1} - G_s| \geq |G_d^{*0} - G_s|$;
- (ii) $U_s(G_c^{*1}) \geq U_s(G_c^{*0})$ and $U_s(G_d^{*1}) \leq U_s(G_d^{*0})$;
- (iii) $U_t(G_c^{*1}) \geq U_t(G_c^{*0}) \forall t$ such that $\text{sign}(G_t - G_c^{*1}) = \text{sign}(G_s - G_c^{*1})$;
- (iv) $U_t(G_c^{*1}) \leq U_t(G_c^{*0}) \forall t$ such that $\text{sign}(G_t - G_c^{*0}) = -\text{sign}(G_s - G_c^{*0})$;
- (v) $U_t(G_d^{*1}) \leq U_t(G_d^{*0}) \forall t$ such that $\text{sign}(G_t - G_d^{*0}) = \text{sign}(G_s - G_d^{*0})$;
- (vi) $U_t(G_d^{*1}) \geq U_t(G_d^{*0}) \forall t$ such that $\text{sign}(G_t - G_d^{*1}) = -\text{sign}(G_s - G_d^{*1})$.

To explain this example, we focus on the case for the receiving community because the case of the sending community is analogous. If public good provision is set by (2), the first order condition $\sum_t N_{tc} U_t'(G_c^*) = 0$ must be satisfied, where N_{tc} is the number of type t in community c . If one more person of type s moves in, holding the other N_{tc} constant, the weight on the $U_s'(G_c^*)$ increases. Unless $U_s'(G_c^{*0}) = 0$, the community must move G_c^* closer to G_s to satisfy the new first order condition. This increases the utility of type s and all types on the same side of G_c^{*1} as s and reduces the utility of all types on the other side of G_c^{*0} . The inequalities are strict unless $U_s'(G_c^{*0}) = 0$.

Proof of the Observation.

Examples B and C provide conditions under which *sorting is associated with increased variation of policy outcomes across communities*.

Example B: If preferences satisfy **A1** and there are two communities and two types of individuals, then migration obeying (3) and (4) increases the differences between the communities' policies. Call the two communities c and d and the two types r and s , where $G_r < G_s$. Let N_i be the total population of type i and let N_{ic} be the number in community c . Given N_{rc} and N_{sc} , G_c^* will be set where $N_{rc}U_r'(G_c^*) = -N_{sc}U_s'(G_c^*)$. Note that $U_r'(G_c^*) < 0 < U_s'(G_c^*)$ and that $d[U_s'(G_c^*)/(-U_r'(G_c^*))]/dG_c^* < 0$. By the implicit function theorem, we can solve for $G_c^* = H[N_{rc}/N_{sc}]$ where $H' < 0$, $H[0] = G_s$ and $H[\infty] = G_r$. By a similar argument, $G_d^* = H[(N_r - N_{rc})/(N_s - N_{sc})]$. If $N_{rc}/N_{sc} > N_1/N_2$, then $G_r \leq G_c^* < G_d^* \leq G_s$. Community c will be the preferred community for type r and community d for type s . Migration obeying (3) and (4), which increases in N_{rc} and N_{sd} , causes greater segregation and widens the differences between the communities' policies: $d|G_d^* - G_c^*|/dN_{rc} > 0$ and $d|G_d^* - G_c^*|/dN_{sd} > 0$.

Example C: *If preferences are quadratic **A2**, then any move obeying (3) and (4) increases the aggregate population-weighted variance of policies.* Under **A2** and (2), the policy in a given community is the mean of members' ideal policies and the aggregate population-weighted mean, χ , is independent of the distribution of types across communities. An agent i of type s will move from community d (initially with N_d^0 members) to community c (with N_c^0 members) if $(G_c^{*0} - G_s)^2 + m_i < (G_d^{*0} - G_s)^2$. Such a move will change the policy in community d from G_d^{*0} to $G_d^{*1} = (G_d^{*0} N_d^0 - G_s) / (N_d^0 - 1)$ and that in c from G_c^{*0} to $G_c^{*1} = (G_c^{*0} N_c^0 + G_s) / (N_c^0 + 1)$. Such a move will not affect the population-weighted mean of policies in c and d , i.e. $N_c^0 G_c^{*0} + N_d G_d^{*0} = (N_c^0 + 1) G_c^{*1} + (N_d^0 - 1) G_d^{*1}$. Nor will it change the aggregate mean or policies in communities other than c and d . But such a move does raise the population-weighted variances of policies:

$$\begin{aligned} & \Sigma_e [N_c^1 (G_c^{*1} - \chi)^2 - N_c^0 (G_c^{*0} - \chi)^2] / N \\ & = [(N_c^0 + 1) (G_c^{*1})^2 + (N_d^0 - 1) (G_d^{*1})^2 - N_c^0 (G_c^{*0})^2 - N_d^0 (G_d^{*0})^2] / N \\ & = [(N_d^0 / (N_d^0 - 1)) (G_d^{*0} - G_s)^2 - (N_c^0 / (N_c^0 + 1)) (G_c^{*0} - G_s)^2] / N > 0 \end{aligned}$$

where the inequality follows from $(N_c^0 / (N_c^0 + 1)) < 1 < (N_d^0 / (N_d^0 - 1))$ and $(G_c^{*0} - G_s)^2 < (G_d^{*0} - G_s)^2$.

**Table 1: Dispersion of Local Policies
Across-MCD/School District and Across-County Indices**

(a) Municipal Per Capita Taxes

Year	MCD (Municipal)-Level			Aggregated to County-Level		
	N	CV	DG	N	CV	DG
1880	634	1.738	0.537	239	1.353	0.412
1890	968	1.321	0.349	268	0.793	0.307
1972	3175	1.100	0.347	304	0.762	0.259
1982	3196	0.864	0.311	305	0.703	0.224
1992	3251	0.844	0.295	305	0.622	0.207

(b) School District Per Capita Taxes

Year	School District-Level			Aggregated to County-Level		
	N	CV	DG	N	CV	DG
1880	634	2.093	0.796	239	0.759	0.313
1890	968	0.974	0.289	268	0.730	0.275
1972	1352	0.573	0.216	275	0.504	0.202
1982	1193	0.675	0.246	276	0.568	0.207
1992	1221	0.696	0.264	276	0.571	0.218

(c) Total Spending (Current Operations) Per Capita

Year	MCD (Municipal)-Level			Aggregated to County-Level		
	N	CV	DG	N	CV	DG
1972	3171	1.003	0.357	304	0.654	0.244
1982	3177	0.847	0.304	305	0.503	0.192
1992	3271	0.778	0.285	305	0.463	0.184

(d) Protection Spending (Current Operations) Per Capita

Year	MCD (Municipal)-Level			Aggregated to County-Level		
	N	CV	DG	N	CV	DG
1972	1906	0.721	0.261	303	0.512	0.202
1982	2257	0.603	0.229	305	0.470	0.188
1992	2336	0.607	0.225	305	0.435	0.179

(e) Education Spending (Current Operations) Per Student Enrollment

Year	School District-Level			Aggregated to County-Level		
	N	CV	DG	N	CV	DG
1972	1346	0.253	0.091	275	0.232	0.085
1982	1191	0.241	0.090	276	0.237	0.084
1992	1225	0.217	0.080	276	0.220	0.077

Notes: The coefficient of variation, CV, is defined in equation (11) of Section III; the reallocation index, DG, is defined in equation (12). These values are based on a 1-in-10 sample of all municipalities, which is further described in the online Data Appendix. County-level taxes/spending are the sum of all municipal taxes/spending plus any county taxes/spending

**Table 2: Dispersion of Preference Proxies
Across-MCD and Across-County Dissimilarity Indices**

Year	Racial Composition (Black)			Nativity (Foreign-Born)		
	N	MCD-Level	Aggregated to County-Level	N	MCD-Level	Aggregated to County-Level
1870	2071	0.724	0.697	2071	0.559	0.511
1930	5071	0.650	0.615	5070	0.503	0.491
1960	3798	0.510	0.477			
1970	3020	0.528	0.499	3039	0.476	0.447
1980	3453	0.564	0.520	3415	0.470	0.458
1990	3456	0.572	0.525	3456	0.482	0.465

Year	Young Population			Old Population		
	N	MCD-Level	Aggregated to County-Level	N	MCD-Level	Aggregated to County-Level
1930	5071	0.106	0.087	5071	0.182	0.161
1960	3798	0.086	0.064	3798	0.149	0.121
1970	3275	0.073	0.045	3275	0.169	0.136
1980	3415	0.069	0.068	3415	0.175	0.145
1990	3456	0.073	0.054	3456	0.176	0.147

Notes: The dissimilarity index is defined in equation (13) of Section III. The sample size (N) refers to the number of MCDs. Empty cells indicate missing data. These results are based on a 1-in-10 sample. Full details of the data as well as precise definition of the categories are in the online Data Appendix.

Table 3: Household Income Heterogeneity Within- and Between-MCDs in the Boston SMSA

Year	N	Within-Indices			Within-/Between- Decomposition			
		CV	GL	GU	I ₁ -W	I ₁ -B	I ₂ -W	I ₂ -B
1949	74	0.805 (0.090)	0.404 (0.043)	0.414 (0.044)	0.290	0.041	0.386	0.033
1979	92	0.788 (0.106)	0.394 (0.039)	0.398 (0.039)	0.261	0.041	0.316	0.039
1989	92	0.809 (0.105)	0.407 (0.040)	0.413 (0.039)	0.279	0.040	0.349	0.038

Notes: The coefficient of variation, CV, is defined in equation (11) of Section III. The Gini indices, GL and GU, are defined in equations (19) and (20), respectively. The Theil indices, I₁ and I₂ are defined in equations (17) and (18). The “W” is within-MCD, and “B” is between-MCD. The within-measures are population weighted averages of each of the MCD indices. Standard deviations are reported in parentheses below the within-measures.

There is no electronic data available prior to 1979 for households. The results for households in 1979 and 1989 are similar if MCDs which are missing data in 1949 are omitted.

Results for families are not reported since there is no family-level data in the 1949 hardcopies (see the online Data Appendix for details). The family-level indices are quite similar to the household-level indices for 1979 and 1989.

To compute the indices, we need to know for each MCD the proportion of people in each income category and the mean income in each group. Because the latter is unavailable, the mid-point of each income interval was used as the mean. For the top-coded income group, a mean of 1.5 times the lower bound was used (several other values were considered and the results do not appear to be sensitive to this choice). For the Gini measures the upper and lower bound of each income interval is also needed. For the upper bound of the top-coded group, 20 times the lower bound was used (again the results are robust to using other values for the top-coded group).

**Table 4: Across-MCD Current Operations Spending in the Boston SMSA
Coefficients of Variations (CV)**

Year	N	G _{Total}	G _{Protection}	G _{Education}
1906	92	0.359	0.490	0.197
1913	92	0.297	0.402	0.218
1923	92	0.267	0.351	0.203
1932	92	0.281	0.346	0.181
1942	92	0.220	0.297	0.160
1955	92	0.236	0.350	0.190
1962	59	0.189	0.311	0.311
1967	59	0.216	0.281	0.356
1972	92	0.293	0.399	0.267
1977	92	0.242	0.365	0.177
1982	92	0.233	0.276	0.206
1987	92	0.243	0.334	0.187
1992	92	0.267	0.340	0.194

Notes: See the online Data Appendix for a list of sources and definitions of these series. The coefficient of variation, CV, is defined in equation (11). All values are population-weighted. In 1962 and 1967 there are no values reported for the 33 municipalities which have populations less than 10,000. The CV's in the remaining years do not change significantly when these 33 municipalities are omitted (because the measure is population-weighted and these are all small communities).

Table 5: Within-County/Between-County Decomposition of the MCD-Level Entropy Index

Year	Racial Composition (Black)				Nativity (Foreign-Born)			
	N	MCD Total	Within-County	Between-County	N	MCD Total	Within-County	Between-County
1870	2071	0.453	0.041	0.411	2071	0.255	0.033	0.222
1930	5071	0.374	0.051	0.323	5070	0.207	0.011	0.196
1960	3798	0.252	0.041	0.211				
1970	3020	0.246	0.044	0.202	3039	0.160	0.016	0.144
1980	3453	0.274	0.044	0.230	3415	0.166	0.010	0.156
1990	3456	0.293	0.055	0.237	3456	0.187	0.013	0.174

Year	Young Population				Old Population			
	N	MCD Total	Within-County	Between-County	N	MCD Total	Within-County	Between-County
1930	5071	0.012	0.004	0.008	5071	0.026	0.008	0.018
1960	3798	0.008	0.002	0.005	3798	0.026	0.008	0.018
1970	3275	0.007	0.003	0.004	3275	0.035	0.011	0.024
1980	3415	0.007	0.003	0.004	3415	0.034	0.010	0.024
1990	3456	0.007	0.003	0.004	3456	0.034	0.012	0.023

Year	Education				Home-Owner Occupation			
	N	MCD Total	Within-County	Between-County	N	MCD Total	Within-County	Between-County
1970	3275	0.045	0.014	0.031				
1980	3415	0.042	0.011	0.032	3422	0.102	0.023	0.080
1990	3455	0.047	0.011	0.033	3456	0.098	0.029	0.070

Year	Family Income: Theil ₁ -Between				Family Income: Theil ₂ -Between			
	N	MCD Total	Within-County	Between-County	N	MCD Total	Within-County	Between-County
1969	3275	0.029	0.007	0.022	3275	0.028	0.007	0.021
1979	3415	0.024	0.006	0.018	3415	0.022	0.006	0.017
1989	3456	0.037	0.011	0.025	3456	0.034	0.010	0.024

Notes: The overall-, within-, and between-entropy terms are defined in equations (15) and (16) of Section III. Empty cells indicate missing data. These results are based on a 1-in-10 sample. Full details of the data as well as precise definition of the categories are in the online Data Appendix.

Table 6: Dispersion of Local Policies Across Counties

Per Capita Education Spending

Year	N	CV	DG
1890	2623	0.663	0.212
1932	3084	0.487	0.183
1957	3091	0.335	0.124
1962	3103	0.302	0.116
1967	3102	0.285	0.104
1972	3106	0.297	0.109
1977	3110	0.270	0.103
1982	3110	0.251	0.087
1987	3110	0.247	0.084
1992	3112	0.249	0.089

Per Capita Taxes and Revenues

Year	N	CV				DG			
		Tax ₁	Tax ₂	Rev ₁	Rev ₂	Tax ₁	Tax ₂	Rev ₁	Rev ₂
1870	2098	1.179	0.933			0.349	0.338		
1880	2302	0.878	0.889			0.282	0.327		
1890	1308			1.015				0.364	
1902	2679			0.745				0.297	
1913	2902			0.868				0.338	
1922	3024	0.695				0.255			
1932	3083	0.677	0.473	0.640	0.463	0.248	0.191	0.234	0.187
1942	2497	0.689		0.755		0.261		0.258	
1957	3087				0.373				0.150
1962	3093		0.464		0.346		0.192		0.139
1967	3095		0.467		0.385		0.187		0.143
1972	3097		0.485		0.419		0.197		0.159
1977	3104		0.524		0.421		0.206		0.154
1982	3103		0.503		0.363		0.183		0.134
1987	3104		0.522		0.372		0.187		0.137
1992	3104		0.497		0.350		0.182		0.129

Notes: The coefficient of variation, CV, is defined in equation (11) of Section III; the reallocation index, DG, is defined in equation (12).

For the top panel, the G variables involve education spending. For the bottom panel, the G's are taxes or revenues (G₁ = just county government G; G₂ = county + sub-county government G). Empty cells are due to missing data. See the online Data Appendix for further details about the data.

Table 7: Income Heterogeneity Within- and Between-Counties

	Year	N	Within-Indices			Within-/Between- Decomposition			
			CV	GL	GU	I ₁ -W	I ₁ -B	I ₂ -W	I ₂ -B
Families	1949	311	0.808 (0.132)	0.390 (0.046)	0.396 (0.047)	0.261	0.031	0.327	0.035
	1969	311	0.815 (0.098)	0.378 (0.034)	0.386 (0.033)	0.254	0.022	0.291	0.021
	1979	311	0.776 (0.075)	0.376 (0.031)	0.379 (0.031)	0.244	0.018	0.284	0.017
	1989	311	0.833 (0.086)	0.395 (0.036)	0.397 (0.036)	0.273	0.025	0.314	0.024
Households	1949	311	0.896 (0.124)	0.433 (0.039)	0.439 (0.041)	0.322	0.028	0.413	0.031
	1979	311	0.857 (0.074)	0.414 (0.027)	0.417 (0.027)	0.294	0.018	0.346	0.017
	1989	311	0.909 (0.082)	0.428 (0.030)	0.431 (0.030)	0.319	0.025	0.372	0.023

Notes: The coefficient of variation, CV, is defined in equation (11) of Section III. The Gini indices, GL and GU, are defined in equations (19) and (20), respectively. The Theil indices, I₁ and I₂ are defined in equations (17) and (18). The “W” is within-counties, and “B” is between-counties. The within-measures are population-weighted averages of each of the county indices. Standard deviations are reported in parentheses below the within-measures.

As described in the online Data Appendix, these values are based on a random, 1-in-10 sample. Household data for 1969 is not available in electronic form.

See Table 3 for additional comments.

Figure 1: Real Transportation Costs

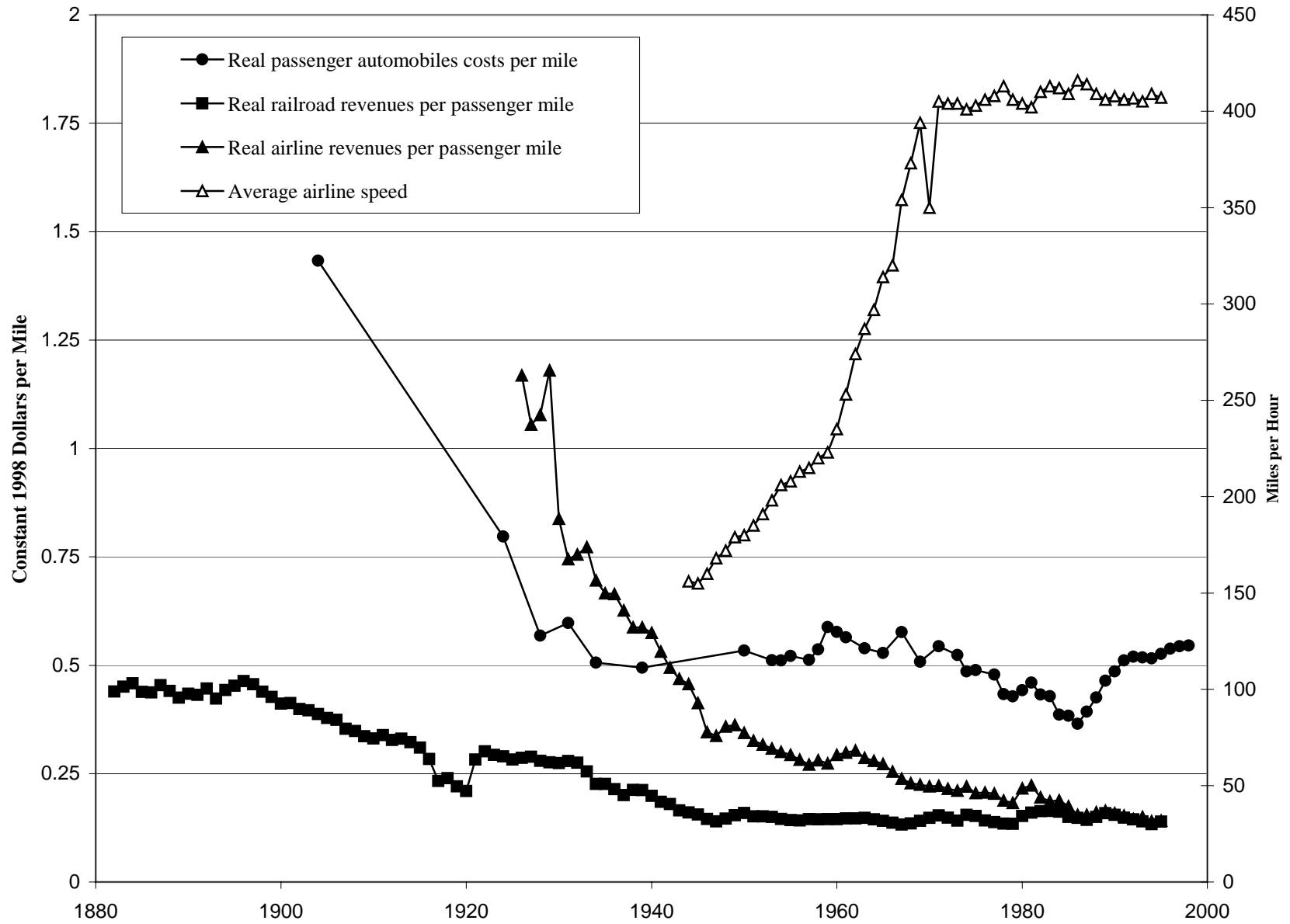
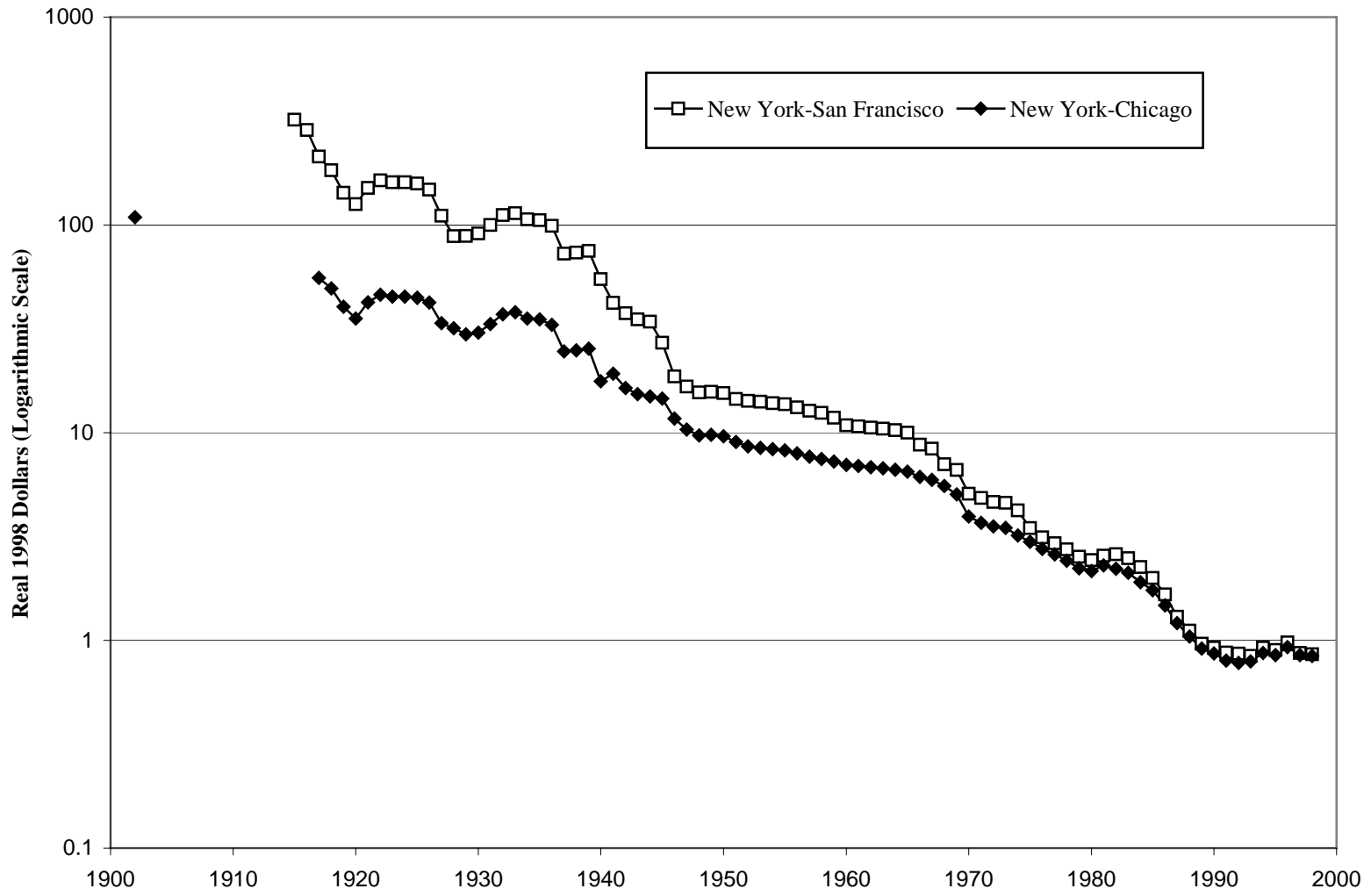
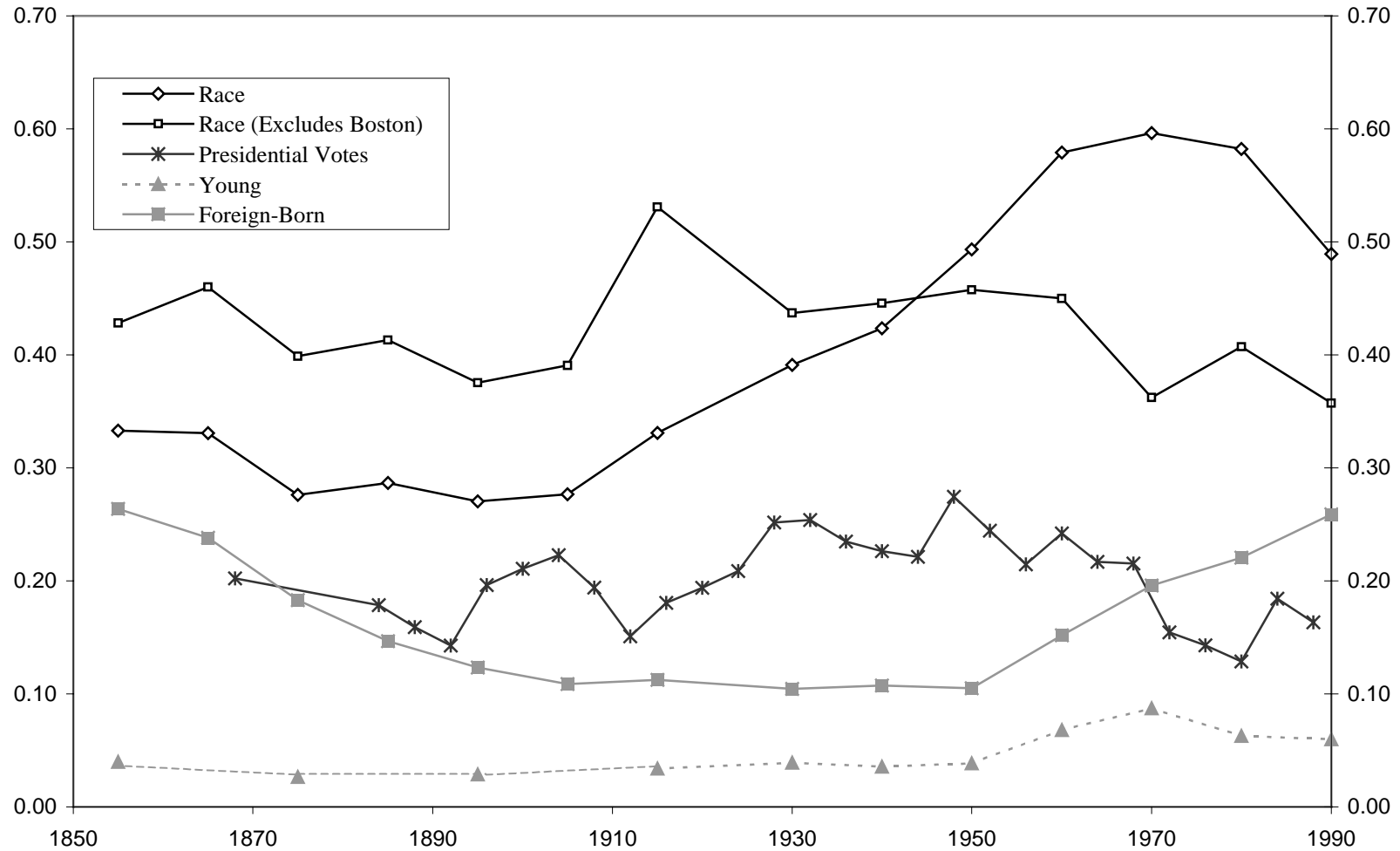


Figure 2: Real Cost of a Three-Minute Phone Call

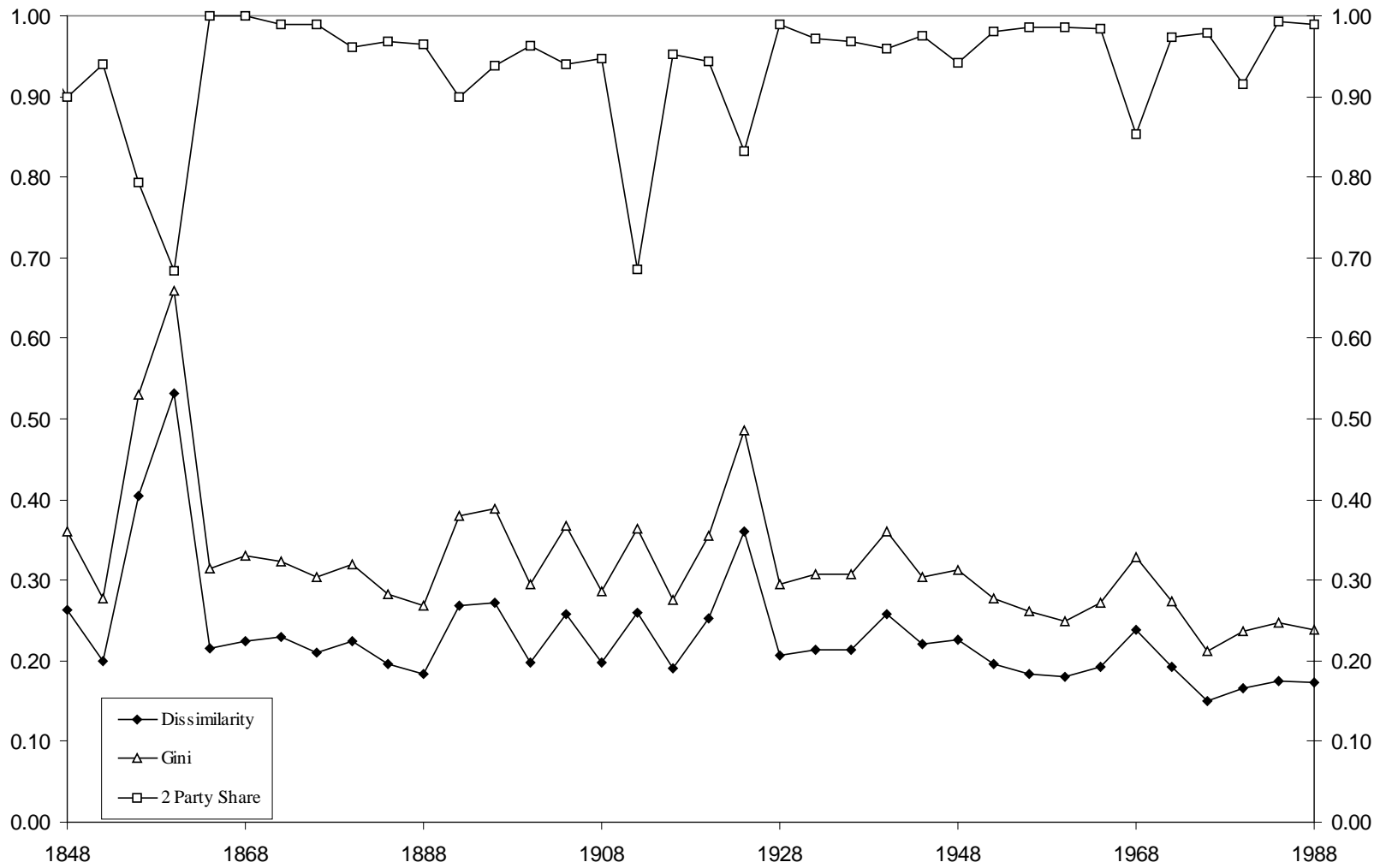


**Figure 3: Demographic Heterogeneity Across Minor Civil Divisions
in the Boston SMSA (Dissimilarity Index)**



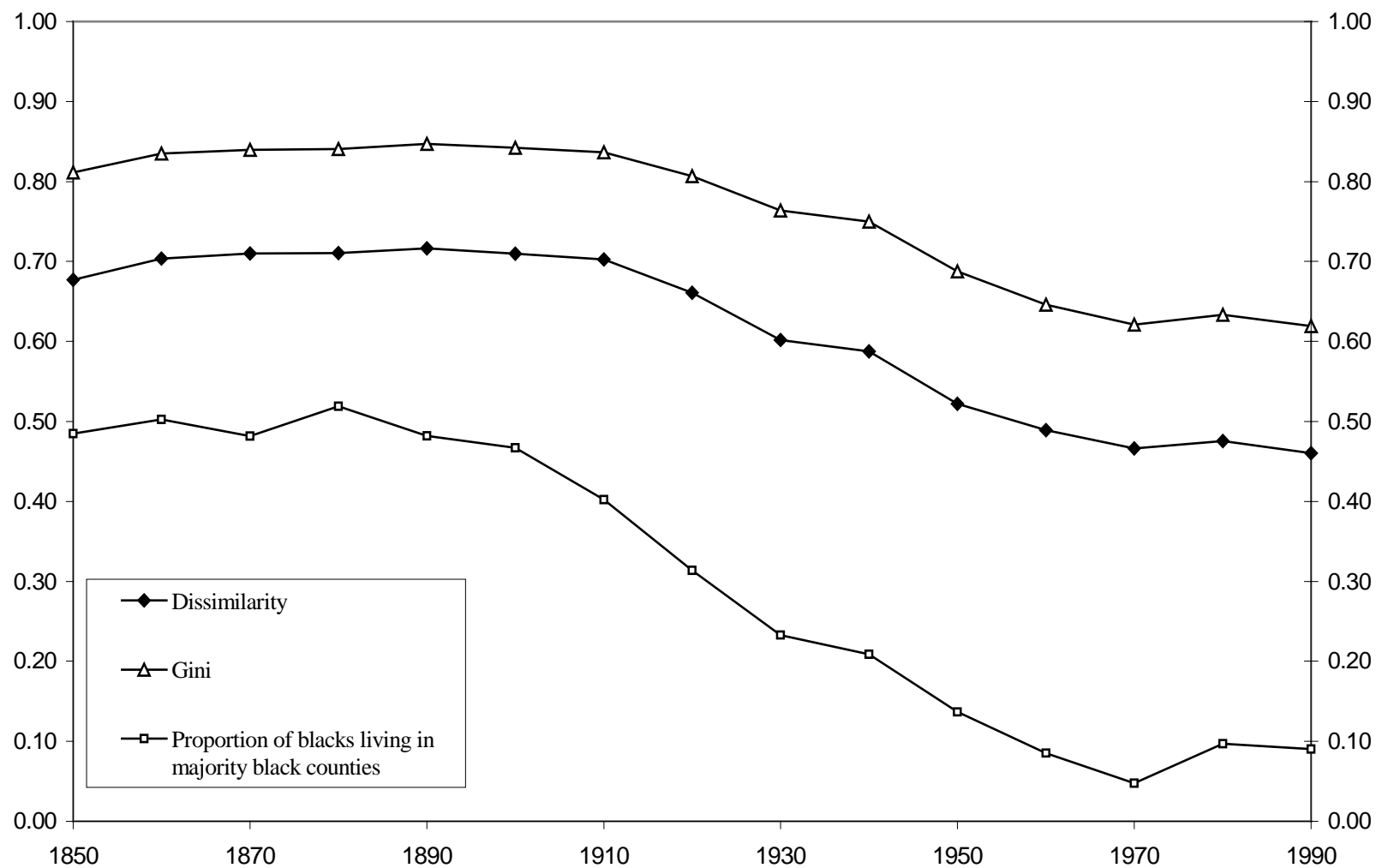
Year (Sample Size): 1855 (79), 1865 (79), 1875 (88), 1885 (90), 1895 (91; 90 for Foreign-Born), 1905 (92), 1915 (92), 1930 (92), 1940 (92), 1950 (75), 1960 (83; 92 for Race), 1970 (91; 92 for Race), 1980 (92), 1990 (92). Presidential vote share is every four years (1868-1988) and includes 92 MCDs.

Figure 4: Heterogeneity of Presidential Vote Shares Across US Counties



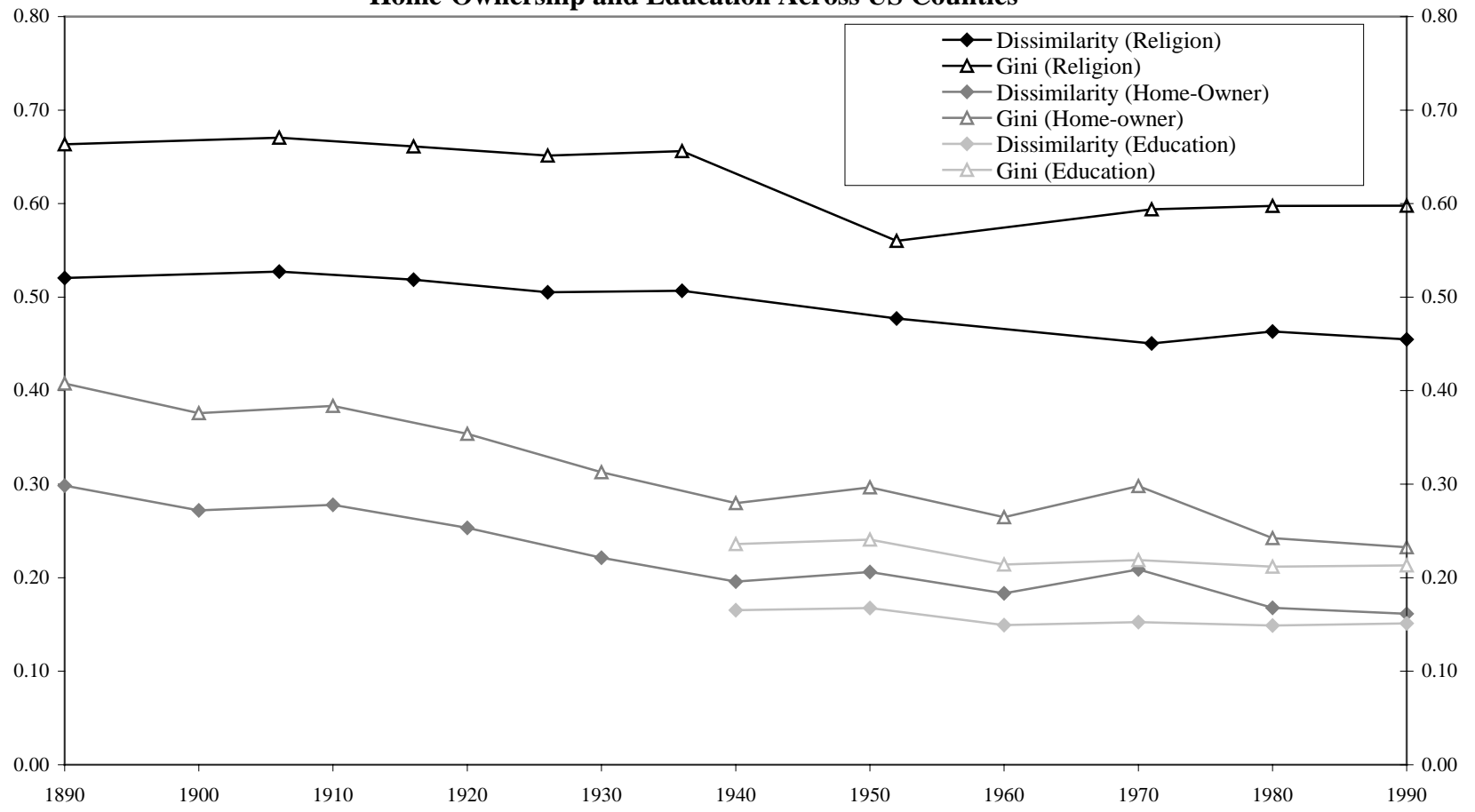
Year (Sample Size): 1852 (1551), 1860 (1864), 1872 (2177), 1880 (2315),
 1892 (2667), 1900 (2730), 1912 (2970), 1920 (3031), 1932 (3091), 1940 (3067),
 1952 (3097), 1960 (3101), 1972 (3105), 1980 (3111), 1988 (3113)

Figure 5: Heterogeneity in Black Population Share Across US Counties



Year (Sample Size): 1850 (1596), 1860 (2030), 1870 (2185), 1880 (2400),
 1890 (2743), 1900 (2777), 1910 (2950), 1920 (3064), 1930 (3100), 1940 (3097),
 1950 (3100), 1960 (3108), 1970 (3111), 1980 (3114), 1990 (3116)

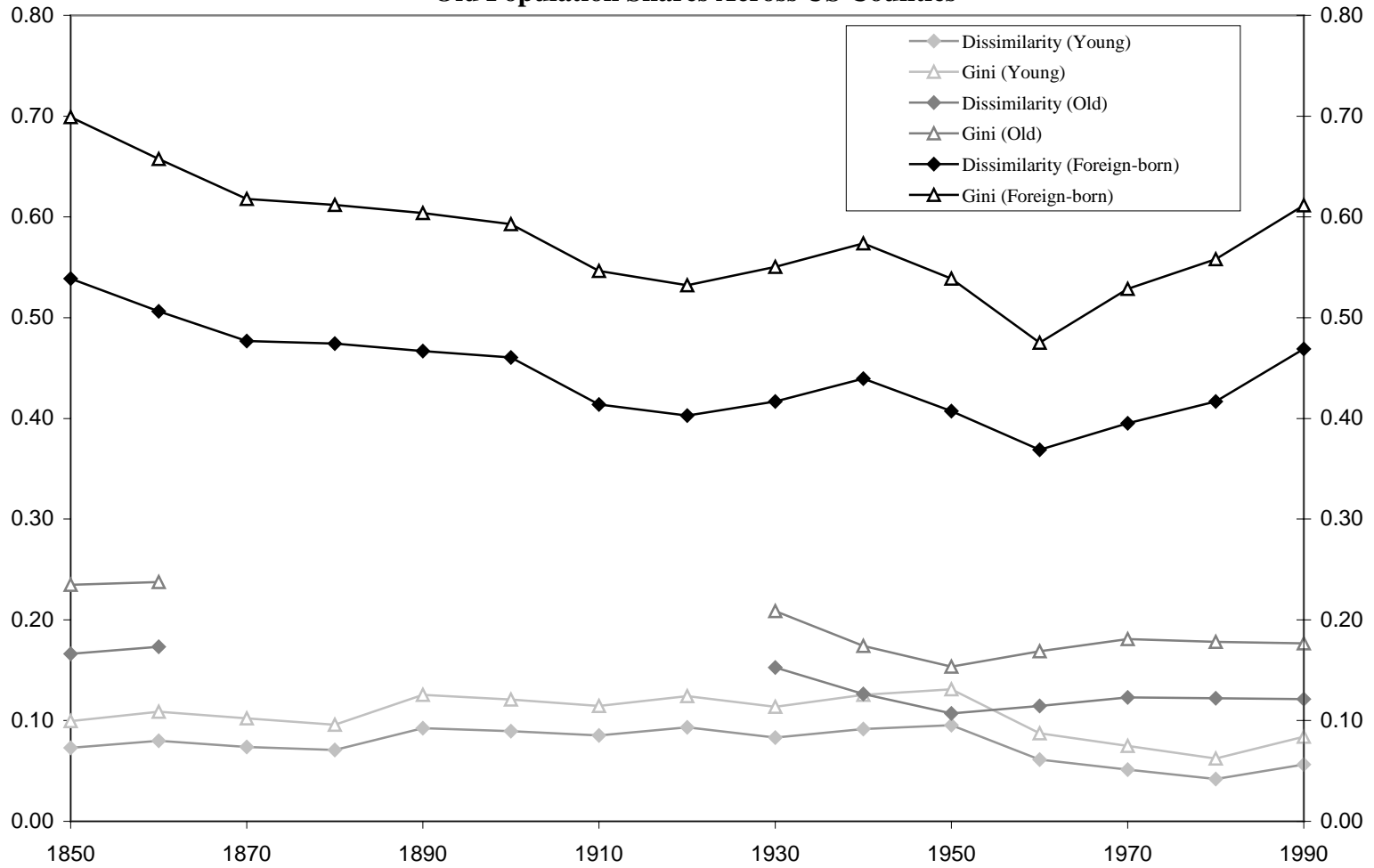
**Figure 6: Heterogeneity of Religion Denominational Shares,
Home-Ownership and Education Across US Counties**



Year (Sample Size with R denoting religious data): 1890 (2753, R: 2677), 1900 (2825), 1906 (R: 2767), 1910 (2949), 1916 (R: 2948), 1920 (3064), 1926 (R: 3064), 1930 (3100), 1936 (R: 3096), 1940 (3097), 1950 (3099), 1952 (R: 3072), 1960 (3103), 1970 (3109), 1971 (R: 3077), 1980 (2753, R: 3068), 1990 (3110, R: 3080)

Note: Unaffiliated and affiliates of minor denominations are excluded.

Figure 7: Heterogeneity of Foreign-Born, Young, and Old Population Shares Across US Counties



Year (Sample Size): 1850 (1607), 1860 (2055), 1870 (2230), 1880 (2421), 1890 (2780), 1900 (2832), 1910 (2955), 1920 (3071), 1930 (3102), 1940 (3099), 1950 (3102), 1960 (3114), 1970 (3112), 1980 (3115), 1990 (3117)

Note: there is no county-level data for the old population share between 1870 and 1920. For foreignborn, 1850, 1860: includes slave population; 1910-1930: white only; 1950: >= 21 years only.

