

NBER WORKING PAPER SERIES

A HISTORICAL TEST OF THE TIEBOUT HYPOTHESIS:
LOCAL HETEROGENEITY FROM 1850 TO 1990Paul W. Rhode
Koleman S. StrumpfWorking Paper 7946
<http://www.nber.org/papers/w7946>NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2000

We thank David Blau, Judith Blau, Dennis Epple, Claudia Goldin, Thomas Mroz, Wilbert van der Klauuw, Jacob Vigdor and participants at the 2000 AEA meetings, 1999 NBER Development of the American Economy Summer Institute and UNC-Duke seminars for comments and suggestions. The views expressed in this paper are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 2000 by Paul W. Rhode and Koleman S. Strumpf. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

A Historical Test of the Tiebout Hypothesis: Local Heterogeneity from 1850 to 1990
Paul W. Rhode and Koleman S. Strumpf
NBER Working Paper No. 7946
October 2000
JEL No. D7, H7, N3, R5

ABSTRACT

The Tiebout hypothesis, which states that individuals will costlessly sort themselves across local communities according to their public good preferences, is the workhorse of the local public finance literature. This paper develops a test of the Tiebout hypothesis using historical variation in mobility costs. Our extension of the Tiebout model to incorporate such costs yields the following comparative statics: as mobility costs fall, the heterogeneity across communities of individual public good preferences and, under some standard assumptions, of public good provision must (weakly) increase. Given mobility costs have fallen over time, a natural test of the Tiebout hypothesis is to take these predictions to the data — here all US counties over the 1850-1990 period. Contrary to the predictions, we find decreasing heterogeneity between counties in policy outcomes (local education spending and total taxes or revenues) and in a wide variety of proxies for public good preferences (age groups, education levels, election outcomes, home ownership, income, race, and religious affiliation). Using the Boston SMSA as a case study, we show that the heterogeneity trends are similar at the municipal and county levels. These results suggest that forces working in opposition to Tiebout sorting have dominated individual location decisions over the past century.

Paul W. Rhode
Department of Economics
UNC, Chapel Hill
CB#3305
Chapel Hill, NC 27599
and NBER
prhode@email.unc.edu

Koleman S. Strumpf
Department of Economics
UNC, Chapel Hill
Chapel Hill, NC 27599

I. Introduction

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 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.¹ One of its central assumptions is that individuals are perfectly mobile. While this assumption simplifies the theoretical analysis, real-world individuals who change communities often face substantial moving costs. This paper utilizes historical changes in mobility costs to develop a new test of the Tiebout hypothesis.

We first generalize a standard Tiebout model to incorporate mobility costs and derive the following comparative static: as mobility costs fall, preference heterogeneity among residents *across* communities (weakly) increases. Under some standard assumptions (such as restricting individual preferences to a quadratic loss function), the variation of policies across communities will also increase. The empirical section begins by documenting the dramatic reduction in mobility costs over the last century. This suggests a natural test of the Tiebout hypothesis is to see whether heterogeneity across communities increases in the historical record. For this task, we assemble a vast dataset comprising all of the counties in the United States (except Alaska) over the 1850-1990 period. The county is the smallest government unit for which consistent detailed data are available for such a long period.

Almost all of our empirical results stand in opposition to the Tiebout prediction of increasing heterogeneity across communities. We first consider measures of local policy outcomes. Our analysis shows that the heterogeneity of local per capita education spending has declined significantly. The coefficient of variation across counties 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. These trends are robust to adopting

¹See for example Bewley (1981), Dowding et al (1994), Epple and Sieg (1999), Kollman et al (1997) and the references therein.

various heterogeneity measures (such as the proportion of money which must be reallocated to create a uniform distribution across counties) and to controlling for state-specific factors and outliers. To test the comparative static prediction regarding preference heterogeneity, we consider numerous proxies for public good preferences including electoral behavior, race, religious affiliation, age, nativity, and over the post-World War Two period, education, home ownership and income. Almost every preference proxy exhibits diminishing heterogeneity across communities over our sample period.² Two of the more graphic examples are that the proportion of blacks living in black majority counties decreased 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 the South, by rural-urban migration, or by states where counties have few fiscal responsibilities.

Finally, we address two possible criticisms. One concern is that our preference proxies may lose their salience over the sample period. To the contrary, we find that the proxies can explain roughly the same proportion of variation in local fiscal policies and vote outcomes in the beginning and end of the sample. A second concern is that our county-level analysis is inappropriate because Tiebout sorting occurs only over smaller jurisdictions such as municipalities. We show that most counties have important fiscal responsibilities, implying at least a weakened version of Tiebout sorting should hold. As a more formal test, we re-evaluate the government policy and preference proxy measures using the 92 municipalities in the Boston SMSA over the past century. All of the heterogeneity measures follow similar patterns at the municipal- and county-level in this sample, suggesting that our county-level analysis is not masking sub-county Tiebout sorting.

In total, these results suggest that the Tiebout model cannot be the complete explanation of individual locational choices and that Tiebout sorting has been historically overwhelmed by forces reducing across-community heterogeneity. This implies that any theoretical or empirical model which adopts a pure Tiebout framework is misspecified. (The final section suggests several alternative motives for individual moves which might

²Heterogeneity is measured with the dissimilarity index and the Gini coefficient. For income, the within- and between-Theil measure is also used.

explain the data in this paper). While such a conclusion may not seem surprising, several an ongoing literature takes a literal interpretation of the Tiebout model. In a recent empirical paper, Epple and Sieg (1999) assume that all locational choices are the result of costless sorting over community-specific housing prices and other “factors.” In the empirical implementation, they assume these factors are public good provisions and estimate the underlying preference parameters under the maintained hypothesis of a Tiebout equilibrium. In addition, most theoretical locational choice models assume that community choice is driven exclusively by local public goods and taxes. Some prominent recent examples include Fernandez and Rogerson (1998) and Nechyba (2000), who consider education quality/spending, Epple and Romer (1991), who investigate redistribution, and Glomm and Lagunoff (1999) and Nechyba (1997), who analyze generic local public goods. In addition, these models all assume no mobility costs.³ The latter assumption has also played an important implicit role in empirical papers (e.g. Cutler and Glaeser, 1997). Given that our work suggests that non-Tiebout motives must matter, a more general focus is needed.⁴

This paper develops a sharp test of the Tiebout hypothesis. In contrast, most of the existing literature investigates the extent of heterogeneity within communities, the motives for household mobility, and the degree to which fiscal policies are capitalized into property value (see Dowding et al, 1994). But these results do not provide satisfactory tests of the Tiebout hypothesis because, for example, it is unclear how large a deviation from perfect sorting is necessary to reject the hypothesis.⁵ Our comparative static provides a more meaningful test of Tiebout because it implies a direction of change—to greater sorting—that is empirically refutable. We recognize that no single piece of evidence presented here is convincing by itself, but the lack of strong support for

³Mobility costs have been added to other locational choice models. Some examples are Carrington et al (1996), Hercowitz and Pines (1999), and Kennan and Walker (2000).

⁴Some fruitful approaches here could utilize club theory, which allows individuals to explicitly care about their neighbor’s type (Scotchmer, 1994), and computational models, which can be constructed to deal with complex and non-analytical migration motives (Kollman, Miller and Page, 1997). Bayer (2000) improves upon the empirical literature by allowing residential choice to depend upon employment location and community racial composition (i.e. neighbors’ “type”).

⁵An implicit assumption in many empirical tests of Tiebout is that the system is in a consistent state of disequilibrium perhaps due to random shocks. These frictions are needed to explain why individuals move or why communities are heterogeneous. This is problematic since most Tiebout models consider static equilibria.

Tiebout sorting in any of the dozen or so measures we analyze constitutes a serious challenge to the hypothesis.

The 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 Cutler, Glaeser and Vigdor (1999) who analyze the segregation patterns of blacks in urban areas between 1890 and 1990. Second, this paper contributes to the growing literature on the efficiency implications of heterogeneity (e.g. Benabou, 1996). Alesina and La Ferrara (2000) show that community heterogeneity reduces participation in various social groups. Goldin and 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, thereby informing future work on this topic.

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 III documents the long-run decline in selected measures of moving costs. Section IV investigates historical changes in the variation of local policy outcomes while Section V examines trends in the dispersion of population characteristics that proxy for public good preferences. Section VI addresses some critiques of our tests and investigates heterogeneity trends at the municipal level. Alternative explanations and implications are briefly explored in the concluding section.

II. 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. 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 , allocated 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, $G_c \in \mathcal{G}$ where \mathcal{G} is a compact set. 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).

We will assume that agents only care about G_c in their community.⁶ Further assume that each agent belongs to a fixed type t characterized by the continuous utility function, $U_t(G_c)$. Let G_t be the unique ideal $G_c \in \mathcal{G}$ for type t , and presume there are T types where $T \leq N$. In some of the results derived below, we will consider special assumptions. In order of increasing restrictiveness they are:

Assumptions:

A1. Single-peaked preferences: $G_c \in \mathcal{G}$ 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 $U_t'(G_t) = 0$.⁷

A2. Quadratic preferences: $G_c \in \mathcal{G}$ and $U_t = -(G_t - G_c)^2$.

Community decisions.

Suppose that each community c chooses G_c^* to maximize the utility of its current residents. Thus,

⁶That 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.

⁷Single-peaked preferences can be understood as an individual maximizing a utility function containing a public good G and a private goods x subject to a budget constraint including a tax for the public good. See Romer and Rosenthal (1977).

$$(1) \quad G_c^* = \operatorname{argmax}_{G_c \in \Sigma_j U_{t_j}(G_c)}$$

This G_c^* exists given our assumptions. (1) is equivalent to majority rule with side payments.⁸ Denote the set of community public goods as $\mathbf{G}=(G_1^*, G_2^*, \dots, G_C^*)$. Note that some of the communities, z , may be empty, implying $G_z^* \in \emptyset$. In the migration decisions, assume each agent treats any empty community as setting G_z^* equal to her ideal G_t .

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, so long as each agent's expected order in the sequence is finite. Refer to each agent's turn to move as her *location 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.⁹

Assume that the mobility decisions are myopic. That is, each agent i takes the prevailing policies, \mathbf{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.

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:

$$(2) \quad G_c^* = \operatorname{argmax}_{G_e^* \in \mathbf{G}} U_{t_i}(G_e^*); \text{ and}$$

$$(3) \quad U_{t_i}(G_c^*) > U_{t_i}(G_d^*) + m_i.$$

In equilibrium, no individual will ever move from her current community no matter how many times her decision event occurs.

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 .

⁸Imagine individual locations are fixed. Assume that the majority rule outcome, G_C' , does not maximize aggregate welfare. The result is obvious if G_C' is Pareto dominated by G_C^* , so assume that some individuals ("losers") are better off under G_C' than with G_C^* . Then by the definition of argmax : (i) some individuals must be better off under G_C^* ("gainers"); (ii) there must be a Pareto improving transfer from the gainers to losers. This shows that G_C' could not be the majority rule outcome.

Social Welfare.

The aggregate measure of social welfare for any allocation \mathbf{A} is the sum of all agents' utility:

$$(4) \quad W = \sum_c \sum_i U_{t_i}(G_c^*).$$

Note that W is defined without reference to the mobility costs, m_i , because we are mainly interested in using welfare to measure heterogeneity.

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 (1), then W is maximized when each community contains only one type.

Proposition 1 makes it clear that population heterogeneity 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 total welfare via (1), 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 will then equal the negative of the population-weighted variance of the ideal policies:

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

Obviously, if preferences over public goods differ across types, per capita welfare in community c would be maximized if its population were homogeneous. In this example, the degree of heterogeneity of the community's population can naturally be measured by the population-weighted variance. In the general case, appropriately measuring or even defining the degree of heterogeneity is more difficult. The important point is that in the Tiebout model *population heterogeneity is inversely related to social welfare*.

⁹We are assuming that mobility costs for individual i are constant across communities. The results below

By assuming that preferences satisfy the single-peaked condition **A1** and that public good provisions are set via (1), we can show that sorting is a self-reinforcing 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 solve through the general equilibrium implications of her move on the subsequent public good provisions or the movements of others.¹⁰ 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 restrictions on individual preferences are needed for this result.

Proposition 2: When individual moves obey (2) and (3) and policies are set via (1),

- a. Any individual move strictly increases W .
- b. If m_i falls, then individual i either stays or moves and if she moves, then W increases. The 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 is,

$$(6) \quad \sum_{j/i} [U_{t_j}(G_{d/i}^*) - U_{t_j}(G_d^*)] \geq 0$$

where the inequality follows from the definition of argmax in (1) (over d/i versus d alone). Second, the net effect among the initial residents of region c is,

$$(7) \quad \sum_{j/i} [U_{t_j}(G_{c+i}^*) - U_{t_j}(G_c^*)] \leq 0$$

where the inequality follows from the definition of argmax in (1). Finally, the effect on agent i has two components. From her myopic comparison of d and c ,

will not change if these costs vary with some measure of "distance".

¹⁰For a sense of the complexity of this process, see Kollman et al (1997).

$$(8) \quad U_{t_i}(G_c^*) - U_{t_i}(G_d^*) > 0$$

where the inequality follows from (2) and (3). The other component considers how her move will change G in c ,

$$(9) \quad U_{t_i}(G_{c+i}^*) - U_{t_i}(G_c^*).$$

But (7)+(9) equal,

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

by definition of the argmax in (1). Thus the total effect (6)+(7)+(8)+(9) is positive.

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.

Comments:

(i) The intuition of the argmax argument in (6), (7), and (10) is as follows: A community cannot be made worse off by adjusting 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 $G_{d/i}^*$ as under G_d^* . Similarly, in aggregate the residents of c (including i) are at least as well off under G_{c+i}^* as under G_c^* . The change in G_c may harm its initial residents, but the gain to i more than offsets their losses.

(ii) Proposition 2 also holds in a world of Leviathan governments where the \mathbf{G} is fixed. Here, (6), (7) and (9) are each zero, but (8) is positive by (2) and (3).

(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 decreases).¹¹

¹¹In 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.

Observation: Under **A1** with $C=T=2$ or under **A2**, when local policies are set via (1) then sorting is associated with increased variation of policy outcomes across communities.

A formal demonstration of the result is contained in Appendix A. Consider first when **A1** with $C=T=2$ holds.¹² 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 (1) and **A2**, the policy in a given community is the mean of residents' ideal policies and the aggregate population-weighted mean policy is independent of the distribution of types across communities. Any move obeying (2) and (3) 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 according to Tiebout. A reduction in mobility costs has a non-negative effect on social welfare. Because social welfare in the Tiebout model is positively related to sorting according to type, falling mobility costs increase heterogeneity between communities. We also find that sorting increases the variation of local policy outcomes under several variants of the model. The remainder of the paper tests these predictions empirically.

¹²The observation also holds in a generic case when the population is initially completely diffuse, implying policies are identical across communities. At 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.

III. Documenting Declining Mobility Costs

The conventional wisdom holds that mobility costs have fallen over time. 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. Few migrants suffer such a loss today.¹³ And, as we argue below, communication improvements have reduced one of the key costs of moving, the lost contacts with one’s friends and extended family in the home community.

There is clear evidence that physical moving costs have fallen over the last century. The most obvious change is the spread of the personal automobile. In 1900, 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 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 cost of driving the more reliable, comfortable cars of today average 54.9 cents. 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 four million today.¹⁴ An

¹³D. Galenson, "Settlement and Growth of the American Colonies" Ch. 4 of *Cambridge Economic History of the United States*, Vol. 1,

¹⁴Motor Vehicle Manufacturers Association, *Motor Vehicle Facts and Figures: 1999*, p. 51; Hiram Percy Maxim, "Some Data on the Cost of Operating Automobiles for Commercial Purposes" *Scientific American* Supplement No. 1479, 7 May 1904, pp. 23694-95. US Federal Highway Administration, *Highway Statistics*, various years. This 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

important consequence of the spread of the automobile was to weaken the link between work and residence locations, 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 1995.¹⁵ The costs of in-town moves have declined as well. For example, when L. S. and Anna 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.¹⁶

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

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>).

¹⁵It 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 to the price decrease. Over the 1994-96 period, the average billed shipment weighed just under three tons (5919 pounds) and traveled 1261 miles. “Revenue Need and the Cost of Transporting Household Goods Personal Effect Shipments,” <http://www.avatar-moving.com>. For 1936 rates, see Household Goods Rates of A. D. Rymers & Sons, 12 MCC 541-544 (1939). For estimates of the average weight of shipments in the past, see William Ashby, et al. *Motor Truck Red Book 1940 Edition* (New York: Traffic Publishing Co. 1940) p. 24.

¹⁶1945 information is from http://www.uhaul.com/corporate/company_history.html; downloaded 16 June 2000; current rental rates are based on a 19 June 2000 phone conversation with I-5 Auto Sales, Woodland WA.

encourages more dispersed production activity, implying people are less tied to a particular community for employment reasons (rather than for local policies). 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.¹⁷

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-trend 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.¹⁸ 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.¹⁹

IV. Trends in the Dispersion of Local Taxes and Revenues

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

¹⁷The reduction in postal rates, especially across country, was also notable. In 1860, during the Pony Express period, it cost \$10 in specie 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. "History of the US Postal Service," USPS Web Site, <http://www.usps.gov/history>.

¹⁸The data on migration in 1940 are from IPUMS sample, <http://www.ipums.umn.edu/~pmigration/>; 1970 data are from the 1980 Census of Population Vol. 1 Characteristics of the Population Ch. C General Social and Economic Characteristics pt. 1 United States Summary PC80-1-C1 (Dec. 1983) p. 1-18; 1990 are from <http://venus.census.gov/cdrom>. Data on residence of state of birth are from the US Historical Statistics, Series C1-14; 1980 Census of Population Vol. 1 Characteristics of the Population Ch. C General Social and Economic Characteristics pt. 1 United States Summary PC80-1-C1 (Dec. 1983) p. 1-18; and from <http://venus.census.gov/cdrom>.

¹⁹Consider 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.

outcomes and in the sorting of population types across localities. We now examine such trends over the 1850-1990 period using counties-- the smallest government for which consistent historical data are available-- as our unit of analysis.²⁰ Whenever 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.²¹

The first test involves variation in local policy outcomes. We initially consider per capita education spending, one of the most prominent local policies. As discussed more fully in the Data Appendix, each county's data includes all direct education spending within its boundaries. Table 1 presents two heterogeneity measures over the 1890-1992 period. The first measure is the population-weighted coefficient of variation (CV). This measure increases with the standard deviation across counties while controlling for changes in the mean (education spending grew significantly as income increased). The second measure (DG) calculates the proportion of total spending in each year which would have to be re-allocated across counties to yield a uniform, per capita distribution (this is related to the dissimilarity index which is discussed in the next section).²² Table 1 shows that both measures markedly *decrease* over time. The values in 1992 are less than half of the 1890 values.

We consider two robustness checks (but omit the formal results in the interest of brevity). The first is to control for outliers by using the ratio of the inter-quartile difference to the median (this is analogous to the CV but is less sensitive to outliers). This ratio has a similar downward trend. The second is to control for state-fixed effects. This is useful because intergovernmental grants and income, key determinants of local

²⁰It is worth noting that the growth in the number of counties by itself will increase the measured degree of heterogeneity across counties. The number of counties rose from 1626 in 1850, to 2867 in 1900, 3065 in 1920 and 3141 in 1990. Much of this growth was due to splitting counties rather than incorporating new area (for a graphical depiction, see <http://www.ac.wvu.edu/~stephan/Animation/us.gif>). Increasing the number of communities induces a "statistical sorting," which means that even if the spatial distribution of types remains constant we should observe greater heterogeneity of either individual type or policy outcome. To see this point, consider first the extreme case where every person lives in his own county. By definition this case will have maximal heterogeneity since each county is composed of exactly one type of person. In contrast, when there is only one county, there is minimal heterogeneity. (See Dowding et al, 1994 for a detailed discussion of this issue.) This biases our indices in favor of *accepting* the Tiebout model, which predicts a trend towards greater heterogeneity across counties. Finding the opposite pattern will be strong evidence against the Tiebout model.

²¹Data sources and detailed descriptions of all variables are contained in the Data Appendix. Alaska is omitted due to inconsistencies in its county codes.

spending, vary widely across states. (Note that intergovernmental grants mitigate Tiebout sorting, thereby casting doubt on rigid use of Tiebout models.) To create heterogeneity measures controlling for state fixed effects, we substitute state-level averages for the national average in the formulae. The modified measures continue to decline, indicating that even within states there is decreasing heterogeneity of education spending.

The second set of local policy outcomes we consider are real per capita taxes and revenues. These measure the overall level of government activity in the county. 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. The main difference between taxes and revenues is intergovernmental grants, which were typically small before 1945.²³ Table 2 presents the two heterogeneity measures discussed above for the 1870-1992 period. As with education spending, there is a sharp drop in the dispersion across counties of all four fiscal measures. As an example, the CV for Tax_2 fell by nearly one-half between 1870 and 1992. These downward trends are still apparent when we use state fixed effects or use the ratio of the inter-quartile difference to the median as our heterogeneity measure (results omitted).²⁴

In conclusion, we find that local fiscal policies have converged substantially across counties over the last century. This conflicts with the prediction of the generalized Tiebout model that local policies should become more heterogeneous as mobility costs decline. The rationale for the theoretical result was that falling mobility costs induce greater sorting of individual preference types. We next investigate whether heterogeneity of types across counties has in fact increased as predicted.

²²Formally, the second measure is $0.5G^{-1}\sum_j P_j|G_j - G|$ where G_j is per capita spending in county j , G is mean per capita spending for all counties, and P_j is the share of total population in county j .

²³For example, in 1902 local government taxes are 83.2 percent of local government revenues (U.S. Bureau of the Census, 1948). While intergovernmental grants grew in importance, local government taxes are still 39.4 percent of local government revenues in 1992.

²⁴Another concern is that there may be returns to scale in government services. Our county-level regression analysis finds no consistent relationship between any of the per capita fiscal data and population and population-squared (results omitted).

V. Trends in the Dispersion of Individual Types

The Tiebout framework presumes that an individual's location decision depends on his policy preferences. With the secular drop in mobility costs, the extended Tiebout model predicts that policy preferences should become more heterogeneous across communities. Such preferences are not directly observable, leading us to adopt the strategy of examining numerous characteristics that proxy for individual types.²⁵ 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 we argue in the next section, even if the observable characteristics are noisy signals or available categories are too coarse, our measures of population heterogeneity should remain informative.²⁶

A. Proxies for Types

This sub-section motivates the preference proxies we use (details on their construction and sources are in the Data Appendix). The first variable we consider is party vote shares in presidential elections. The idea is that individuals vote for the party whose platform is closest to their own ideal policy, implying those voting for a particular party have similar preferences. If all individuals in a local area vote for the same party, they are relatively alike and the area is homogeneous. If they split their votes, the area is heterogeneous.²⁷

Second, it is often observed that members of racial groups share economic interests and maintain strong common party affiliations. For example, the General Social

²⁵An 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.

²⁶There is also evidence that several of our type measures are transmitted from parents to children. Piketty (1995) cites the extensive literature showing that political preferences have an important hereditary component even after controlling for income and social class. Newport (1979) finds that in the mid-1970s over two-thirds of individuals maintain their childhood religion.

²⁷It 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 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.

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. There are also notable racial differences in the GSS over political ideology and attitudes towards government redistribution.²⁸ We use the black population share to proxy for these beliefs.

Third, we consider 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 policy preferences.²⁹ 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.

We consider five other demographic proxies (some of which are only available over a limited period). The first proxy involves age groups. 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. Alternatively, the old population share (those at least 65 years old) is used, since the elderly are more likely to oppose education spending.³⁰ The second proxy is the fraction of the adult population that is married, which is also related to family structure and demand for school spending. Our third proxy is the proportion of homes which are owner occupied. Home owners are typically wealthier and have greater civic involvement in the community. The fourth proxy is education level which is likely to be related to income, wealth, and attitudes toward government. 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. Fifth, the foreign-

²⁸The General Social Survey (1999) is a micro dataset of individual attitudes collected over the 1972-96 period. The variables we consider here are “race,” “partyid,” “polviews,” “religid,” “nateducy,” and “govunemp.”

²⁹Based on the 1990 National Election Study, Legee and 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. 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. Noll (1990) also documents the historical link between religion and politics using largely non-quantitative analysis.

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 ($N=1,181$) consider 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.

born represents another distinctive population with important ramifications for local politics. The foreign-born share is of additional interest because immigrants may sort across communities for non-Tiebout motives, for example, to take advantage of social networks or ports of entry.

Finally we consider income, 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, 17 income groups in 1979 and 25 groups in 1989. Although the county-level data cover only fifty years, we still expect to see increased sorting by income across local communities if the generalized Tiebout hypothesis holds.

B. Heterogeneity Measures

To measure the heterogeneity of the discrete types across communities, we employ the dissimilarity index and the Gini coefficient.³¹ These measures, which are commonly used in the segregation literature, have 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. 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.³² Third, the measures weight the counties by their population.

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

³¹See Duncan and Duncan (1955), Massey and Denton (1988), and Reardon (1998). In all cases, we also calculated an entropy index, which yielded essentially identical results and is omitted for space reasons.

³²More 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).

‘surplus’ and ‘deficit’ communities. In multiple (≥ 2) type comparisons, the dissimilarity index, D , and the Gini coefficient, GC , are defined as:

$$(11) \quad D = \frac{1}{2} \frac{\sum_t \sum_j N_j |P_{tj} - P_t|}{(N \sum_t P_t (1 - P_t))}.$$

$$(12) \quad GC = \frac{1}{2} \frac{\sum_t \sum_k \sum_j N_k N_j |P_{tk} - P_{tj}|}{(N^2 \sum_t P_t (1 - P_t))}.$$

where, N_j is the total population in community j , N is the total population, P_{tj} is the share of type t in community j and P_t is the share of type t in the total population.

Because the income distribution data is in ordered categories, we can perform a more detailed within- and between-decomposition. We employ two additively separable Theil measures,

$$(13) \quad I_1 = -\sum_j P_j \sum_s P_{sj} \log(\mu_{sj} / \mu_j) + -\sum_j P_j \log(\mu_j / \mu)$$

$$(14) \quad I_2 = \sum_j P_j \sum_s P_{sj} \log(\mu_j / \mu_{sj}) + \sum_j P_j \log(\mu_j / \mu)$$

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 (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.³³

C. Empirical Results

This section investigates the historical evolution of the heterogeneity indices of the type proxies. Figure 3 plots the heterogeneity measures 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.) Contrary to the prediction of the generalized Tiebout model, there is a gradual downward trend, especially after 1892.³⁴ For example, the Gini trend line has a slope of 0.01040 per decade over the entire period, and of

³³Because the available data is grouped in income ranges, both lower- and upper-bound Gini's are computed for each community j ,

$$GL_j = (2 \mu_j)^{-1} \sum_s \sum_t P_{sj} P_{tj} |a_{sj} - a_{tj}| \quad GU_j = GL_j + (\mu_j)^{-1} \sum_s P_{sj}^2 (a_{sj} - a_{s-1})(a_s - a_{sj})(a_s - a_{s-1})^{-1}$$

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

³⁴The 1860 election was highly unusual because four major parties- Republicans, Democrats, Southern Democrats, and Constitutional Unionists- participated in the sectionally-divided contest.

0.01357 from 1892 on (the index decreases from 0.38 to 0.24 between 1892 and 1988). Heterogeneity also falls (to a lesser degree) in the post-World War Two period. The convergence of county election results could be due to the declining importance of third parties. But Figure 3 shows there is no such trend, and more importantly, the heterogeneity measures calculated only for the two major parties also fall (figure omitted). Nor are these patterns explained by the Democratic party's slow loss of control of the South. Figure 4, which plots the heterogeneity measures dropping the South from the sample, displays the same general reduction in voting heterogeneity.³⁵

The heterogeneity indices for the black population share decline more noticeably. Figure 5 shows the dissimilarity index and Gini coefficient 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 (figure omitted). Our results run counter to Cutler, Glaeser and Vigdor (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.

The data on religious affiliation, displayed in Figure 6, also reveal counties are becoming more alike over time.³⁶ (The series has a break in 1890 when the unit of

³⁵The South includes Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia.

³⁶One caveat to these results is that the group sizes do not sum to one because some individuals are affiliated with religions not in the sample or are unaffiliated. When these two groups are lumped together, the series is essentially flat. We are skeptical of the validity of this series since it groups together two quite different groups of individuals, the unaffiliated and the non-sampled.

observation shifts from seatings to church members.) A similar pattern is present when the South is omitted. Figure 7 presents data for the age groups. 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 1850.³⁷ The overall decline in heterogeneity among the elderly is particularly noteworthy because it runs counter to the tremendous growth of retirement communities in Florida and Arizona.

Figure 8 shows the heterogeneity trends for home-owner occupation and education levels for roughly the last fifty years. The home-ownership indices each fall by a quarter over the 1930-90 period while the education indices decline slightly between 1940 and 1990. Figure 9 shows information about the foreign-born population share. There is a slight downward trend in heterogeneity over the whole sample but a noticeable rise between 1960-90. The increase at the end of the sample is due to the disproportionately rapid growth of Hispanics in California, Texas, Florida, New Jersey, and New York. The one proxy which exhibits increasing heterogeneity is the proportion married (figure omitted). The dissimilarity index has a small increase from 0.06 in 1930 to 0.10 in 1990.

It is also interesting to consider the *absolute level* of sorting. Of all the proxies, only race in the pre-1940 period displays a high degree of heterogeneity, that is, a dissimilarity index exceeding the conventional standard of 0.6 (see Cutler, Glaeser and Vigdor, 1999, p. 458). This also seems to be at odds with Tiebout's predictions of sorting by type.

Finally, Table 3 presents income inequality/heterogeneity measures for 1949, 1979 and 1989. The generalized Tiebout model predicts that decreasing mobility costs will induce better sorting by individuals according to income groups, leading to decreasing income inequality within counties and increasing inequality between counties. We first create two within-county measures. We calculate a CV and Gini index for the income distribution in *each* sample county and then average the county values using population-weights. Contrary to the prediction that the within-county measures should

³⁷While there is no county-level electronic data for the elderly between 1870 and 1940, 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.

fall, these measures in 1989 were generally close to their 1949 levels. (Note that the measures fell slightly between 1949 and 1979, but the increases over the next decade more than offset these changes.) We next create two Theil decomposition measures to account for inequality both within- and between-counties. The empirical results in the second panel of Table 3 also undermine the generalized Tiebout model because the within-county inequality measures are roughly constant while the between-county measures *decrease*.

In conclusion, this evidence indicates a reduction in heterogeneity of preference types across counties over the twentieth century. The actual decline in heterogeneity is even greater than our figures suggest due to the statistical sorting induced by the growing number of counties (cf. footnote 20). This is in conflict with the prediction of the generalized Tiebout model since mobility costs have fallen over this period.

VI. Potential Criticisms

This section addresses two possible criticisms of applying our empirical results to test the Tiebout hypothesis: the use of preference type proxies and the use of counties rather than smaller governments.³⁸ We consider each of these issues in turn. One concern is that we are only using proxies rather than actual policy preferences in Section V. But even if the observable characteristics are noisy signals of the true types or available categories are coarse, our measures of population heterogeneity should remain informative. Consider the case where the available proxies are coarse, and consequently combine several distinct types. According to Tiebout's reasoning, each of the types will sort into separate homogeneous communities. Our heterogeneity measures using the coarse proxies will still register the sorting.

Now take the case where the proxies are noisy signals. As a simple example developed in Appendix A shows, as long as the conditional probabilities linking the

³⁸It has also been suggested to us that the observed decline in heterogeneity across counties was driven by the shift from a rural to urban society. The idea here is that rural areas are closer to Tiebout's homogeneous communities while cities tend to be a more diverse. To investigate this conjecture, we repeated our analysis with a sample restricted to those counties where at least 85 percent of the population was rural in 1990. (Roughly 30 percent of the counties meet this criterion.) For all measures but one, the same downward heterogeneity trends are apparent. (The one exception is race, which remains roughly constant over the 1890-1990 period.) Overall, these results run counter to the prediction that these rural counties should have constant or increasing heterogeneity over time.

proxies to the true types are distinct and constant, *when the unobservable types sort, there is also increased sorting of observable proxies*. This suggests an explanation for decreasing observed heterogeneity that is compatible with the Tiebout model, namely that these conditional probabilities have become less distinct over time. For example, Baptist affiliation could provide a weaker signal about political preferences in 1990 than in 1850. Thus, the observed types might appear less sorted over counties even if the true types are becoming more sorted.³⁹

To investigate whether our proxies have lost their salience, we estimate the regression,

$$(15) \quad Y = \mathbf{t}^{\text{proxy}} + e$$

where Y is either taxes per capita, revenues per capita, or the presidential vote share and $\mathbf{t}^{\text{proxy}}$ is the vector of our proxies available over the full sample. We evaluate this equation separately for each year at the county level using least squares with population weights. If the proxies became less relevant over time they should do a poorer job explaining the variation in local policies or presidential vote shares. No such trends are apparent. In regressions explaining local taxes, the R^2 equals 0.56 in 1870 and lies in the 0.4-0.7 range over the 1962-92 period. In regressions explaining presidential vote shares, the R^2 actually rises, climbing from 0.16 in 1852 to 0.44 in 1988. Moreover, the decreasing salience argument cannot be used to explain all of the data within the Tiebout framework. It is difficult to understand the reduced heterogeneity across counties in presidential voting, education spending, local taxes, and revenues, since these are direct measures of preferences and policy outcomes.

A second concern is that the county is too coarse a unit of analysis and that Tiebout sorting applies only to smaller government units, such as municipalities and school districts, which provide most local public services. We offer four responses to this critique. First, Tiebout (1956, p. 418) originally argued that his sorting argument applied to all local governments and “with less force to state governments.” (We have replicated

³⁹Suppose the true type is a linear combination of the proxy and some unobserved component, $t = \mathbf{t}^{\text{proxy}} + (1 - \alpha)\mathbf{t}^{\text{unobs}}$ with $\alpha \in [0,1]$, $\mathbf{t}^{\text{proxy}} \perp \mathbf{t}^{\text{unobs}}$. Say that initially the proxy is a perfect measure of type but eventually becomes uninformative ($\alpha_{\text{initial}} = 1$, $\alpha_{\text{final}} = 0$), and that there is always complete sorting of the true t 's across communities. The observed proxies will be perfectly sorted initially but by the end of the sample will be completely unsorted.

our heterogeneity analysis at the state level and found the same basic patterns.) Second, counties often do have important fiscal responsibilities. Over the sample period, the county government share of total local taxation ranged between 24 percent (1992) and 37 percent (1870). While many county expenditure programs are low profile, counties typically control or strongly influence highway and police spending policies. Given voters care about such policies and taxes, Tiebout sorting should occur at the county level. In fact, our results indicate reduced sorting of type proxies even in states where counties have major fiscal responsibilities.⁴⁰ Also, it is worthwhile recalling that counties are the most important local governments in many southern and southwestern states. Third, this critique does not explain the observed reduction in heterogeneity across counties. If Tiebout sorting is occurring at a sub-county level, there must be increased mixing within counties of different types of municipalities and school districts over time.

Fourth, we can offer a direct test of the critique by replicating our analysis on the 92 municipalities in the Boston SMSA (1980 definition).⁴¹ 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. Massachusetts counties have few fiscal responsibilities and municipalities typically directly provide public schooling, one of the most prominent local public goods. This means any Tiebout sorting occurring at the sub-county (municipal) level should be apparent here. In addition, there is clear evidence of a secular decline in intra-Boston SMSA transportation costs.⁴² Finally, detailed municipal demographics, election outcomes, and government finances are available for the Boston area dating back to the late-nineteenth century.

Three results are apparent in Table 4 and Figure 10. First and most importantly, the same heterogeneity trends are found at the municipal and county levels. (This is not surprising given both sets of indices are principally driven by high-population areas.)

⁴⁰Our formal test is to limit the sample to counties in states that have high per capita county spending. We use the 1992 data reported in the US Bureau of the Census (1997) as our benchmark and define a state as “high spending” if per capita county spending is greater than \$850. The U.S. average is \$691. The high spending states are California, Florida, Hawaii, Maryland, Nevada, New York, North Carolina, Tennessee, Virginia and Wyoming (Alaska is also a high spending state but it is omitted in all samples). We repeated all of our former analysis on these counties and found the same downward trend in heterogeneity though it is not as steep as in the full sample.

⁴¹The Data Appendix contains details on the data sources.

⁴²For example, see the discussion of the development of the Boston mass transit system in George Sanborn, *The Chronicle of the Boston Transit System*. www.mbta.com.

This suggests that our general focus on county-level data is not masking sub-county Tiebout sorting. Second, the municipal-level political and government expenditure measures are not consistent with increased Tiebout sorting. As Table 4 shows, the population-weighted CV of total government spending, measured by the per capita current operation budget, falls over the 1906-92 period. We also investigate two major spending categories, protection (basically police plus fire services) and education, which typically comprise over one-half of total spending. The CV for per capita protection spending falls while the CV for per capita education spending remains roughly constant.⁴³ Figure 10 shows no clear trend in heterogeneity across municipalities of presidential election votes over 1868 to 1988. There is even a trend towards reduced heterogeneity over the last fifty years.⁴⁴ Third, the demographic variables often do show growing heterogeneity between communities as the generalized Tiebout model predicts (figures omitted). But none of the dissimilarity indices ever reach 0.6, the conventional standard for a high degree of heterogeneity.⁴⁵ The population of the Boston SMSA is far from the level of sorting that Tiebout would predict.

In conclusion, it appears that counties do provide an appropriate test of the generalized Tiebout model. Not only do counties have important fiscal responsibilities but they manifest similar trends in heterogeneity as do municipalities (at least for the Boston area). And despite selecting a case study where municipalities have significant fiscal powers, we do not find strong evidence for Tiebout's prediction of increased sorting at a sub-county level.

⁴³There is a spike up in education spending heterogeneity in the 1960s which stems from Boston's relative reduction in education spending. 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 Sheehan, 1984) and changes in the state school aid formula (see Weiss, 1970).

⁴⁴The same pattern remains if we only consider votes for the two largest parties in each year.

⁴⁵It is more difficult to say what a "small" degree of heterogeneity is for income. However, our (unreported) within- and between-Theil indices are comparable to those in Mookherjee and Shorrocks (1982) who decompose income inequality between age groups for the entire UK over 1965-1980. They are also similar to our earlier county-level values for the US as a whole. Strong Tiebout sorting should have induced smaller within- and larger between-inequality for the Boston municipal-level sample.

VII. Conclusion

This paper considers a new test of the Tiebout hypothesis related to the secular decline in mobility costs. The augmented Tiebout model predicts greater heterogeneity across communities in both resident preferences and government policies as movement becomes easier. 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 counties) have become more alike one another. These results call into question the local public economics literature that adopts a rigid Tiebout framework to explain community composition and mobility choices.

Although a full investigation of the forces causing the decline in heterogeneity across communities must be reserved for future work, we briefly consider several alternatives. One possibility involves changing production relationships. For example, the complementarities between different types of labor may have increased. If it has grown increasingly important to reside near others with different characteristics (and correlated preferences), then Tiebout sorting may not occur. A second factor could be changing attitudes. There may be a growing fraction of the population who prefer to live in more diverse communities. Such demand for heterogeneity may be a normal good, so income growth would promote greater mixing of different types of individuals. A related point involves the dramatic changes in racial attitudes in the post-World War Two period. The fear of being a minority, subject to exploitation or simply to having one's political preferences ignored, may have decreased, leading members of small groups to spread out. A third mechanism involves immigrants. It is common to think of first-generation migrants clustering together with others of their own type (language, religion, ethnic capital). As immigration rates fell (until the most recent period), these "ports of entry" have declined in importance, leading to less apparent sorting. However, the increased segregation of the foreign-born over the last 30 years seems to run contrary to this explanation.

Finally, the rapid growth of the federal government since the New Deal may have made local-level policies less significant factors in individual residential decisions. The alternative motives discussed above may have become more important as the independent

role of local governments has waned. But this point raises concerns regarding the contemporary movement to devolve political decision-making to a more decentralized level. Tiebout is often invoked, explicitly or implicitly, to justify decentralized provision of government policies. Tiebout argued that because the residents of each community share relatively similar preferences, local provision can avoid the free-rider problem associated with public goods. The results in this paper cast serious doubt on this presumed advantage of decentralization. Whatever benefits devolution provides in the long run, our analysis leads us to expect local conflict and population re-sorting in the short run.

Regardless of the underlying mechanisms, it is clear that in aggregate the forces driving individual location decisions are neither consistent with nor simply independent of Tiebout motives. Over the past century, these forces have worked powerfully to counteract Tiebout sorting. While Tiebout predicts that heterogeneity across communities should increase, we find the opposite for a wide variety of local policy outcomes, elections, and demographic measures. Incorporating these new results into models of local public finance will be an important and challenging avenue for future work.

References

- Alesina, Alberto and Eliana La Ferrara (2000). Participation in Heterogeneous Communities. *Quarterly Journal of Economics*, forthcoming.
- Bayer, Patrick (2000). Tiebout Sorting and Discrete Choices: A New Explanation for Socioeconomic Differences in the Consumption of School Quality. Yale University working paper.
- Benabou, Roland (1996). Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance. *American Economic Review*. 86: 584-609.
- Bewley, Truman (1981). A Critique of Tiebout's Theory of Local Expenditures. *Econometrica*. 49: 713-40.
- Carrington, William, Enrica Detragiache and Tara Vishwanath (1996). Migration with Endogenous Moving Costs. *American Economic Review*. 86: 909-930.
- Cutler, David M. and Edward L. Glaeser (1997). Are Ghettos Good or Bad? *Quarterly Journal of Economics*. 112: 827-72.
- Cutler, David M., Edward L. Glaeser and Jacob L. Vigdor (1999). The Rise and Decline of the American Ghetto. *Journal of Political Economy*. 107: 455-506.
- Duncan, O. and B. Duncan (1955). A Methodological Analysis of Segregation Indices. *American Sociological Review*. 20:210-17.
- Dowding, Keith, Peter John and Stephen Biggs (1994). Tiebout: A Survey of the Empirical Literature. *Urban Studies*. 31: 767-97.
- Epple, Dennis and Thomas Romer (1991). Mobility and Redistribution. *Journal of Political Economy*. 99: 828-858.
- Epple, Dennis and Holger Sieg (1999). Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*. 107: 645-81.
- Fernandez, Raquel and Richard Rogerson (1998). Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform. *American Economic Review*. 88: 813-833.
- Gastwirth, Joseph L. (1972). The Estimation of the Lorenz Curve and the Gini Index. *Review of Economics and Statistics*. 54: 306-16.

- Glomm, Gerhard and Roger Lagunoff (1999). A Dynamic Tiebout Theory of Voluntary vs. Involuntary Provision of Public Goods. *Review of Economic Studies*. 66: 659-677.
- Goldin, Claudia and Lawrence F. Katz (1999). Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910 to 1940. *Journal of Interdisciplinary History*. 29: 683-723.
- Hercowitz, Zvi and David Pines (1997). Migration Between Home Country and Diaspora: An Economic Analysis. *Journal of Public Economics*. 65: 45-59.
- Iannaccone, Laurence R. (1998). Introduction to the Economics of Religion. *Journal of Economic Literature*. 36: 1465-95.
- Kennan, John and James Walker (2000). Geographical Wage Differentials, Welfare Benefits and Migration. Working paper.
- Kollman, Ken, John H. Miller and Scott E. Page (1997). Political Institutions and Sorting in a Tiebout Model. *American Economic Review*. 87: 977-92.
- Leege, David C. and Lyman A. Kellstedt, ed. (1993). *Rediscovering the Religious Factor in American Politics*. M.E. Sharpe, Armonk, NY.
- Massey, Douglas S., and Nancy A. Denton (1988). The Dimensions of Residential Segregation. *Social Forces*. 67(2): 281-315.
- Mookherjee, Dilip and Anthony Shorrocks (1982). A Decomposition of the Trend in UK Income Inequality. *Economic Journal*. 92: 886-902.
- Nechyba, Thomas (1997). Local Property and State Income Taxes: The Role of Interjurisdictional Competition and Collusion. *Journal of Political Economy*. 105: 351-385.
- Nechyba, Thomas (2000). Mobility, Targeting, and Private-School Vouchers. *American Economic Review*. 90: 130-146.
- Newport, Frank (1979). The Religious Switcher in the United States. *American Sociological Review*. 44: 528-52.
- Noll, Mark A., ed. (1990). *Religion and American Politics: From the Colonial Period to the 1980s*. Oxford University Press, New York.
- Piketty, Thomas (1995). Social Mobility and Redistributive Politics. *Quarterly Journal of Economics*. 110: 551-84.
- Reardon, Sean (1998). Stata module to compute multiple-group diversity and segre-

- gation indices. ideas.uqam.ca/ideas/data/Softwares/bocbocodeS375001.html.
- Romer, Thomas and Rosenthal, Howard (1977). Bureaucrats Versus Voters: On the Political Economy of Resource Allocation by Direct Democracy. *Quarterly Journal of Economics*. 93: 563-87.
- Sacks, Seymour (1972). *City Schools/Suburban Schools: A History of Fiscal Conflict*. Syracuse University Press, Syracuse.
- Scotchmer, Suzanne (1994). Public Goods and the Invisible Hand, Chapter 4 in *Modern Public Finance* edited by J. Quigley and E. Smolensky, Harvard University Press.
- Sheehan, J. Brian (1984). *The Boston School Integration Dispute: Social Change and Legal Maneuvers*. Columbia University Press, New York.
- Shorrocks, A. F. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica*. 48: 613-26.
- Tiebout, Charles (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*. 64: 416-24.
- Weiss, Steven (1970). *Existing Disparities in Public School Finance and Proposals for Reform*. Federal Reserve Bank of Boston, Research Report 46.

Appendix A

This appendix contains the example of sorting with single-peaked preferences, proofs of cases where sorting is associated with increased policy variation, and an example mapping observed characteristics into unobserved true types.

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 (1). 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 (1), 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 (2) and (3) 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 (2) and (3), 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 (2) and (3) increases the aggregate population-weighted variance of policies. Under **A2** and (1), 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 $G_d^{*1} = (G_d^{*0} N_d - G_s^0) / (N_d^0 - 1)$ 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:

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

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

Example of Proxies as Noisy Signals

Example D: Suppose there are two true preference types, say 'yes' and 'no' opinions regarding a particular policy. Suppose that the 'yes' camp makes up a fraction, p , of the population and the 'no' camp the remainder, but that these fractions are unobservable. Suppose, instead, we can observe the proportion of the population, s , possessing some characteristic that plays no direct role in the sorting behavior. Denote the probability of choosing 'yes' conditional on the presence of the characteristic as q and conditional on its absence as r . These probabilities must satisfy the equation: $p = sq + (1-s)r$. Note that either (i) $p = q = r$; (ii) $q > p > r$; or (iii) $q < p < r$.

If the observable types have any distinct opinions (case (ii) or (iii) holds), then the distribution of the characteristic will be related to the unobservable type. Without loss of generality, let $q > r$. This means that individuals possessing the characteristic are more likely to favor the policy. We know that the fraction of the population possessing the characteristic is $s = (p - r) / (q - r)$. Simple calculus shows $ds/dp > 0$. This means an increase in the fraction of individuals favoring the policy implies an increase in the fraction possessing the correlated characteristic.

Data Appendix (not for publication, to be made available on the web).

Mobility Costs: Sources of Figures 1-2

Figure 1: Transportation Costs

Bureau of the Census, *Historical Statistics of the United States on CD-ROM: Colonial Times to 1970 - Bicentennial Edition*, Susan B. Carter et al., ed., Cambridge University Press 1997), Vol. I, pp. 210-11; Vol. II, pp. 727, 769-70; *1998 Economic Report of the President*, H. Doc. 105-176, Washington, DC. GPO, Table B-60. US Department of Transportation, *Transportation Statistics Annual Report*, various years.

Figure 2: Real Cost of Communications

Historical Statistics of the United States, Vol. I, pp. 210-11; Vol. II, pp. 784, 790; Federal Communications Commission, *1999 Statistics of Communications Common Carriers*, Table 7.1-AT&T Historical Rates at Years' End; *1998 Economic Report of the President*, H. Doc. 105-176, Washington, DC. GPO, Table B-60.

County Data

The variables below are considered at the county level. All counties are included though data for Alaska is omitted because of inconsistencies in its county codes. The number in parentheses following each data source indicates the years for which that source is used.

Policy outcomes

Education Spending

Frequency: Sporadic (1890, 1932); every five years (1957-1992). Sources: US Census Office, 1895 (1890), U.S. Bureau of the Census, 1935 (1932), ICPSR study 7736 (1957-1972), U.S. Bureau of the Census, 1998 (1977-1992).

Notes: In all cases the values sum up education spending by all governments within each county and include capital expenditures. The exact definitions of the spending categories are: for 1890, "ordinary expenditures on public common schools;" for 1932 "[school] government-cost payments operation and maintenance;" for 1957-1992 "direct general expenditures for education" which include spending on all education institutions within the county. See also the note in the taxes and revenues section below regarding New York area counties.

Taxes and Revenues

Frequency: Approximately every 10 years (1870-1942); every five years (1957-1992). Sources: the Census of Government. ICPSR study 3 (1870-1880), US Census Office, 1895 (1890), U.S. Bureau of the Census, 1908 (1902), U.S. Bureau of the Census, 1915 (1913), U.S. Bureau of the Census, 1924 (1922), U.S. Bureau of the Census, 1935 (1932), U.S. Bureau of the Census, 1944 (1942), ICPSR studies 7736, 8256, 9251 (1957-1987), U.S. Bureau of the Census, 1998 (1992).

Notes: (i) Data for the five counties in the greater-New York area (Bronx, Kings, New York, Queens, Richmond) are aggregated into one county. This is necessary since the Census puts all city spending in just one county, but shifts the identity of this county over

time. All results are qualitatively similar if instead these five counties are excluded. (ii) Rhode Island is excluded because its counties have no fiscal responsibilities. (iii) Only certain variables are available each year. The four possible categories are: taxes collected by counties, Tax_1 ; taxes collected by all local governments—county, minor civil divisions, school districts, etc.—within the county, Tax_2 ; revenues—including inter-governmental grants but excluding inter-local government grants—collected by counties, Rev_1 ; and revenues collected by all local governments within the county, Rev_2 . The available data are: 1870-80: Tax_1 , Tax_2 ; 1890-1913: Rev_1 (“gross receipts”); 1922: Tax_1 ; 1932: Rev_1 ; 1942: Tax_1 , Rev_1 ; 1957: Rev_2 ; 1962-92: Tax_2 , Rev_2 .

Preference Proxies

Party vote shares in presidential elections

Frequency: every 4 years. Sources: ICPSR studies 8611 (1850-1972) and 13 (1976-88). Each election includes vote shares for Democrats and Republicans (except 1852 when Whigs are included) and significant minor parties (Socialists are included in 1912-32). In addition, each election included an “all others” category to ensure the vote shares sum to 1. For certain observations in ICPSR study 8611, the sum of the party vote shares is less than or greater than 1 (these discrepancies were confirmed by Erik Austin, the data archive director at ICPSR). In such cases the vote share are re-allocated in proportion to their stated values with the resulting shares now summing to 1.

Black population share

Frequency: every 10 years. Sources: ICPSR studies 3 (1850-1970), 9693 (1980) and U.S. Census Bureau, 1994b (1990). The data for 1850-60 are the sum of the population categories “free colored” and “slave.” The data for 1870-80 are for the population category “colored.” The data for 1890-1970 are for the population category “negro.” The data for 1980-90 are for the population category “black.”

Religion shares

Frequency: every 10 years for 1850-70, 1906-36, 1980-90. Additional observations in 1890, 1952, 1971. Sources: ICPSR studies 3 (1850-70, 1890), 8 (1906-52), and Roper Center (1952, 1971, 1980, 1990). The data for 1850-70 is for the number of accommodations or sittings; the data for 1890 is for both the number of seatings and members; the data for the remaining years is for the number of members. No data was collected in 1880. The data for 1952 is from both ICPSR study 8 and Roper Center. Because the particular groups included in the files change over time (due largely to schisms and mergers), the data was aggregated into the following 27 denominational families: Adventist, Baptist, Black Baptist, Catholic, Communal, Congregationalist, Disciples of Christ, Eastern Liturgical, Episcopalian, Friends, Fundamentalist, Holiness, Jewish, Lutheran, Mennonite, Metaphysical, Methodist, Black Methodist, Moravian, Mormon, New Age, Pentecostal, Presbyterian, Reformed, Unitarian-Universalist, Members of other bodies, and Non-classified. This classification scheme is based on the list in Melton (1978). Some denominations are not present in all years. All values are normalized by the county population. In one of our classification schemes we also generate an “unaffiliated” or non-sampled group to ensure the groups sum to 1. In certain years the sum of the group shares exceed 1 due to the manner in which the data is collected (these cases were confirmed by consulting the hard-copy versions of the various

studies). When this happens, the groups are re-allocated in proportion to their stated values with the resulting shares now summing to 1. We uncovered some errors in the data which we corrected by consulting hard-copy versions of the statistics. A full list of these discrepancies as well as details of our classification scheme are available upon request.

Young population share

Frequency: every 10 years. Sources: ICPSR studies 3 (1850-1950, 1970), 7736 (1960), 9693 (1980), U.S. Bureau of the Census, 1994b (1990). Generally, the young are defined as those 5-20 years old. Because of data coding, this had to be modified in several years. 1850-60: young = 5-19 years old; 1870: young = 5-18 years old; 1880: young = 5-17 years old; 1910: young = 6-20 years old; 1920-30: young = 7-20 years old; 1980: young = 6-17 years old + 0.5 18-24 years old.

Old population share

Frequency: every 10 years (with gaps). Sources: ICPSR studies 3 (1850-60), U.S. Bureau of the Census, 1932 (1930), 7736 (1950-70), U.S. Bureau of the Census, 1998b (1980-90). The old are defined as those at least 65 years old. In 1850 and 1860, half of the individuals in the age category 60-69 year olds were counted as old as well as all individuals 70 years and older. No electronic data exists for this age group between 1870 and 1940.

Marriage rate

Frequency: every 10 years (with gaps). Sources: ICPSR study 3 (1930, 1950), U.S. Bureau of the Census, 1964 (1960), U.S. Bureau of the Census, 1973 (1970), U.S. Bureau of the Census, 1998b (1980-90). No marriage data are available for 1940. The Census defines these groups for those who are at least 14 years old (in 1950-70) and for those who are at least 15 years old (in 1930, 1980-90). These age categories are used to normalize the marriage rate and also serve as the population weight in the heterogeneity measures. The data for 1970 only include males.

Owner-Occupied Homes

Frequency: every 10 years. Sources: U.S. Bureau of the Census, 1933 (1930), ICPSR studies 3 (1940), 7736 (1950-70), U.S. Bureau of the Census, 1998b (1980-90). This variable is defined as the percentage of all homes which are owner occupied. Housing data for 1930 is unavailable in tape form. Instead we use values for the variable "families" listed in the hardcopies (the prefatory remarks of the 1930 Census reads: "Since a home is defined as the living quarters occupied by a family, the number of homes is always the same as the number of families").

Education

Frequency: every 10 years. Sources: ICPSR study 3 (1940-50), U.S. Bureau of the Census, 1964 (1960), ICPSR study 7736 (1970), U.S. Bureau of the Census, 1998b (1980-90). Our categories are those with only a high school degree and those with a college degree or more. The Census defines these groups for those who are at least 25 years old (this age category is used to normalize the education groups and also serves as

the population weight in the heterogeneity measures). No education data are available prior to 1940.

Foreign-born population share

Frequency: every 10 years. Sources: ICPSR studies 3 (1850-1950), 7736 (1960), 9694 (1970), 9693 (1980), U.S. Bureau of the Census, 1994b (1990). The data for 1850-60 include both free and slave population. The data for 1910-30 are just for white foreign-born (these values are normalized to the total white population); in 1900 and 1940 the white and overall foreign-born shares are comparable. For 1950 the data is just for foreign-born aged 21 years or older (these values are normalized to the total population aged 21 years or older).

Income Distribution

Frequency: 1949, 1979, 1989. Sources: U.S. Bureau of the Census, 1952 (1949), U.S. Bureau of the Census, 1998 (1979, 1989). Observation unit: we separately consider both families and households (“families and unrelated individuals” in the 1949 Census). Data note: the distribution is reported as the number of families/households whose income falls in a given income interval. There are 14 listed income groups for 1949, 17 income groups in 1979 and 25 groups in 1989. The Census did not begin collecting income data until 1940 and detailed, county-level income distribution data is not available in electronic form prior to 1970 (records in the *County and City Data Book* generally include only 3 income groups). We collected a random, 1-in-10 sample ($N=311$) for 1949 from hardcopies in U.S. Bureau of the Census, 1952. To maintain comparability, the same 1-in-10 sample counties are used in the later years.

Boston SMSA Data

All data are for the 92 municipalities in the 1980 definition of the Boston SMSA.

Party vote shares in presidential elections

Frequency: every 4 years. Availability: 1868, 1884-1988. Source: Massachusetts General Court (various years).

Current expenditure

Frequency: Approximately every 10 years (1906-55); every five years (1962-92). Sources: Massachusetts Department of Corporations and Taxation, various years (1906-1955), U.S. Bureau of the Census, 1963 and 1968 (1962, 1967), ICPSR studies 69, 8118, 8394, 9484 (1972-87), U.S. Bureau of the Census, 1994 (1992). Category definitions: Total is “total maintenance spending” (1906-55) which is basically current expenditures, “total spending other than capital outlay” (1962, 1967), the sum of all current operations categories (1972-92); Protection is “protection of life and property” current expenses (1906-42), “public safety” current expenses (1955), police plus fire protection current expenses (1962-92); Education is “education” current expenses (1906, 1913), “schools” current expenses (1923-55), “education other than capital outlay” (1962-67), “local schools current operations” (1972-82), “elementary/secondary education current

operations” (1987, 1992). The 1962 and 1967 data are supplemented with observations in Massachusetts Department of Corporations and Taxation (various years).

Demographics (population, race, nativity, age)

Frequency: Approximately every 10 years (1875-1915, 1930-90). Sources: Massachusetts, Bureau of Statistics and Labor, 1877, 1888, 1900, 1910, 1918 (1875-1915), U.S. Bureau of the Census, 1932, 1943, 1952b, 1962, 1973, 1982, 1992 (1930-90). Note that the 1950 and 1960 data are aggregated from tract-level statistics. Age data does not include 1885 or 1905. See the discussion for the county-level, national sample for a fuller description of the demographic variables.

Income

Frequency: 1949 and 1989. Sources: U.S. Bureau of the Census, 1952b (1949), U.S. Bureau of the Census, 1992 (1989). Observation unit: households (“families and unrelated individuals” in the 1949 Census). Data note: the distribution is reported as the number of households whose income falls in a given income interval. There are 14 listed income groups for 1949 and 25 groups in 1989.

Data References

All ICPSR files are available at www.icpsr.umich.edu/. Census and Massachusetts documents with no on-line or CD-ROM reference were used to manually input data.

General Social Survey (1999). *GSS 1972-1996 Cumulative Datafile*. Survey Documentation & Analysis Archive. csa.berkeley.edu:7502/.

ICPSR study 3 (1992). *Historical, Demographic, Economic, and Social Data: The United States, 1790-1970*. Investigator: Inter-university Consortium for Political and Social Research.

ICPSR study 8 (1992). *Censuses of Religious Bodies, 1906-1936 and NCC 1952*. Investigator: Inter-university Consortium for Political and Social Research.

ICPSR study 13 (1995). *General Election Data for the United States, 1950-1990*. Investigator: Inter-university Consortium for Political and Social Research.

ICPSR study 69 (1975). *Census of Governments 1972: Finance Files*. Investigator: U.S. Department of Commerce. Bureau of the Census.

ICPSR studies 7736, 8256, 9251 (various years). *County and City Data Book*. Investigator: U.S. Department of Commerce. Bureau of the Census.

ICPSR study 8118, 8394, 9484 (various years). *Census of Governments: Finance [Summary] Statistics*. Investigator: United States Department of Commerce. Bureau of the Census.

- ICPSR study 8611 (1992). *Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972*. Investigators: Jerome M. Clubb, William H. Flanigan, and Nancy H. Zingale.
- ICPSR studies 9693, 9694 (1992). *Census of Population and Housing, 1970/1980 [United States]: Extract Data*. Investigator: Terry K. Adams.
- IPUMS. *Integrated Public Use Microdata Series*. www.ipums.umn.edu.
- Massachusetts, Bureau of Statistics and Labor (1877). *Census of Massachusetts: 1875*. Vol. I. *Population and Social Statistics*. Boston, Wright and Potter.
- Massachusetts, Bureau of Statistics and Labor (1888). *Census of Massachusetts: 1885*. Vol. I. *Population and Social Statistics*. Boston, Wright and Potter.
- Massachusetts, Bureau of Statistics and Labor (1900). *Census of the Commonwealth of Massachusetts: 1895*. Vol. IV. *Population and Social Statistics*. Boston, Wright and Potter.
- Massachusetts, Bureau of Statistics and Labor (1910). *Census of the Commonwealth of Massachusetts: 1905*. Vol. I. *Population and Social Statistics*. Boston, Wright and Potter.
- Massachusetts, Bureau of Statistics and Labor (1918). *Decennial Census, 1915*. Vol. IV. *Population and Social Statistics*. Boston, Wright and Potter.
- Massachusetts, Department of Corporations and Taxation (various years). *Report on the Statistics of Municipal Finances*. Commonwealth of Massachusetts, Boston.
- Massachusetts, General Court (various years). *Manual for the General Court*. Commonwealth of Massachusetts, Boston.
- Melton, J. Gordon (1978). *The Encyclopedia of American Religions*. McGrath Publishing Company, Wilmington NC.
- Roper Center (various years). *Churches and church membership in the United States*. Unpublished data tape at the Social Science Data Center of the University of Connecticut, Storrs, CN.
- U.S. Bureau of the Census (1908). *Wealth, Debt, and Taxation, 1902*. Washington DC, US GPO.
- U.S. Bureau of the Census (1915). *Wealth, Public Debt, and Taxation: 1913*, Vol. II, Part VII. *County Revenues, Expenditures, and Public Properties*. Washington DC, US GPO.

- U.S. Bureau of the Census (1924). *Wealth, Public Debt, and Taxation: 1922, Taxes Collected*. Washington DC, US GPO.
- U.S. Bureau of the Census (1932). *Fifteenth Census of the United States: 1930. Population*, Vol. III, Pts 1-2. *Report by States*. Washington DC, GPO.
- U.S. Bureau of the Census (1933). *Fifteenth Census of Population*. Vol. VI. *Families*. Washington DC, GPO.
- U.S. Bureau of the Census (1935). *Financial Statistics of State and Local Governments: 1932*. Washington DC, US GPO.
- U.S. Bureau of the Census (1943). *Sixteenth Census of the United States: 1940. Population*, Vol. II, *Characteristics of the Population*. Part 3: *Alabama-Michigan*. Washington DC, GPO.
- U.S. Bureau of the Census (1944). *Census of Governments, County Finances: 1942 Compendium*. Washington DC, US GPO.
- U.S. Bureau of the Census (1948). *Historical Review of State and Local Finances. State and Local Government Special Studies No. 25*. Washington DC, US GPO.
- U.S. Bureau of the Census (1952). *Census of Population, 1950: Characteristics of the Population*. Washington DC, US GPO.
- U.S. Bureau of the Census (1952b). *U.S. Census of Population: 1950* Vol. III, *Census Tract Statistics*, Chapter 6: *Boston. Bulletin P-D6*. Washington DC, GPO.
- U.S. Bureau of the Census (1962). *U.S. Census of Population and Housing: 1960. Census Tracts: Boston, Mass. Final Report PHC(1)-18*. Washington DC, GPO.
- U.S. Bureau of the Census (1963). *1962 Census of Government*. Washington DC, GPO.
- U.S. Bureau of the Census (1964). *Census of Population, 1960*. Vol. I: *Characteristics of the Population*. Washington DC, GPO.
- U.S. Bureau of the Census (1968). *1967 Census of Government*. Washington DC, GPO.
- U.S. Bureau of the Census (1973). *1970 Census of Population*, Vol. I, *Characteristics of the Population*. Washington DC, GPO.
- U.S. Bureau of the Census (1982). *1980 Census of Population*, Vol. I, *Characteristics of the Population*. Chapter B: *General Population Characteristics*. Part 23: *Massachusetts*. Washington DC, GPO.

- U.S. Bureau of the Census (1992). *Census of Population and Housing, 1990: Summary Tape File 3A*. <http://venus.census.gov/cdrom/lookup/>.
- U.S. Bureau of the Census (1994). *1992 Census of Governments, CD-ROM*. Washington DC, GPO.
- U.S. Bureau of the Census (1994b). *County and City Data Book: 1994*. Washington DC, US GPO. fisher.lib.Virginia.EDU/ccdb/.
- U.S. Bureau of the Census (1997). *Census of Governments, Vol. 4, no 3: Finances of County Governments*. www.census.gov/prod/www/abs/gc92fin.html.
- U.S. Bureau of the Census (1998). *USA Counties 1998*. CD-ROM. Washington DC, US GPO.
- U.S. Census Office (1895). *Eleventh Census: 1890. Wealth, Debt, and Taxation, Part II, Valuation and Taxation*. Washington DC, US GPO.

Table 1: Dispersion of Per Capita Education Spending Across Counties

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

Table 2: Dispersion of Per Capita Taxes and Revenues Across Counties

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

Definitions: $CV = \text{Std Dev}(\mathbf{G})/\text{Mean}(\mathbf{G})$ where \mathbf{G} is the vector of per capita spending and where each county receives a population-weight. $DG = 0.5G^{-1}\sum_j P_j|G_j - G|$ where G_j is per capita spending in county j , G is mean per capita spending for all counties, and P_j is the share of total population in county j . For Table 1, the G 's involve education spending. For Table 2, the G 's are taxes or revenues ($G_1 =$ just county government G ; $G_2 =$ county + sub-county government G). Empty cells are due to missing data. See the Data Appendix for further details about the data.

Table 3: 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
	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

See Section V.B. for definitions of the indices. As described in the Data Appendix, these values are based on a random, 1-in-10 sample.

“W” is within, and “B” is between. The within-measures are population weighted averages of each of the county indices. Standard deviations are reported in parentheses below the within-measures.

To compute the indices, we need to know for each county 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: Municipal Spending (Current Operations) in Boston SMSA
Coefficients of Variations**

Year	Municipal-Level			Aggregated to County-Level				
	N	G _{Total}	G _{Protection}	G _{Education}	N	G _{Total}	G _{Protection}	G _{Education}
1906	92	0.359	0.490	0.197	5	0.289	0.402	0.107
1913	92	0.297	0.402	0.218	5	0.202	0.291	0.136
1923	92	0.267	0.351	0.203	5	0.130	0.207	0.081
1932	92	0.281	0.346	0.181	5	0.202	0.204	0.102
1942	92	0.220	0.297	0.160	5	0.138	0.159	0.082
1955	92	0.236	0.350	0.190	5	0.174	0.267	0.104
1962	59	0.189	0.311	0.311	5	0.139	0.241	0.189
1967	59	0.216	0.281	0.356	5	0.173	0.208	0.198
1972	92	0.293	0.399	0.267	5	0.170	0.333	0.115
1977	92	0.242	0.365	0.177	5	0.152	0.266	0.034
1982	92	0.233	0.276	0.206	5	0.106	0.148	0.046
1987	92	0.243	0.334	0.187	5	0.133	0.254	0.042
1992	92	0.267	0.340	0.194	5	0.150	0.270	0.052

See the Data Appendix for a list of sources and definitions of these series. 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).

Figure 1: Transportation Costs

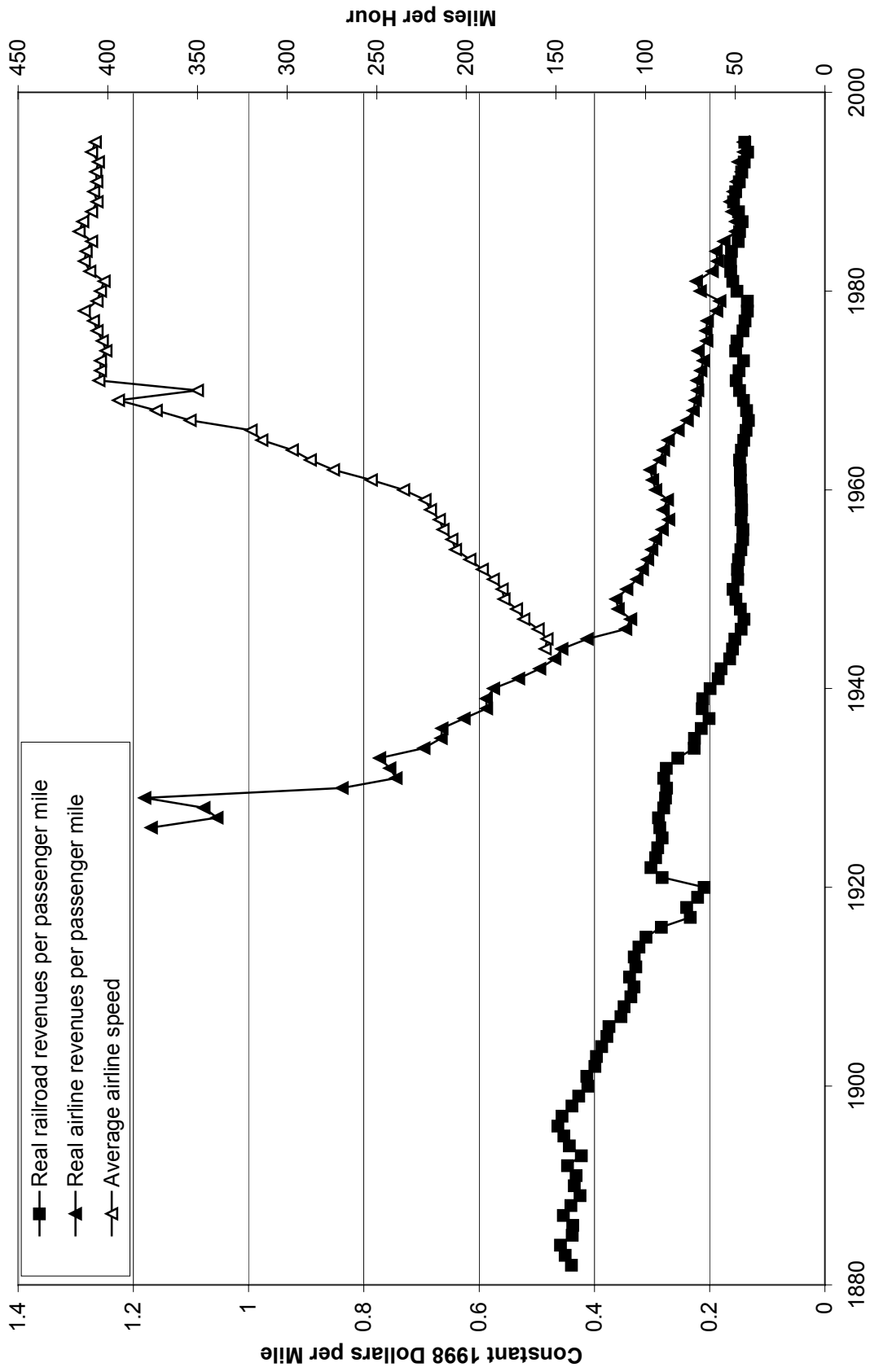


Figure 2: Real Cost of a Three-Minute Phone Call

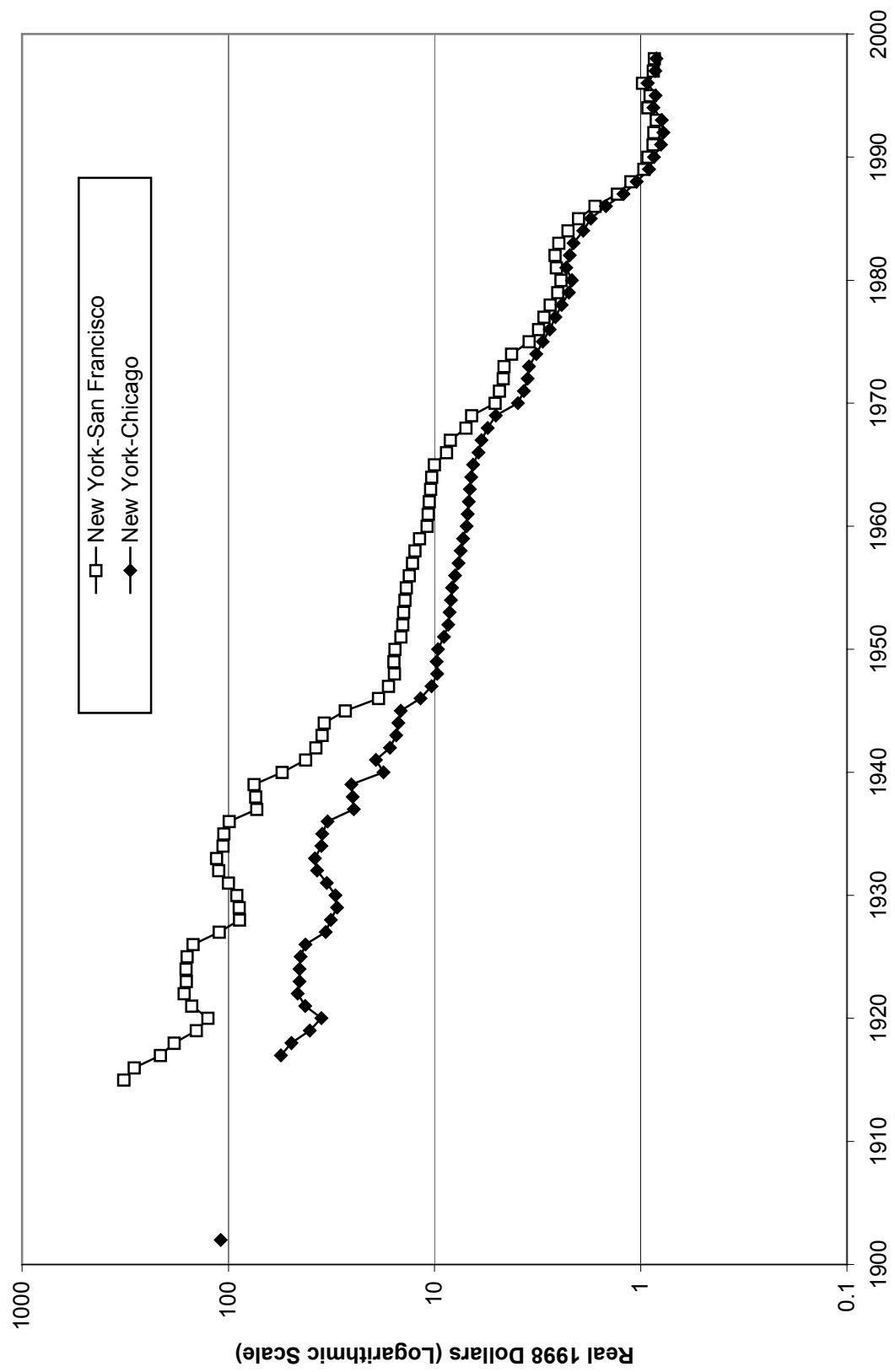
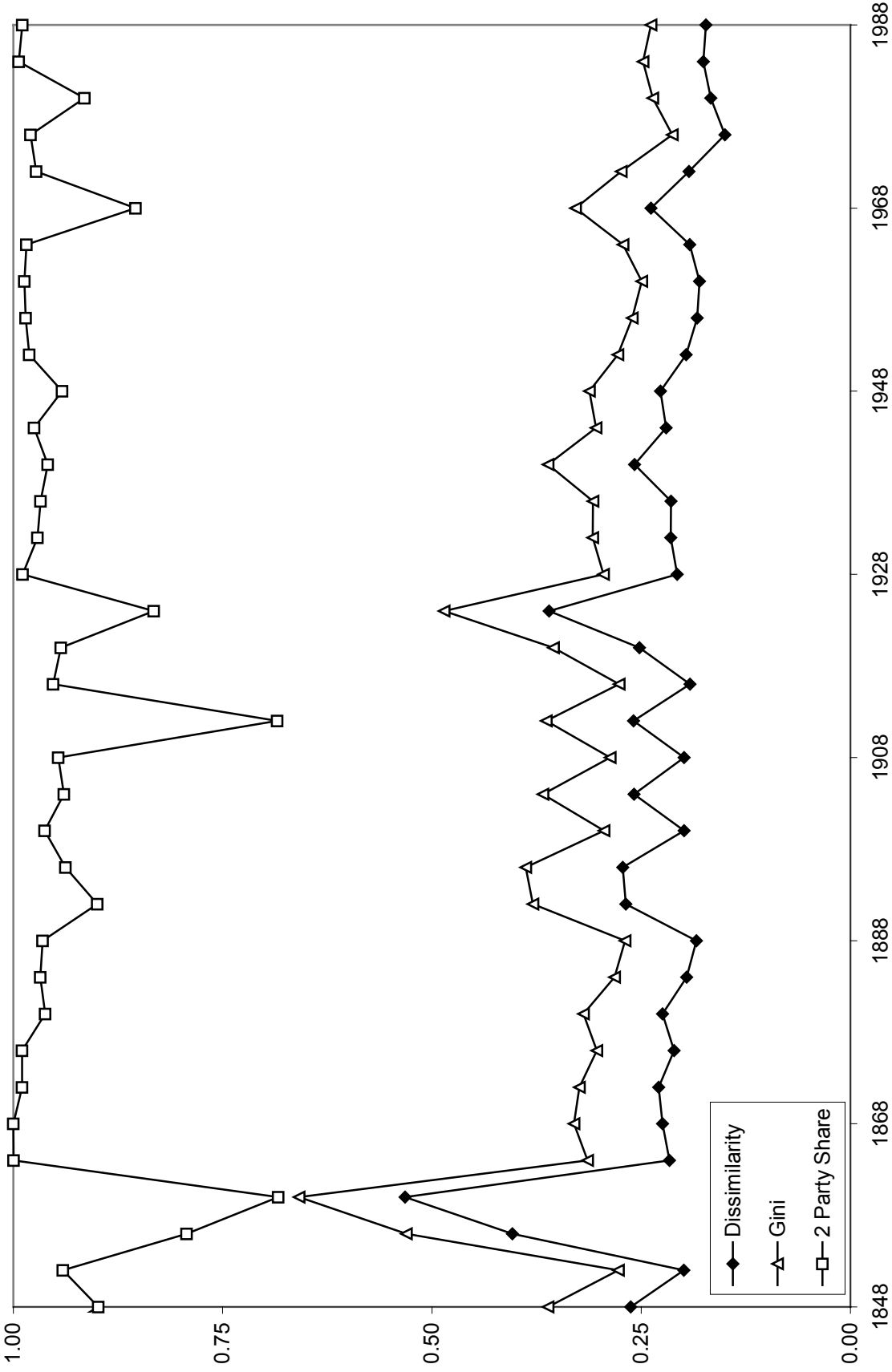
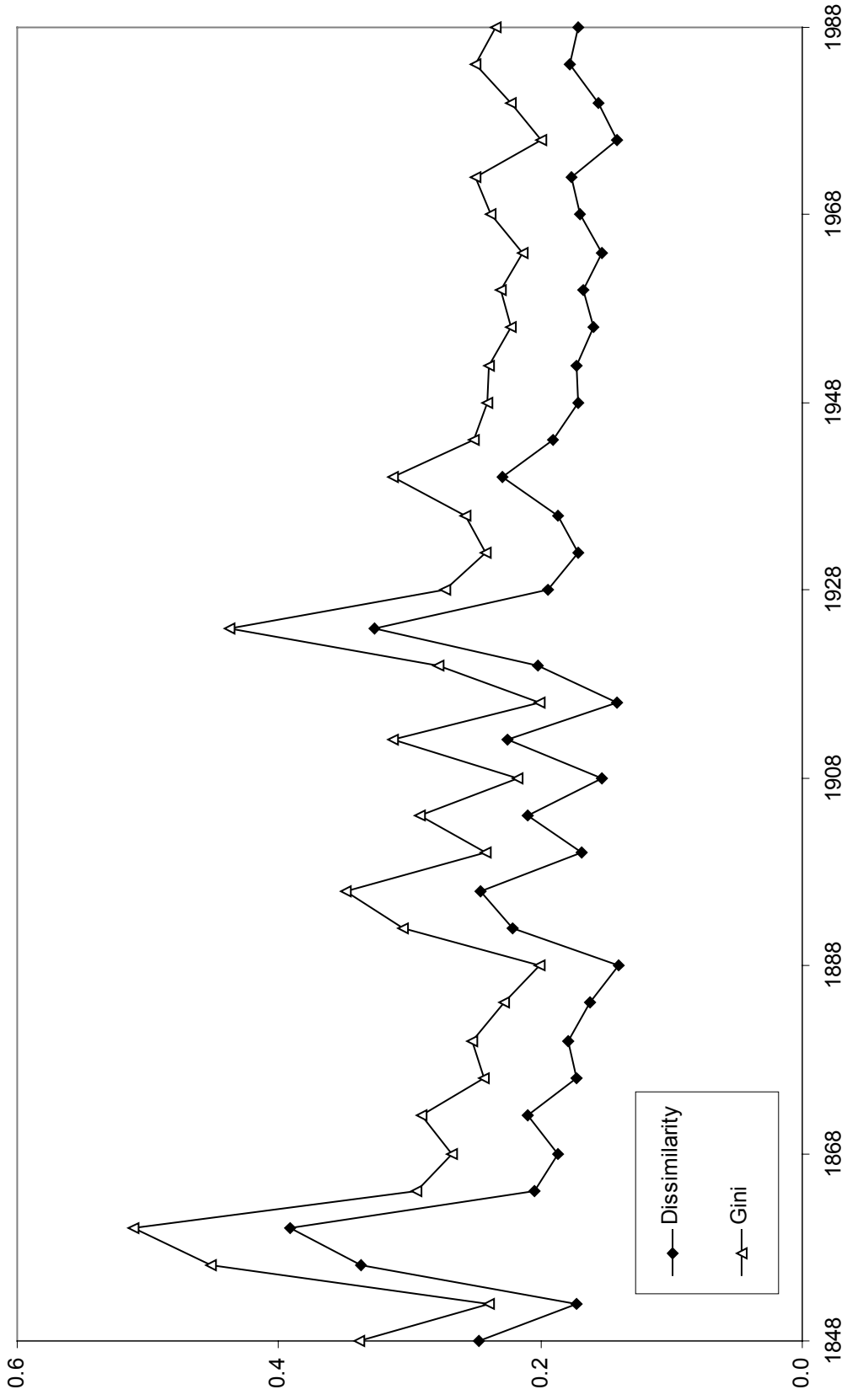


Figure 3: Presidential Vote Shares



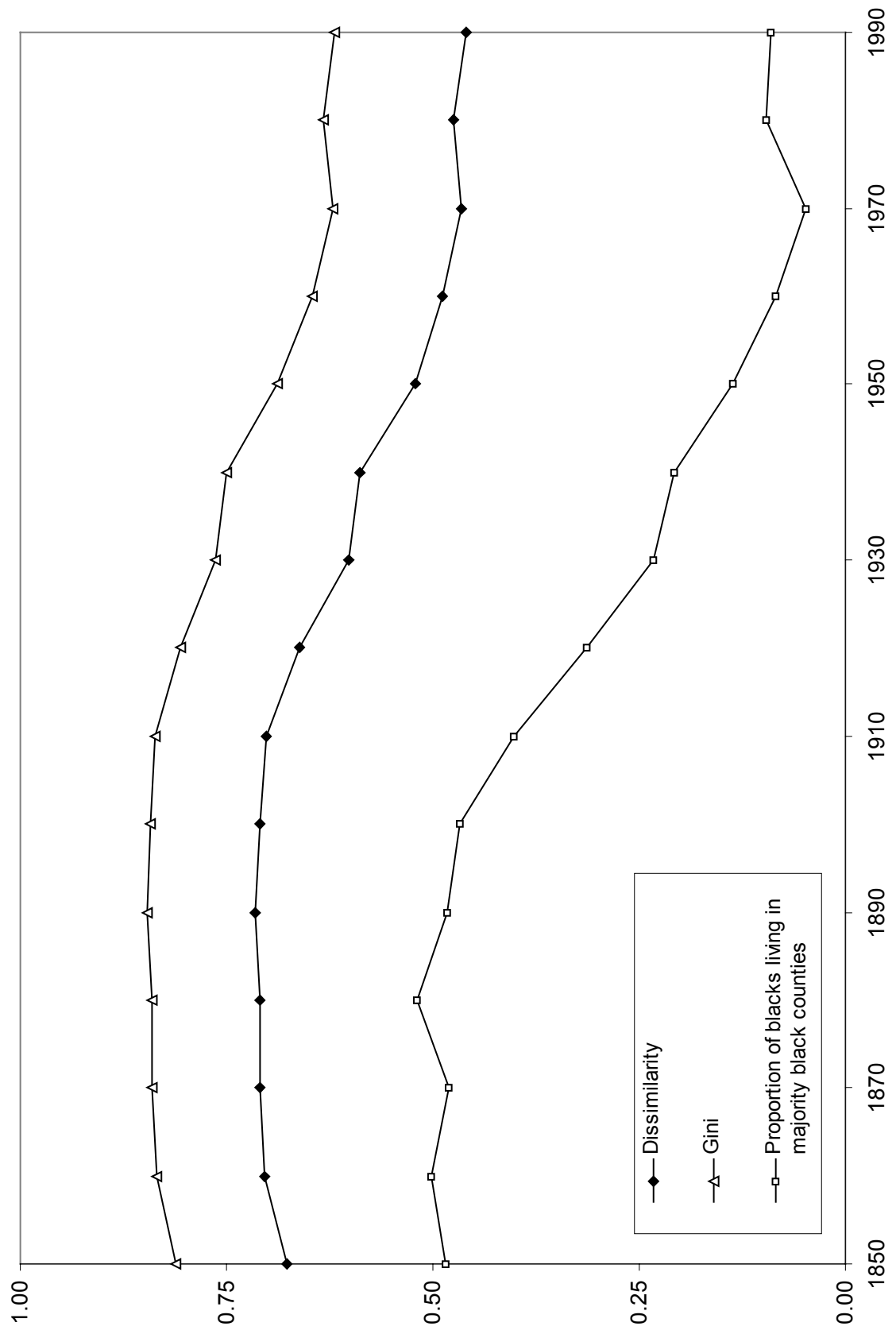
Year (Sample Size): 1852 (1551), 1860 (1864), 1872 (2177), 1880 (2315), 1892 (2667), 1900 (2730), 1912 (2970), 1920 (3031), 1932 (3091), 1940 (3067), 1952 (3097), 1960 (3101), 1972 (3105), 1980 (3111), 1988 (3113)

Figure 4: Presidential Vote Shares (excluding the South)



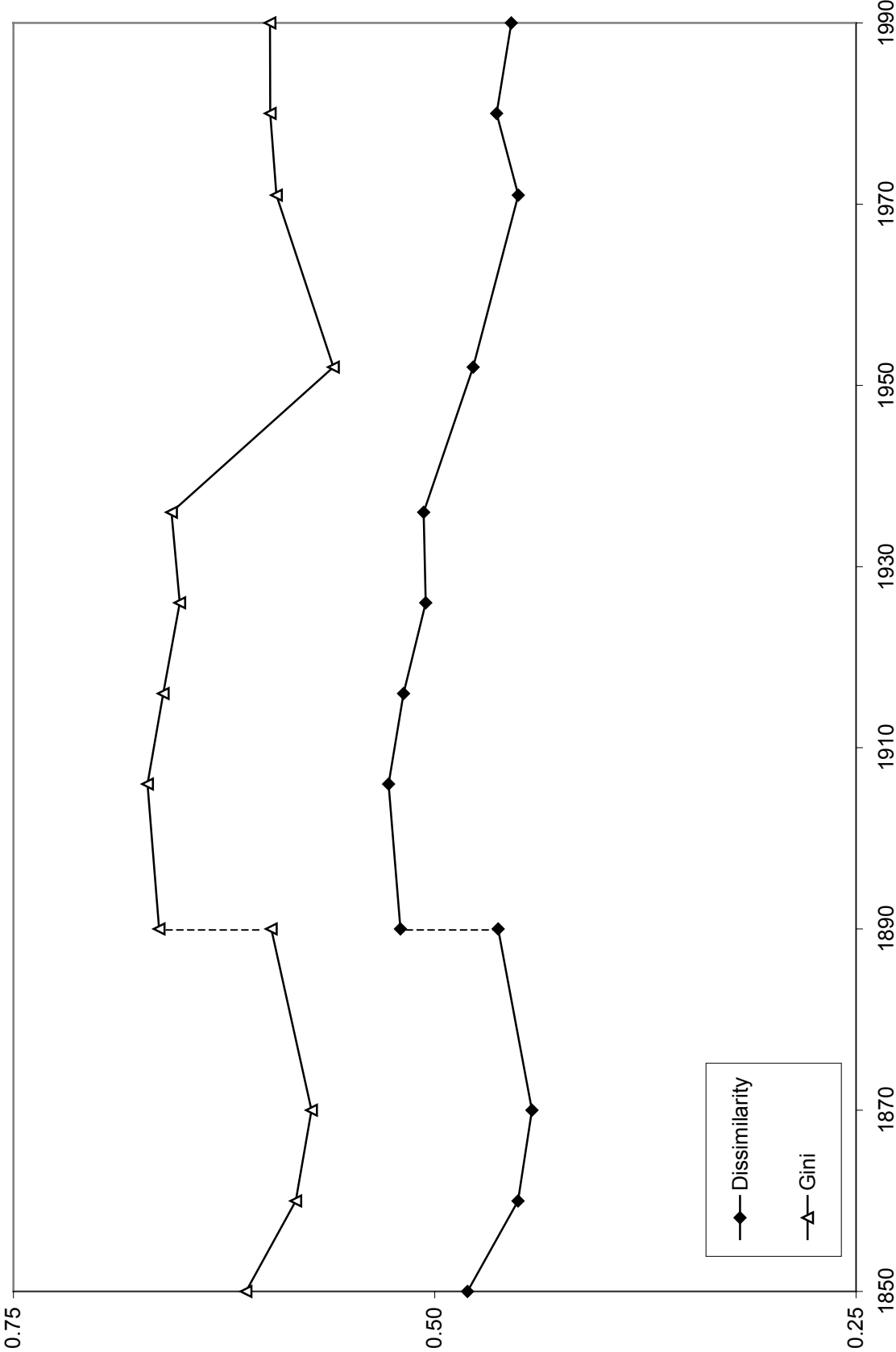
The South: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia

Figure 5: Black Population Share



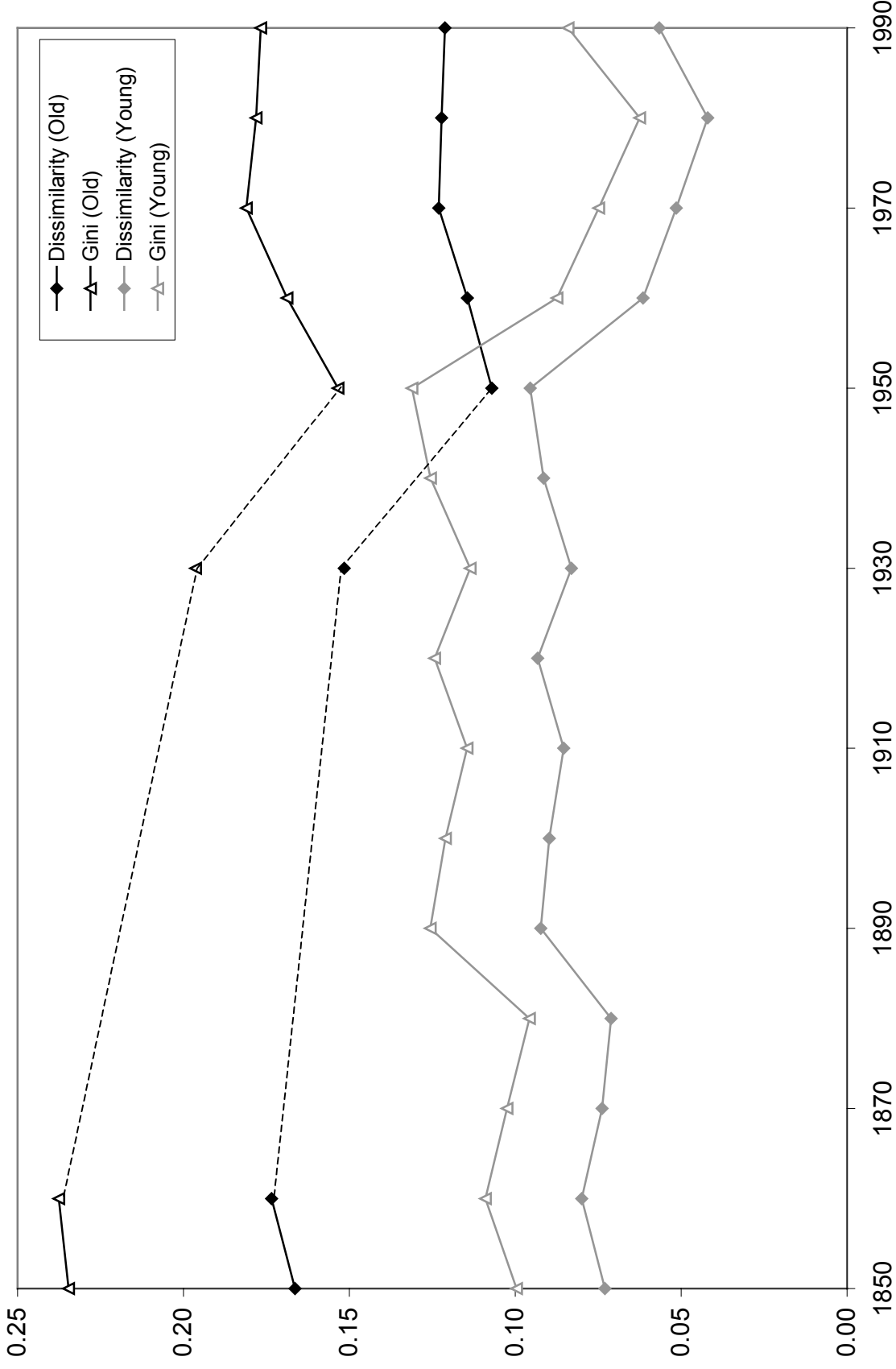
Year (Sample Size): 1850 (1596), 1860 (2030), 1870 (2185), 1880 (2400), 1890 (2743), 1900 (2777), 1910 (2950), 1920 (3064), 1930 (3100), 1940 (3097),

Figure 6: Religion Denominational Shares



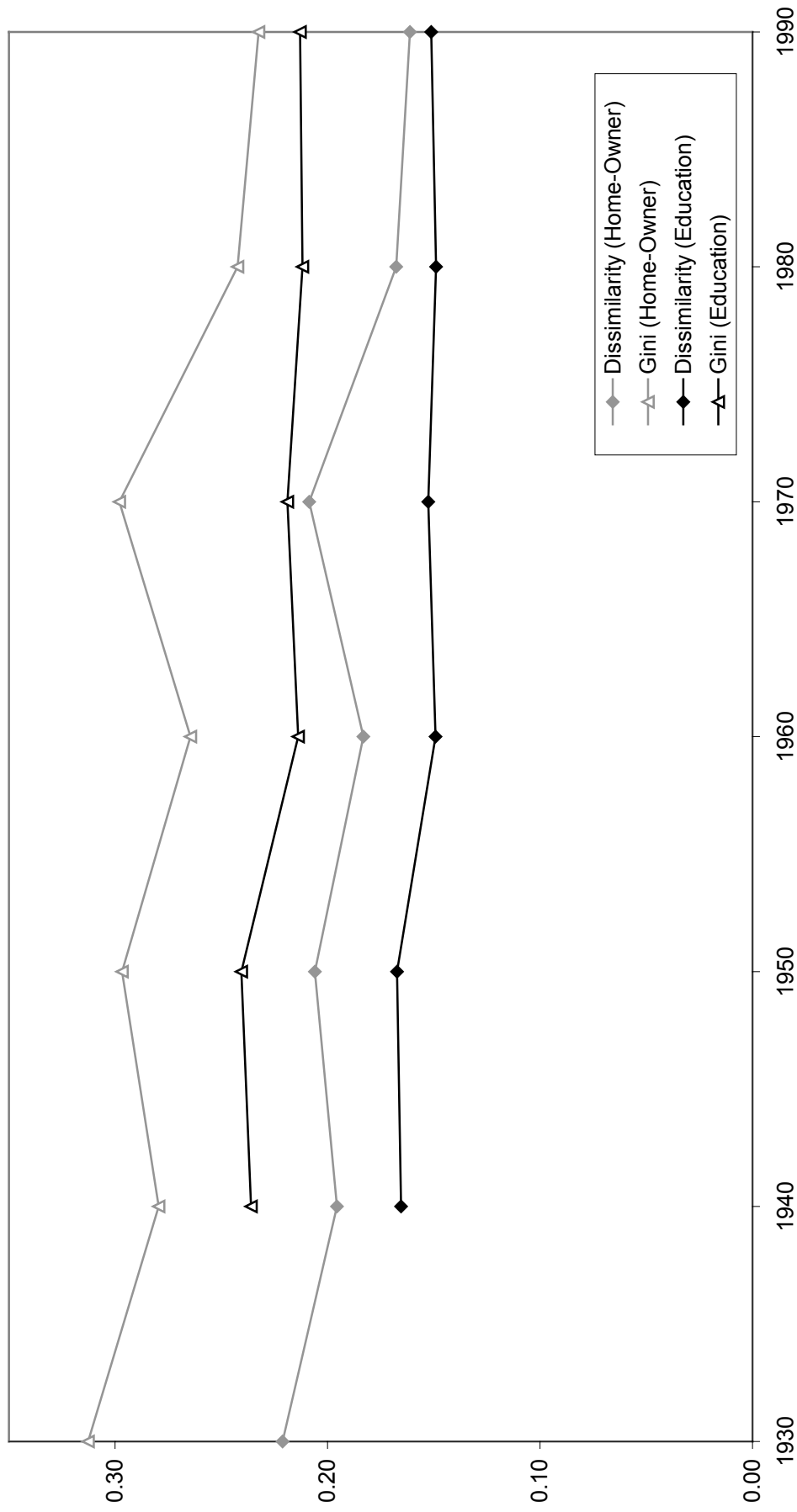
Note: 1850-1890 data is based on *seatings* while 1890-1990 is based on *members* (both sources are available for 1890). Unaffiliated and affiliates of minor denominations are excluded.

Figure 7: Young and Old Population Shares



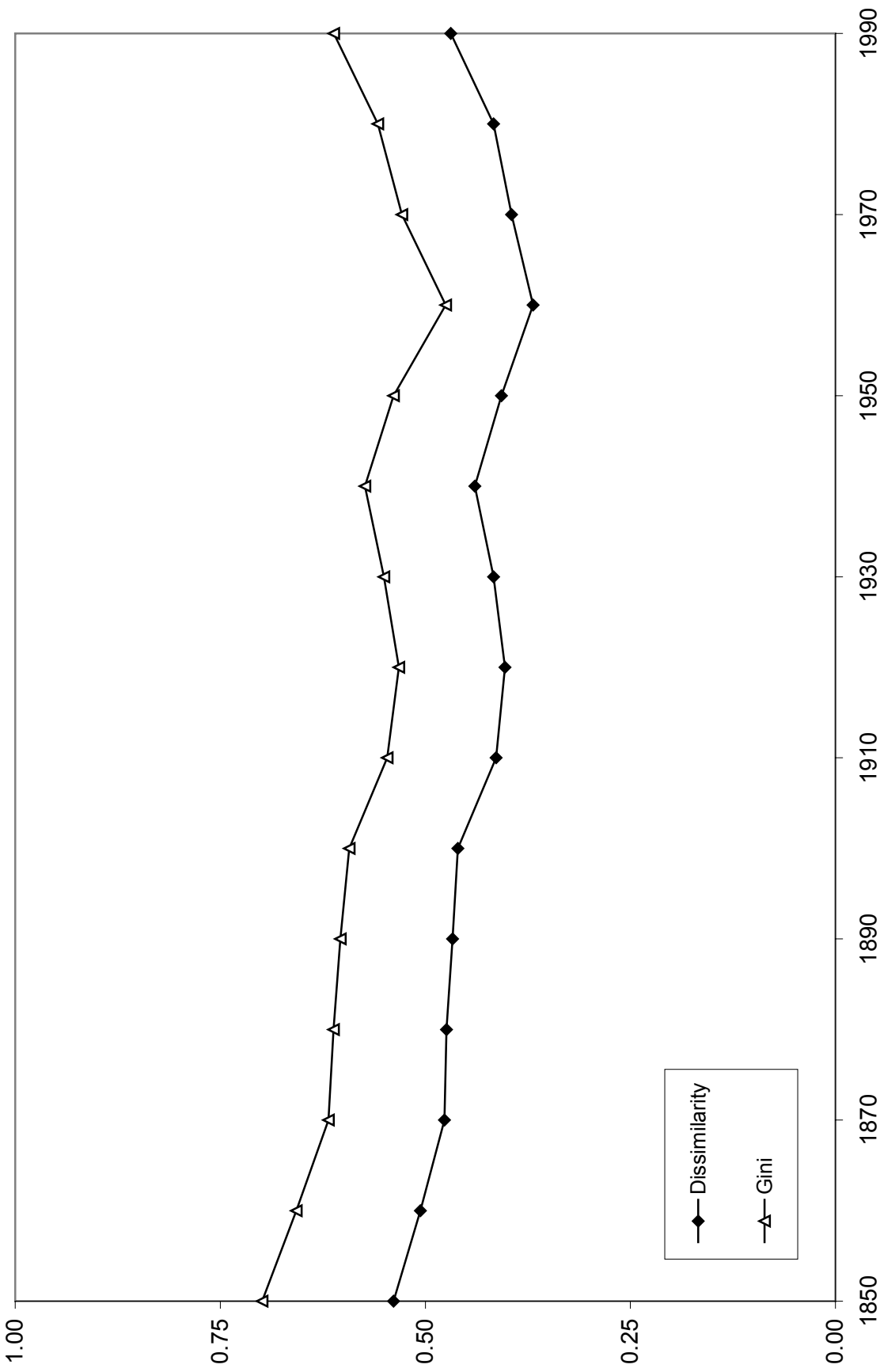
Year (Sample Size): 1850 (1607), 1860 (2055), 1870 (2230), 1880 (2421), 1890 (2780), 1900 (2832), 1910 (2955), 1920 (3071), 1930 (3102), 1940 (3099), 1950 (3102), 1960 (3114), 1970 (3112), 1980 (3115), 1990 (3117)

Figure 8: Home-Owner Occupation and Education Levels



Year (Sample Size): 1930 (3100), 1940 (3097), 1950 (3099), 1960 (3103), 1970 (3109), 1980 (3108), 1990 (3110)

Figure 9: Foreign-Born Population Share



1850, 1860: includes slave population; 1910-1930: just white population; 1950: just >= 21 years old. See "Young and Old Population Share" figure for sample size.

**Figure 10: Presidential Vote Shares in Boston SMSA
(Municipal and County Level)**

