

Political Jurisdictions in Heterogeneous Communities

Alberto Alesina

Harvard University, National Bureau of Economic Research, and Centre for Economic Policy Research

Reza Baqir

International Monetary Fund

Caroline Hoxby

Harvard University and National Bureau of Economic Research

We investigate whether political jurisdictions form in response to the trade-off between economies of scale and the costs of a heterogeneous population. We consider heterogeneity in income, race, ethnicity, and religion, and we test the model using American school districts, school attendance areas, municipalities, and special districts. We find strong evidence of a trade-off between economies of scale and racial heterogeneity; we also find evidence of a trade-off between economies of scale and income heterogeneity. Conversely, we find little evidence that ethnic or religious heterogeneity shapes jurisdictions. To clarify the direction of causality between heterogeneity and jurisdictions, we exploit shocks to racial heterogeneity generated by the two world wars.

We thank Steve Levitt, Antonio Rangel, David Romer, two anonymous referees, and participants in seminars at Massachusetts Institute of Technology, Stanford, Harvard, Yale, the University of Stockholm, the World Bank, the National Bureau of Economic Research, and the MacArthur Network on Social Interactions and Inequality for very useful comments. Alesina gratefully acknowledges financial support from the National Science Foundation. Hoxby gratefully acknowledges financial support from the MacArthur Foundation Social Interactions and Economic Inequality Network.

[*Journal of Political Economy*, 2004, vol. 112, no. 2]
© 2004 by The University of Chicago. All rights reserved. 0022-3808/2004/11202-0001\$10.00

I. Introduction

The largest country in the world, China, has 1.284 billion inhabitants; the smallest, Palau, has 19,409.¹ The largest county in the United States (Los Angeles, Calif.) has 9,519,338 inhabitants; the smallest (Loving, Texas) has 67.² The largest school district in the United States has 1.1 million school-aged children; the smallest has one.³ What determines the size of political jurisdictions?

A conventional answer to this question is historical contingency, a catchall for leaders, wars, migrations, and many other circumstances. While historical contingency is certainly important, in this paper we argue that there is a fundamental trade-off that shapes jurisdictions. This is the trade-off between the benefits of larger scale and the costs of a more heterogeneous population. Heterogeneity can be costly if different individuals have different policy preferences, so that they must compromise in order to share a jurisdiction. Heterogeneity is also costly if individuals prefer to interact with people like themselves, regardless of preferences over public policies.

Consider, for example, the problem of dividing an area into school districts. Large districts have economies of scale because they can provide libraries, sports facilities, and administration on a districtwide basis. On the other hand, in large districts, many families have to mix their children and agree on common educational policies. If families in an area are homogeneous, an increase in size may be purely beneficial (unless there is a point at which diseconomies of scale set in). If, instead, an increase in size implies an increase in heterogeneity, there may be a trade-off.

In this paper, we test whether such a trade-off exists using data on local jurisdictions in the United States. While we believe that this trade-off may exist for many political jurisdictions including countries, focusing on local jurisdictions is instructive. Local jurisdictions such as school districts and municipalities are interesting in their own right because they affect social outcomes and policies. Also, local jurisdictions are more numerous and change more frequently than larger jurisdic-

¹ Source: Central Intelligence Agency (2003). The data refer to 2002. Palau is the smallest country with a full membership in the United Nations. Three others are smaller but do not have a U.N. seat.

² Source: 2000 Census of Population. Yellowstone National Park (which is a county) actually has the smallest population of any county in the United States, but its population is artificially limited.

³ The largest districts in the United States are the New York City School District (which includes all five boroughs), the Los Angeles Unified District, and the Chicago School District. There are approximately 50 school districts in the United States that have between one and three students in a typical year. They include districts such as Maine's Magalloway District and Montana's Two Dot District. These enrollment numbers are taken from the U.S. Department of Education (2003).

tions, providing better opportunities for empirical tests. Finally, different types of local jurisdictions vary in the degree to which people must interact and make joint decisions: this helps us distinguish between the two reasons why heterogeneity may be costly.

Specifically, we test whether a trade-off between economies of scale and heterogeneity helps to explain the number and size of school districts, municipalities, special districts, and school attendance areas in the United States. We use counties as our basic areas because they almost never consolidate or break apart and local jurisdictions rarely traverse county lines. We concentrate on heterogeneity in income, race, ethnicity, and religion. While other types of heterogeneity, such as age heterogeneity, may be relevant (and, in fact, we include them in our estimation), we emphasize the results for the aforementioned types of heterogeneity. The reason is that a vast sociological and political literature argues that they are the main fault lines of preferences and political conflict in the United States (see, e.g., Huckfeldt and Kohfeldt 1989; Hacker 1992; Wilson 1996).

Our results suggest that people are willing to give up economies of scale in order to avoid being in a jurisdiction with significant racial or income heterogeneity. The trade-off between economies of scale and racial heterogeneity tends to be larger in magnitude and more robust empirically than the trade-off between economies of scale and income heterogeneity. This result is striking because the one *obvious* reason for people to care about the population in their jurisdiction is that people with different incomes face different tax burdens but receive about the same level of local public goods.

Readers may find it hard to envision how local jurisdictions respond to heterogeneity because they can recall few, if any, jurisdictions being created in their area. However, refusal to consolidate is the main mechanism by which jurisdictions respond. Through consolidation, the number of local jurisdictions in the United States fell 12-fold between 1900 and 2000. In most states, a heterogeneous area could end the twentieth century with a large number of jurisdictions simply by refusing to consolidate. In addition, the creation of new jurisdictions does account for some response.

Our paper is related to four strands of literature. One is recent work on the endogenous formation of political jurisdictions. In particular, Alesina and Spolaore (1997) argue that the trade-off between size and heterogeneity is an important explanation of the number and size of nations; Bolton and Roland (1997) focus instead on income differences. The second strand of literature studies the effects of racial heterogeneity on local policies, particularly policies that involve redistribution. Alesina, Glaeser, and Sacerdote (2001) and Luttmer (2001) describe how racial divisions affect redistributive policies in the United States. Cutler, El-

mendorf, and Zeckhauser (1993), Poterba (1997), and Goldin and Katz (1999) find evidence that suggests that racial heterogeneity affects local policies toward public education. Alesina, Baqir, and Easterly (1999) argue that, in racially fragmented areas, individuals are less willing to pool their fiscal resources to provide public goods. Glaeser, Scheinkman, and Shleifer (1995) provide evidence that racial heterogeneity affects city growth. The third strand is the literature on the formation of local governments in the United States. Burns (1994), for instance, argues that land developers (who presumably had an interest in maximizing the value of their land) were important in the creation of jurisdictions. Finally, our paper can be seen as a test of the Tiebout (1956) model, in which households sort themselves among local jurisdictions according to their preferences for local public goods and taxes. Previous tests of the Tiebout model have always taken the number of jurisdictions as given, but this is a restriction not envisioned by Tiebout, who assumed that adjustment would occur both through household mobility and through endogenous formation of jurisdictions.

The paper is organized as follows. In Section II, we model the hypothesis that there is a trade-off between economies of scale and heterogeneity. Our empirical strategy and data are described in Section III. In Section IV, we present our results on school districts and attendance areas. Section V contains results on municipalities and special districts. Section VI discusses the actual mechanisms by which local jurisdictions consolidate or break apart. Section VII provides final comments and conclusions.

II. The Theory

A. *A Simple Model of Jurisdictions*⁴

Consider a political jurisdiction that has a population of size M . With an eye to the empirical work that follows, let us call this jurisdiction a “county.” There are T types of individuals, and T is an even integer. The types are located at a distance h from each other. The mass of individuals of each type is denoted m , so that $M = mT$. Without loss of generality, the left-most and right-most individuals are located at a distance of $h/2$ from the borders of the county. We denote density by d and note that $d = m/h$.

For now, we consider only one spatial dimension in order to keep the model simple, and we interpret the distance between individuals as a general measure of their difference—which may be ideological, geographic, taste-based, or income-based. Individuals actually differ on mul-

⁴ This model builds on Alesina and Spolaore (1997).

multiple dimensions simultaneously, but the single-dimension model makes the predictions clear. We relax the single-dimension assumption below, and we take multiple dimensions into account in our empirical work.⁵

The assumption that each individual's location is fixed is natural if location represents tastes or ideology. It is less natural if location represents geography because individuals can move in response to changes in jurisdictional boundaries. In our empirical work, we address such endogenous mobility, treating it as a causality problem.

Each individual has the following utility function:

$$U_i = g(A - al_i) + y - t_i, \quad g > 0, a > 0, A > 0. \quad (1)$$

In equation (1), l_i is the distance of individual i from the public good, y is income, and t_i is the tax paid by individual i . Thus $y - t_i$ is private consumption. The linearity of this utility function simplifies the algebra but does not affect the qualitative nature of the results. The utility an individual derives from the public good is decreasing with his distance from it, where, remember, distance captures both a geographical and an ideological dimension. For example, a county might contain a white population that prefers a traditional school located in the suburban area, a Hispanic population that prefers a school with bilingual education that is located in an urban area, and a black population that prefers a school that teaches black history and is located in the urban area.

We are interested in the number of school districts, say, into which this county splits. Each school district provides a public school, and residents of a district attend the school and pay taxes to finance it.⁶ Thus a school and two borders characterize a school district.⁷ (Below, we discuss the possibility of multiple schools in a district.) The cost of each school is given by

$$k = \bar{k} + k_1 Sm, \quad (2)$$

where k is cost, \bar{k} is fixed cost, k_1 is variable cost, and S is the number of types of the school district, so that Sm is the population being served by the district. We have economies of scale since average costs are de-

⁵ Under certain conditions, partial analysis of a single dimension of a multidimensional model would yield similar results (see, e.g., Epple and Platt 1998). Researchers have used a few approaches to maintain clarity: unidimensional heterogeneity (most common), partial analysis of single dimensions in multidimensional models, or strong restrictions on the correlations among variables on which people differ. Calibrated computational models are useful for prediction but do not yield clear intuition.

⁶ For simplicity, assume that each household uses the school to the same degree.

⁷ It is immediate to show that disjoint school districts would not be equilibrium.

creasing in the district's size.⁸ By the budget constraint of each school district, we have

$$\int_s t_i = \bar{k} + k_1 Sm. \quad (3)$$

RESULT. A social planner maximizing the sum of individual utilities would locate the school in the middle of each school district and would choose the following number N of equally sized school districts:

$$N = \begin{cases} \frac{T}{2} \sqrt{\frac{gahm}{\bar{k}}} & \text{if } \frac{T}{2} \sqrt{\frac{gahm}{\bar{k}}} > 1 \\ 1 & \text{otherwise.} \end{cases} \quad (4)$$

The proof is a straightforward generalization of proposition 1 of Alesina and Spolaore (1997).⁹ Since individuals have linear utility, the social planner is indifferent to the distribution of taxes; obviously an income tax would produce the same tax burden for everyone since everyone's income is the same. Note that in order to equalize utilities among individuals, the social planner would choose to draw the borders of the school district between two adjacent individuals; this also implies that every individual strictly belongs to one and only one school district. Several comments are in order.

1. The optimal number of school districts is increasing in the benefits of the public good (captured by the parameter g). In more colorful terms, the more people like schooling, the more they are willing to pay to avoid having a school that is far away, in terms of distance or tastes.
2. The optimal number of school districts is increasing in the disutility of distance (captured by the parameter a). That is, the more people dislike sharing the same public goods with others who have different preferences, the larger the optimal number of jurisdictions and the smaller their size.

⁸ Diseconomies of scale may set in for districts of very large size. See below.

⁹ For instance, in the simplest case (homogeneous density and incomes), the proof is as follows. A social planner would locate the school in the middle of each district to minimize the average (and, thus, total) distance from the school. Since there is a uniform distribution of individuals, the districts will be of equal size. Thus, in a district with population Sm , the average distance from the public good is $\bar{l} = (h/4)S$. The social planner's problem is to maximize $g(A - a\bar{l}) + y - (k/Sm) - k_1$ subject to $\bar{l} = (h/4)S$ and $N = T/S$. Because utility is linear, the social planner is indifferent to the distribution of taxes. The solution to this problem is given by eq. (4). We ignore integer problems, but we point out that each county must have at least one school district.

3. The higher \bar{k} is (which captures the importance of economies of scale), the lower the optimal number of jurisdictions and the larger their size. Each additional person makes costs per person fall greatly in a jurisdiction with a small population, but an additional person has little effect on costs per person in a jurisdiction with a large population. More precisely, the importance of economies of scale declines with the population of a jurisdiction, and the decline is nonlinear in the population.
4. The higher heterogeneity is (captured by the parameter h), the larger the number of school districts. That is, the more heterogeneous the ideologies or tastes of a given population are, the larger the number of districts. If one interprets h as a measure of distance, then the more sparsely populated a county is, the larger the number of districts.
5. The total number of jurisdictions is increasing in the total size of the population. It increases linearly in the number of types, but it increases with the square root of the mass of population m at each type.
6. Consider the case in which the parameter values are such that the county has only one school district. Suppose that heterogeneity increases. The optimal number of school districts may still be one. The reason is that the change in parameter values may be insufficient to push the county past the threshold that makes two school districts optimal. The integer problem is similar. The smaller the population of a county is, the less likely that a given change in heterogeneity will pass the threshold at which creating a new district is optimal.

An interesting question is whether proposition 1 is reproduced by a decentralized equilibrium, in which households choose how many districts to have in their county without the help of a social planner. The answer depends on voting rules and on the availability of interpersonal transfers.¹⁰ Alesina and Spolaore (1997) show, however, that even if the optimal number of jurisdictions cannot be sustained by a voting mechanism, the equilibrium number of jurisdictions has the comparative statics discussed above.

¹⁰ Actually, *within* jurisdictions, we expect house prices to differ to compensate individuals who are arbitrarily located farther from the public good, given the number of jurisdictions. We do not need to invoke this result because we characterized counties as line segments; thus, for an optimal number of jurisdictions, there are optimal boundary lines. If we had characterized counties as circular lines, individuals would fight over *where* boundaries should be drawn for a given number of jurisdictions. We would then need to allow house prices to differ to quell such disputes.

B. *On Geographic and Preference Heterogeneity in a Single Dimension*

So far, we have mapped all types of heterogeneity into a single dimension to keep the model transparent. However, we must relax this restriction in one particular way in order to derive predictions with empirical relevancy because population density is extremely variable in the United States.

Note that, with total population held constant, equation (4) implies that the number of jurisdictions decreases as density increases. If one substitutes T for M/m and $h = m/d$, one gets

$$N = \frac{M}{2} \sqrt{\frac{ga}{dk}}. \quad (5)$$

It is important to see the intuition of equation (5): with M and T held constant, an increase in density d is equivalent to a reduction in h . Thus a reduction in h , with M and T held constant, implies both an *increase* in density and a *reduction* in heterogeneity of preferences. This is where it is troubling to have the single-dimensional line capture both geography and preferences. In the real world, cities have high density (people live close to each other), but cities also have great heterogeneity of preferences, race, income, and so on. Our single-dimensional model cannot handle this reality because it imposes the idea that wherever density is high, heterogeneity of preferences is automatically low. To handle this issue, we need to use a bidimensional model so that geographical distance and ideological distance are not perfectly correlated.

We sever the correspondence between preference heterogeneity and geographical distance by breaking each county into J parts, each part having a mass of population m_j and T_j types of individuals for $j = 1, \dots, J$. Total population is $M = \sum_{j=1}^J M_j = \sum_{j=1}^J m_j T_j$. Assume that the m are increasing with j and h is constant throughout the county. Thus a lower subscript identifies more sparsely populated parts of the county, since remember that $d = m/h$. If one ignores the integer problem, it would be optimal to choose a different N_j for each part of the county that has a different density.

The solution is

$$N = \sum_{j=1}^J N_j = \frac{1}{2} \sqrt{\frac{gah}{k}} \left(\sum_{j=1}^J T_j \sqrt{m_j} \right). \quad (6)$$

Equation (6) has the same basic comparative statics as equation (4), the solution in the simpler case. However, equation (6) allows us to take separate account of the geographic density of an area when we evaluate the effect of preference heterogeneity. The equation also allows us to evaluate counties that have areas of low density and high density—a city

in one corner, a rural area in another. Equation (6) guides our investigation, and we derive our estimating equation directly from it.

C. Discussion and Extensions

1. More than One School in a District

So far, we have identified a jurisdiction with a public good; that is, each district has exactly one school. More generally, if heterogeneity increases, households can build another school within their district. Building a new school and creating a new district are very different choices; the former choice is “cheaper,” in terms of institutional transaction costs, but it does not allow different groups of people in the district to independently control or finance their schools. Legally, all schools in a district must have the same contract with teachers, the same spending per pupil, and so on.

How could we extend the model to allow a district to have multiple schools? Building a new school should have a lower fixed cost than creating a new school district, but building a new school limits the diversification among different groups (compared to creating a new district). If we interpret the line of the model as an ideological spectrum, we can capture these phenomena by assuming that two schools in the same district cannot be too far from each other. That is, multiple schools in a district have to be closer to the ideological middle than they would optimally be if they were schools in separate districts. In short, if heterogeneity increases, residents have two choices: the more radical (but more expensive) choice of creating a new district and the less independent (but cheaper) choice of building a new school. We examine both choices in our empirical analysis.

2. Multiple Public Goods

A school district or special district provides only one type of public good, but municipalities typically provide several public goods, such as policing, fire protection, and roads. The model captures the determination of municipalities if the public good is interpreted as a bundle of local goods and services. The mere fact that municipalities provide multiple public goods suggests that there is a trade-off between economies of scope and heterogeneity that is similar to the trade-off between economies of scale and heterogeneity. If there were no economies of scope, it would be optimal to have a special district for each local public good. Note, however, that different types of public goods imply a very different level of interpersonal interaction, from the very high level of interaction in schools to the very low level in garbage collection. Thus one should

expect the effect of, say, racial preferences to be stronger for jurisdictions providing public goods with higher levels of interpersonal contacts. Below we present evidence consistent with this observation.

3. What Explains Preferences for Homogeneity?

There are two reasons why individuals might prefer homogeneity. One is that individuals who share an ethnic background, race, income, or religion may have more similar preferences over public policies than those who do not. The other reason is that people may actually have preferences similar to those of people in other groups, but they may nevertheless prefer to interact with people in their own group.¹¹ So, for instance, a white person may prefer a mainly white school not because the curriculum is different from that of mainly black schools, but simply because he prefers to interact with individuals of his own race. Our model fits either source of preferences equally well. In writing the exposition of the model, we emphasize the first reason (similar preferences) because it is simpler to envision. However, if we make h a measure of the disutility of interacting, the model can embody the second reason: imagine people from different ethnic or income backgrounds having more or less disutility of interacting depending on how close their ethnicities or incomes were. Norwegian Americans and Danish Americans, for instance, might be closer to the line than Norwegian Americans and Chinese Americans. Empirically, it is difficult to distinguish in a precise way between the two reasons that people might prefer homogeneity. However, we shall provide some highly suggestive evidence.

4. Diseconomies of Scale

In the model, there are no diseconomies of scale, but some people believe, on the basis of anecdotal evidence, that jurisdictions with very large populations are unwieldy and do suffer from diseconomies. It is unclear whether people who make such claims are really considering scale only, with the heterogeneity of the population held constant. In our investigations, we found little evidence of diseconomies of scale, so we do not pursue the issue further.

¹¹ For a discussion of this second hypothesis and empirical evidence on segregation in the United States, see Cutler, Glaeser, and Vigdor (1999).

III. Empirical Strategy

A. From Theory to Testing

Our empirical strategy is guided by equation (6), which we reproduce here in logs:

$$\ln N = \frac{1}{2} \ln h + \frac{1}{2} \ln g + \frac{1}{2} \ln a - \frac{1}{2} \ln \bar{k} + \ln \sum_{j=1}^J T_j \sqrt{m_j} \quad (7)$$

This expression suggests that the number of jurisdictions should depend on four types of variables: (i) measures of the size and density of the county, (ii) measures of fixed costs, (iii) measures of preferences for public goods, and (iv) measures of heterogeneity of preferences, which are our focus.

Let us begin with the first set of variables. The term $T_j \sqrt{m_j}$ suggests that the model calls for measures of density ($d = m/h$) and measures of total population in parts of the county with different population density. Consider a linear approximation of equation (7) using a multivariate Taylor expansion. Define $\Psi = \ln \sum_{j=1}^J T_j \sqrt{m_j}$:

$$\begin{aligned} \ln N = & \text{const} + \frac{1}{2} \ln h + \frac{1}{2} \ln g + \frac{1}{2} \ln a - \frac{1}{2} \ln \bar{k} \\ & + \sum_{j=1}^J \Psi_{T_j} T_j + \sum_{j=1}^J \Psi_{m_j} m_j + \sum_{j=1}^J \Psi_{m_j T_j} m_j T_j \\ & + \sum_{j=1}^J \Psi_{T_j T_j} T_j^2 + \sum_{j=1}^J \Psi_{m_j m_j} m_j^2 + \dots \end{aligned} \quad (8)$$

In equation (8), Ψ_{T_j} is the derivative of $\Psi(\cdot)$ with respect to T_j , and so on. The last five terms in equation (8) show that the model calls for measures of the population living in parts of the county with different density (remember that $M_j = T_j m_j$, so with m_j held constant, T_j measures the total population of j). The model also calls for a measure of density, m_j (remember that, with h held constant, density $d = m_j/h$).

The last five terms of (8) are an array of population and population density variables that describe the *baseline* for jurisdictional creation in a county. Put another way, variables such as preferences and fixed costs may make a county's number of jurisdictions differ from its baseline, which is determined by population and population density.

We implement equation (8) with four categories of density for each county: low population density (fewer than 1,000 people per square mile), medium population density (between 1,000 and 10,000 people per square mile), high population density (between 10,000 and 50,000 people per square mile), and very high population density (more than 50,000 people per square mile). This gives us 20 population and population density variables ($j = 4$ times five terms). Having these 20 terms

greatly improves the fit of our equation. Essentially, if we do not get the baseline right, it is hard to explain the data. However, we found that setting j bigger than four did not improve the fit.

As a proxy for fixed costs (\bar{k}), we use natural boundaries for jurisdictions: streams. Hoxby (2000) shows that areas with more streams have more jurisdictions, all else equal. In addition, we include state indicator variables to proxy for fixed costs generated by different state laws and regulations. As proxies for the *level* at which the county's population desires the public good (g in the model), we include the county's mean income, share of adults with a high school education, share of adults with a college education, and share of people who are age 65 or older. In some specifications, we also include industry employment shares (the share of employment associated with each industry). It is not obvious that industry composition affects the preferred level of public goods, but we use the industry employment shares to test whether the results are sensitive to including county characteristics that *may* affect public goods. Finally, our main variables of interest are measures of the fragmentation of preferences (h in the model), which we proxy with racial and ethnic fragmentation indices, religious fragmentation indices, and measures of income inequality.

B. The Causal Mechanism through Which the Trade-off Operates

We have presented the model as though an area's population is exogenously determined and the number of jurisdictions responds endogenously. The model, however, really says only that a certain number of jurisdictions is optimal, given a population's heterogeneity. So, if the model were correct and households were mobile across areas, households might migrate to areas that were divided up more optimally. Endogenous mobility would not make the model wrong; it would affect how one thought about the mechanism through which the trade-off worked. We do not think, however, that endogenous mobility is likely to be the mechanism because there is no guarantee that an area with a large number of jurisdictions would attract migrants who have the "right" mix of heterogeneity for the number of jurisdictions. For instance, an area with many (and therefore expensive) jurisdictions might appeal disproportionately to white, high-income households, but their migration would make the area less heterogeneous, thereby working "against" finding a correlation between fractionalization and number of jurisdictions.

We attempt to resolve the issue of causality by two means. First, we assure ourselves that the relationship is associated with *changes* in jurisdictions by looking at panel evidence: are changes in population heterogeneity associated with changes in the number of jurisdictions? If

we were to see the jurisdictions remaining stable while population heterogeneity changed to fit the existing jurisdictions, it would suggest that endogenous mobility was the key mechanism. However, showing that changes in jurisdictions go hand in hand with changes in population heterogeneity is only necessary, not sufficient, for causality. For sufficient evidence, we need changes in heterogeneity that are credibly exogenous. We find such changes in the shocks to certain counties' racial heterogeneity that occurred during World Wars I and II. In the two "Great Black Migrations," Northern war industries drafted black workers from the South to replace their former supply of white workers. The supply of white workers decreased both because white men served in the military and because the wars shut off the flow of white immigrants.

C. Data

We consider three types of jurisdictions: school districts, municipalities, and special districts. We also consider school attendance areas within districts. Our variables are generally measured at the county level, and the key dependent variables are the numbers of jurisdictions (of a given type) in a county. Because we need stable areas that are capable of being divided into jurisdictions, we exclude counties from our sample that are clearly inappropriate. The excluded counties are those with unstable boundaries and those in states that do not allow counties to be divided into multiple local jurisdictions.

Although much of our data come from U.S. Censuses of Population, we use data from more than 50 sources. In order to focus on facts that every reader needs to know, we have relegated some details of our data set construction to the Data Appendix in the online edition of the *Journal*. Table 1 shows summary statistics for our 1990 variables; other summary statistics are shown in the Data Appendix (table A1).

We use the Gini coefficient as our main measure of income heterogeneity. The mean county in our sample has a Gini coefficient of 0.407 and a standard deviation of 0.031. We obtained similar results using the Theil index, the coefficient of variation, and ratios of income deciles.

Our index of racial heterogeneity is the probability that two randomly drawn individuals in a county belong to different races, where the "races" are the five categories used in the 1990 Census of Population: white non-Hispanic, black non-Hispanic, Asian and Pacific Islander, Native American, and Hispanic. Formally,

$$\text{race} = 1 - \sum_i (\text{group}_i)^2, \quad (9)$$

where group_i denotes the share of the population that belong to race

TABLE 1
DESCRIPTIVE STATISTICS FOR COUNTIES

Variable	Mean	Standard Deviation
Number of districts	5.444	7.730
Number of schools	25.781	58.844
Number of municipalities	6.498	6.649
Number of special districts	10.904	17.026
Racial heterogeneity index based on whole population	.188	.176
White ethnic heterogeneity index based on whole population	.728	.088
Hispanic ethnic heterogeneity index based on whole population	.379	.197
Racial heterogeneity index based on school-aged population	.215	.188
White ethnic heterogeneity index based on school-aged population	.697	.107
Hispanic ethnic heterogeneity index based on school-aged population	.306	.212
Gini coefficient for household income	.407	.031
Religious heterogeneity index	.631	.181
School-aged population (1,000s)	17.620	57.700
Population (1,000s)	79.182	263.813
Land area (1,000 sq. mi.)	1.115	3.811
Number of streams in county	70.114	95.281
Mean household income (1,000s)	29.389	6.867
Percentage of adults with at least a high school education	70.622	10.119
Percentage of adults with at least a college education (16 years)	12.812	6.090
Percentage aged 65 or older	14.861	4.420
Population density (1,000 sq. mi.)	.231	1.633
Percentage of employment in agriculture	10.061	10.115
Percentage of employment in mining and resources	2.625	5.576
Percentage of employment in construction	7.129	2.674
Percentage of employment in manufacturing	20.827	12.168
Percentage of employment in transportation	6.582	2.504
Percentage of employment in trade	18.614	3.652
Percentage of employment in finance, real estate, and insurance	3.847	1.658
Percentage of employment in business services	2.713	1.539
Percentage of employment in personal services	3.933	2.047
Percentage of employment in entertainment	6.527	2.741
Percentage of employment in health	2.299	.814
Percentage of employment in education	8.990	3.625
Percentage of employment in other professions	3.077	1.201
Percentage of employment in public administration	5.068	3.596

NOTE.—The table shows unweighted descriptive statistics for the data, in which an observation is a county. A county is in the sample if it is stable and *can* legally have lower-level jurisdictions within it. The data are taken from the Census of Population and Housing, the Census of Governments, Department of Education (1994), National Council of Churches of Christ of the United States (1956), and the U.S. Geological Survey. The online Data Appendix shows population and population density variables for the low, medium, high, and very high density areas within counties.

i. The mean county in our sample has a heterogeneity index of 0.188 with a standard deviation of 0.176.¹²

In theory, our heterogeneity index does not distinguish between counties that are 80 percent white and 20 percent black and “reverse” counties that are 20 percent white and 80 percent black. In practice, however, whites have a plurality in 98 percent of the counties in our sample. Therefore, for all intents and purposes, “more heterogeneity” means “fewer whites” in American data.

We define analogous indices of ethnic heterogeneity within the white and Hispanic populations of the United States. (Ethnic groups within the black, Asian, and Native American populations are too small to be usable.) Specifically, we define an index of white (Hispanic) ethnic heterogeneity as the probability that two randomly drawn white (Hispanic) individuals in a county belong to different primary ancestry groups.¹³ To make the results interpretable, we need to define ancestry groups that have roughly equal distinctness as *groups*. For instance, if a family were the only Botswanan Americans in a city, it is likely that they would not stay distinct but instead seek out people from other countries in sub-Saharan Africa. Thus we collapse small ancestry groups within the white and Hispanic populations on the basis of the language and geographic proximity of their mother countries. For instance, the Scottish are aggregated with the English into the British. Such aggregations are not a science, but they are necessary, and all major ethnic groups in the United States have their own categories. The mean county has a white ethnic heterogeneity index of 0.728 (with a standard deviation of 0.088) and a Hispanic ethnic heterogeneity index of 0.379 (with a standard deviation of 0.197).

We define analogous indices of religious heterogeneity using data on adherence to 17 major Judeo-Christian denominations.¹⁴ The mean county has a religious heterogeneity index of 0.631 (with a standard deviation of 0.181).

¹² Because there is no one best way to measure racial heterogeneity, we experimented with measures other than the one described above. We replaced the racial heterogeneity index with separate percentage black and percentage Hispanic variables (which are the major sources of variation in the racial heterogeneity index), but we found that percentage black and percentage Hispanic have effects that are so similar that we lose little information by using a single index of racial heterogeneity.

¹³ We use the following ancestry/ethnic groups for whites: British, Irish, French, Italian, German, Greek, Portuguese, Swiss/Austrian, Benelux (Belgian, Dutch, Luxembourgian), Scandinavian, Russian/Ukrainian, Hungarian, Polish, other Eastern European, Arab, and other white. We use the following ancestry/ethnic groups for Hispanics: Mexican, Cuban, Puerto Rican, South American, and other Central American.

¹⁴ We use the following religious groups: Baptist, Catholic, Christian Scientist, Eastern/Byzantine Rite Catholic, Congregationalist/related Reformed Christian, Episcopalian, Friends, Jewish, Lutheran, Mennonite/Amish, Methodist, Mormon, Orthodox, Presbyterian, Seventh Day Adventist, Unitarian/Universalist, miscellaneous conservative, and evangelical Christian.

The last five terms of equation (8) tell us that we need the following variables for each density category within a county: population, population squared, population density, population density squared, and population times population density. We have the four density categories described above (low, medium, high, and very high), so we build 20 variables from census block- (or tract-) level data, aggregated to the county level. See the online Data Appendix for details of their construction.

Municipalities are general-purpose governments such as cities, towns, boroughs, and villages. Special district governments are units that are administratively and fiscally independent of municipalities. Most special districts perform a single function or a very limited number of functions, such as fire protection, water supply, drainage, garbage collection, or flood control. The procedures for creating a special district are considerably less demanding than those for creating a municipality or school district. See the online Data Appendix for additional data on jurisdictions.

For our cross-sectional analysis, we use data from 1990 because that year has the most detailed data. For our basic panel analysis, we use data from 1960 and 1990 and take care to make them comparable (see the online Data Appendix). The reason that we use the long interval from 1960 to 1990 is that jurisdictional consolidation and secession were slow during this period, owing to the fact that many obvious jurisdictional changes had already occurred. If we were to study the beginning and end of a single decade during this period, the number of jurisdictional changes would be too small to generate precise estimates. For our panel analyses of the two world wars, we can look at single decades (1910–20 and 1940–50) because jurisdictional changes occurred more frequently. For our panel analyses of the two world wars, we examine only school districts because we were unable to find sufficient information on municipalities and special districts.

IV. Results on School Districts

A. *The Cross-Sectional Relationship between Heterogeneity and School Districts*

In this section showing cross-sectional results, we attempt to establish the pattern and strength of the relationship between population heterogeneity and the number of jurisdictions. We do not, however, insist on a particular direction of causality, so readers should interpret the cross-sectional results as evidence that the trade-off exists but with an unknown mechanism.

Table 2 displays our basic results on the number of school districts

TABLE 2
EFFECT OF POPULATION HETEROGENEITY ON THE NUMBER OF SCHOOL DISTRICTS IN A COUNTY
Dependent Variable: ln(Number of School Districts in a County)

	POPULATION HETEROGENEITY VARIABLES BASED ON							
	Entire Population				School-Aged Children			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Racial heterogeneity	.288 (.096)	.279 (.096)	.280 (.100)	.284 (.102)	.260 (.085)	.228 (.089)	.216 (.087)	.204 (.091)
White ethnic heterogeneity		-.433 (.163)		-.271 (.163)		-.144 (.136)		-.046 (.136)
Hispanic ethnic heterogeneity		.065 (.062)		.053 (.062)		-.015 (.056)		-.010 (.055)
Gini coefficient household income	1.500 (.601)	1.369 (.612)	1.434 (.600)	1.242 (.611)	1.511 (.600)	1.322 (.624)	1.500 (.598)	1.284 (.624)
Religious heterogeneity	.032 (.086)	.041 (.089)	-.024 (.086)	-.009 (.088)	.036 (.086)	-.054 (.092)	-.015 (.086)	-.065 (.091)
ln(mean household income)	.338 (.104)	.295 (.105)	.246 (.129)	.240 (.131)	.322 (.104)	.266 (.108)	.249 (.130)	.204 (.136)

Percentage of adults with at least high school	-.003 (.002)	.001 (.002)	-.001 (.002)	.001 (.003)	-.003 (.002)	-.001 (.002)	-.001 (.003)	.001 (.003)
Percentage of adults with at least college	-.008 (.003)	-.008 (.003)	-.015 (.004)	-.012 (.004)	-.007 (.003)	-.007 (.003)	-.013 (.004)	-.011 (.004)
Percentage of population aged 65 or older	.014 (.003)	.013 (.003)	.015 (.004)	.016 (.004)	.013 (.003)	.012 (.003)	.014 (.004)	.012 (.004)
ln(number of streams)	.028 (.011)	.026 (.011)	.045 (.011)	.045 (.012)	.027 (.011)	.024 (.011)	.043 (.012)	.043 (.012)
20 variables that describe population and pattern of population density	✓	✓	✓	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
16 industry share variables			✓	✓			✓	✓
Observations (counties)	2,718	2,670	2,718	2,670	2,718	2,546	2,718	2,546

NOTE.—Least-squares estimates of cross-sectional data from 1990. Standard errors follow coefficients, in parentheses. An observation is a county. There are fewer observations in regressions that include the Hispanic ethnicity index because some districts have missing information about the ancestry of the Hispanic population. Such information appears to be missing at random. Sources of data are described in the online Data Appendix and include the 1990 Census of Population, Department of Education (1994), and National Council of Churches of Christ of the United States (1956). The 20 variables that describe population and population density are the population of the county in areas of low, medium, high, and very high density; the squares of each of the four preceding variables; the actual population density of the county in its low, medium, high, and very high density areas; the squares of the preceding four variables; and the interactions between the population and population density in each of the four types of areas (low, medium, high, and very high density). See the text and the online Data Appendix for more on these variables. The employment share variables are used for the following industries: agriculture; construction; mining; nondurables manufacturing; durables manufacturing; transportation; communication; retail trade; wholesale trade; business services; finance, insurance, and real estate; personal services; entertainment, health, education, and other professional services.

in each county. The table is structured so that columns 1–4 use heterogeneity indices based on the entire population of each county, and columns 5–8 use heterogeneity indices based on the school-aged population. The school-aged population is important because it determines whom a student could actually meet in school.¹⁵ The specifications shown in the table always include racial, income, and religious heterogeneity, as well as the variables for which the model clearly calls (population variables, population density variables, and proxies for g and \bar{k}). We show specifications with and without ethnic heterogeneity variables so that readers can see for themselves how racial and ethnic heterogeneity interact. Also, we show specifications with and without employment by industry shares to demonstrate that omitted variables are not an obvious source of the results.

The measure of racial heterogeneity has a statistically significant effect on the number of school districts in all specifications. Recall that a two-standard-deviation change in the racial heterogeneity index is 0.36 (0.38 for school-aged children). Thus, if we focus on the specification with the most control variables, we interpret the coefficient on the racial heterogeneity index as follows: a two-standard-deviation (36 percent) increase in the probability that a person will encounter a person of another race in his county raises the number of school districts in his county by 10 percent, all else equal. The corresponding coefficient for school-aged children suggests that there is an 8 percent increase in school districts for a two-standard-deviation increase in the probability of an interracial encounter.

The ethnic heterogeneity indices and industry employment shares change the estimated effect of racial heterogeneity only slightly.

The coefficient on white ethnic heterogeneity is statistically significantly different from zero in only one specification (with no industry shares), and the coefficient on Hispanic ethnic heterogeneity is never statistically significantly different from zero. We conclude that ethnic heterogeneity (as opposed to racial heterogeneity) has little effect on the number of school districts in a county.

Our measure of income heterogeneity (the Gini coefficient) has a positive effect on the number of school districts in a county. Two standard deviations in the Gini coefficient is 0.06, so the interpretation of the coefficient in the most generous specification is as follows: a two-standard-deviation increase in a county's income heterogeneity raises its number of school districts by 8 percent, all else equal. The effect of

¹⁵ The composition of a county's school-aged children may differ from that of its entire population if the county systematically attracts or repels families with school-aged children or has a composition that is shifting over time.

income inequality changes only slightly when ethnic heterogeneity and industry employment shares are included.

The index of religious heterogeneity is never statistically significantly different from zero. Counties with more streams have more districts, probably because streams were natural barriers that affected how a county was initially divided into districts.

The remaining demographic variables are either proxies for g (indicators of how much people desire the local public good) or variables that are intended to reduce the possibility of omitted variables bias. Mean household income is positively associated with the number of school districts. This suggests that higher-income families are willing to pay the extra cost associated with having more districts in order to have districts that are more local. The share of adults with at least a high school education has no statistically significant effect. The share of adults with a college education has a negative effect, but we hesitate to interpret this surprising coefficient because college education is strongly collinear with mean household income. The share of the population who are aged 65 or older has a statistically significant positive effect. We had no expectations for this variable because there is no consensus in the existing literature about the effect of the elderly on local public goods.¹⁶

We do not show coefficients on the 20 population and population density variables because they need to be interpreted as a group and we use them mainly to get a solid baseline so that we can see the effects of other variables. Readers may be interested to know, however, that population does have a positive relationship with the number of school districts in a county. This relationship is very strong when the population is added to areas with a low population density and quite strong when the population is added to areas with a medium or high population density. Adding population to an area with a very high population density has no effect. The reason is probably that it is impractical to divide up a small area into multiple jurisdictions. New York City, for instance, might be predicted to have a jurisdiction for each city block if we did not take its population density into account!

B. Results for School Attendance Areas

Increasing the number of districts in a county is costly, but a demand for separation may be partially satisfied by increasing the number of

¹⁶ Poterba (1997) suggests that the elderly reduce local public goods, especially education. Goldin and Katz (1999) show an opposite effect, using data from the early twentieth century, but their result is not causal. Hoxby (1998) shows that the correlation between elderly and public goods has become more negative over time. She argues that the reason is that the correlation between percentage elderly and a jurisdiction's maturity (i.e., the omitted variables bias) has eroded over time.

school attendance areas. In particular, if people in a district have similar preferences for public goods and merely wish to avoid interacting, separate school attendance areas are preferable to separate districts because they are cheaper and just as effective. If people have heterogeneous preferences, school attendance areas are only a partial solution because they allow only modest differences in preferences to be expressed: major curricular and hiring decisions are made at the district level. Moreover, separate school attendance areas do *not* prevent the redistribution of income through the financing of local public goods. If avoiding redistribution were the only motivation for multiple jurisdictions, income heterogeneity would affect only the number of districts, not the number of attendance areas controlling for the number of districts.

In short, if the number of school attendance areas responds to heterogeneity *controlling for the number of districts*, it is likely that a heterogeneous population has heterogeneous preferences or an unwillingness to interact, not merely an urge to avoid redistribution. We can get similar evidence by looking *within* districts: does a district with a more heterogeneous population divide itself up into more attendance areas, all else equal?

Table 3 investigates these hypotheses. Panel A shows regressions identical to those in table 2, except that the dependent variable is the number of school attendance areas and we control for a county's number of school districts. That is, once the effects of heterogeneity are accounted for in the number of districts, is there remaining pressure on school attendance areas? We use heterogeneity variables based on the school-aged population because the school-aged are the key to student encounters.

The results in panel A of table 3 show that, even after one accounts for the effect of population heterogeneity on the number of school districts, greater population heterogeneity is associated with a greater number of school attendance areas. This suggests that avoiding redistribution is not the only reason why heterogeneity is associated with more jurisdictions. Moreover, nearly all the types of heterogeneity affect the number of school attendance areas, suggesting that heterogeneity raises the number of attendance areas even when it is too meager to raise the number of districts. This is what we expect, given the much higher costs of creating a district. A two-standard-deviation increase in each of the heterogeneity variables raises the number of school attendance areas by the following amounts (given the number of districts): racial heterogeneity, 9 percent; Hispanic ethnic heterogeneity, 6 percent; income heterogeneity, 4 percent; and religious heterogeneity, 15 percent. We focus on the coefficients in column 4 because the specification includes the greatest number of controls.

Panel B of table 3 shows, using *within*-district evidence, that a district

has more attendance areas if it has greater racial, white ethnic, Hispanic ethnic, or income heterogeneity. The coefficients in column 4 can be interpreted as follows. If a school district were to have a two-standard-deviation (34-percentage-point) increase in the probability of interracial encounters, its number of attendance areas would increase by approximately 9 percent. A two-standard-deviation change in within-district income heterogeneity (10 percentage points) would raise the number of attendance areas by about 15 percent. A two-standard-deviation increase in Hispanic ethnic heterogeneity (45 percentage points) would raise the number of attendance areas by 23 percent.¹⁷

In short, it appears that not only racial and income heterogeneity but also religious and Hispanic ethnic heterogeneity raise the number of school attendance areas for a given number of districts. We suspect that racial and income groups have sufficiently different preferences or sufficiently great unwillingness to interact that they affect districts. The trade-offs generated by Hispanic ethnic and religious heterogeneity are apparently too weak to affect districts, but strong enough to affect attendance areas.

C. *Avoiding Interactions versus Avoiding Joint Decisions about Local Public Goods*

If we see that people achieve just as much homogeneity with an inexpensive abundance of school attendance areas as they do with an expensive abundance of school districts, then the desire to avoid interacting with dissimilar people cannot explain all jurisdictional structure. We can infer that people want to make public-goods choices with people whose preferences are similar.

We investigate this hypothesis by forming ratios of actual to possible heterogeneity experienced. First, we form the ratio of the probability that a person would randomly encounter a person of a different race in his *school* to the probability that he would randomly encounter a person of a different race in his *county*. We also form the ratio of the probability that a person would encounter a person of different poverty status in his school to the probability that he would encounter a person of different poverty status in his county.¹⁸

Table 4 shows the results of regressing these ratios of actual to possible

¹⁷ The religious heterogeneity and streams variables are not available at the district level; they are omitted from the equation.

¹⁸ The ratios necessarily rely on school-aged people because we depend on administrative data for analysis of school attendance areas. For poverty status, we use eligibility for free lunch, which is the best available proxy. A student is eligible for free lunch if his or her family is within 125 percent of the federal poverty line.

TABLE 3
EFFECT OF SCHOOL-AGED POPULATION HETEROGENEITY ON THE NUMBER OF SCHOOL ATTENDANCE AREAS

	REGRESSION			
	(1)	(2)	(3)	(4)
A. Dependent Variable: ln(Number of School Attendance Areas in a County)				
Number of school districts in the county	.477 (.014)	.467 (.014)	.442 (.013)	.437 (.013)
Racial heterogeneity based on school-aged children	.325 (.068)	.302 (.068)	.260 (.068)	.249 (.068)
White ethnic heterogeneity based on school-aged children		.134 (.114)		.142 (.110)
Hispanic ethnic heterogeneity based on school-aged children		.228 (.044)		.146 (.042)
Gini coefficient household income	.984 (.427)	1.036 (.429)	.707 (.407)	.722 (.411)
Religious heterogeneity	.571 (.061)	.528 (.062)	.422 (.058)	.408 (.059)
ln(mean household income)	-.007 (.074)	-.005 (.074)	-.193 (.088)	-.198 (.088)
Percentage of adults with at least high school	-.002 (.002)	-.003 (.002)	-.001 (.002)	-.002 (.002)
Percentage of adults with at least college	-.003 (.002)	-.003 (.002)	-.008 (.003)	-.008 (.003)
Percentage of population aged 65 or older	-.015 (.002)	-.017 (.002)	-.017 (.002)	-.018 (.003)
ln(number of streams)	.036 (.008)	.028 (.008)	.039 (.008)	.036 (.008)
State fixed effects	✓	✓	✓	✓
20 variables that describe population and pattern of population density	✓	✓	✓	✓
16 industry share variables			✓	✓
Observations	2,718	2,546	2,718	2,546
B. Dependent Variable: ln(Number of School Attendance Areas in a School District)				
Racial heterogeneity based on school-aged children	.357 (.043)	.180 (.046)	.357 (.042)	.260 (.046)
White ethnic heterogeneity based on school-aged children		.229 (.064)		.053 (.063)
Hispanic ethnic heterogeneity based on school-aged children		.625 (.027)		.506 (.026)
Gini coefficient household income	1.032 (.109)	1.354 (.163)	.893 (.109)	1.460 (.164)
ln(mean household income)	.117 (.028)	.094 (.036)	-.042 (.032)	-.089 (.042)
Percentage of adults with at least high school	.001 (.001)	-.001 (.001)	-.001 (.001)	-.001 (.001)
Percentage of adults with at least college	-.001 (.001)	-.002 (.001)	.001 (.001)	-.002 (.001)
Percentage of population aged 65 or older	-.650 (.089)	-1.290 (.118)	-.604 (.088)	-1.280 (.120)
State fixed effects	✓	✓	✓	✓

TABLE 3
(Continued)

	REGRESSION			
	(1)	(2)	(3)	(4)
20 variables that describe population and pattern of population density	✓	✓	✓	✓
16 industry share variables			✓	✓
Observations	14,718	12,265	14,718	12,265

NOTE.—See the note to table 2. In panel A, the number of school districts is controlled for.

heterogeneity on a county's number of school districts and school attendance areas.¹⁹

We find that the availability of more school attendance areas has a very strong effect on the homogeneity a person experiences, but the availability of more school districts has little or no effect. For instance, the coefficients in the first row of table 4 suggest that people who live in a county with twice as many districts, all else equal, reduce their probability of interracial encounters by 0.9 percent and their probability of cross-poverty encounters by 0.1 percent. Not only are these point estimates small; the first is statistically significantly different from zero at only the 10 percent level and the second is not statistically significantly different from zero at any conventional level. (Note that these reductions are percentages of whatever a person's base probabilities were.)

The coefficients in the second row of table 4 suggest that people who live in a county with twice as many school attendance areas, all else equal, reduce their probability of interracial encounters by 3 percent and their probability of cross-poverty encounters by 4 percent. Both of these coefficients are highly statistically significantly different from zero.

In short, it is the number of attendance areas and not districts that affects the homogeneity that a student actually experiences. Students in a heterogeneous district can attend schools that are just as homogeneous as those attended by students in a homogeneous district. The heterogeneous districts just have to be broken into numerous attendance areas. The fact that the number of *districts* responds to population heterogeneity (table 2), even though student homogeneity is achieved mainly through attendance areas (table 4), indicates that both motivations for avoiding heterogeneity are probably operative: People avoid making joint decisions with people who have different preferences *and* people avoid interacting with people outside their group.

Many previous researchers who have not used attendance area data

¹⁹ The regressions also include the other explanatory variables in table 2 because those variables might make people live homogeneously even if they did not affect jurisdictional formation.

TABLE 4
EFFECT OF THE NUMBER OF SCHOOL DISTRICTS AND SCHOOL ATTENDANCE AREAS ON
THE HETEROGENEITY THAT PEOPLE ACTUALLY EXPERIENCE

	DEPENDENT VARIABLE	
	Ratio of Actually Experi- enced Racial Heterogeneity over Possible Racial Heterogeneity* (1)	Ratio of Actually Experi- enced Poverty Heterogeneity over Possible Poverty Heterogeneity† (2)
ln(number of districts in county)	-.009 (.005)	-.001 (.006)
ln(number of school at- tendance areas in county)	-.030 (.006)	-.040 (.008)
Gini coefficient house- hold income	-.637 (.122)	-.071 (.027)
Religious heterogeneity	-.019 (.018)	.052 (.023)
ln(mean household income)	-.068 (.026)	.012 (.032)
Percentage of adults with at least high school	.001 (.001)	.001 (.001)
Percentage of adults with at least college	.003 (.001)	-.002 (.001)
Percentage of population aged 65 or older	.287 (.073)	.151 (.107)
ln(number of streams)	-.002 (.002)	-.001 (.003)
12 other variables de- scribing population and population density	✓	✓
State fixed effects	✓	✓
16 industry share variables	✓	✓
20 variables that describe population and pat- tern of population density	✓	✓
Observations (counties)‡	2,294	1,141

NOTE.—See the note to table 2.

* This ratio is the probability that a student meets a student of a different race in his school, divided by the probability that he would meet a person of a different race if his school had the same composition as his county.

† This ratio is the probability that a student meets a student with a different free-lunch eligibility in his school, divided by the probability that he would meet a person with a different free-lunch eligibility if his school had the same composition as his county.

‡ The numbers of observations differ from those in table 2 only because some districts and schools do not report their racial composition or free-lunch eligibility.

have misleadingly concluded that the number of *districts* in an area is the key to the heterogeneity that students experience at school. This is an incorrect conclusion based on inadequate data.

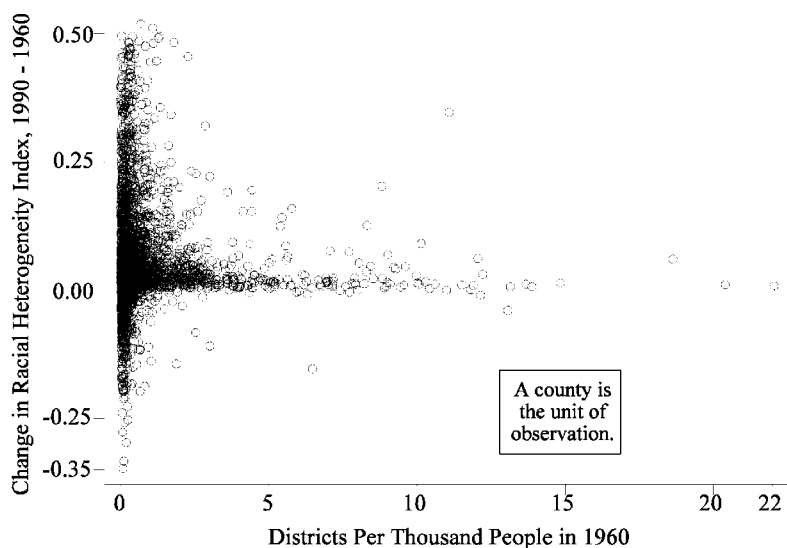


FIG. 1.—Does an amplitude of districts cause racial heterogeneity to change?

D. Panel Evidence: Changes over Time in the Number of School Districts

We now take a stab at clarifying the causal mechanism through which the trade-off operates: does population heterogeneity affect the number of jurisdictions, or is it the other way around? At a minimum, jurisdictions must change if they are responding to heterogeneity. That is, in this subsection, we test a necessary condition for causality, not a sufficient condition. We investigate whether jurisdictions change with population heterogeneity using the interval from 1960 to 1990. Over this period, *increases* in the number of school districts in a county were extremely rare. What we mainly test, therefore, is whether consolidation was slower in counties that were more heterogeneous.²⁰

1. Panel Data in Figures

Before we look at regression results, consider a few figures. In figure 1, we attempt to show whether counties with an amplitude of districts in 1960 attract a racially heterogeneous population between 1960 and

²⁰ Consolidations and secessions are legally easy only when they benefit both potential jurisdictions. Thus many districts can find a district to be a partner and share fixed costs, but few parts of districts that would want to secede can secede, since secessions would usually hurt at least a substantial minority of the population of the area that would be seceded from. See Sec. VII for a description of the legal environment for consolidation and secession.

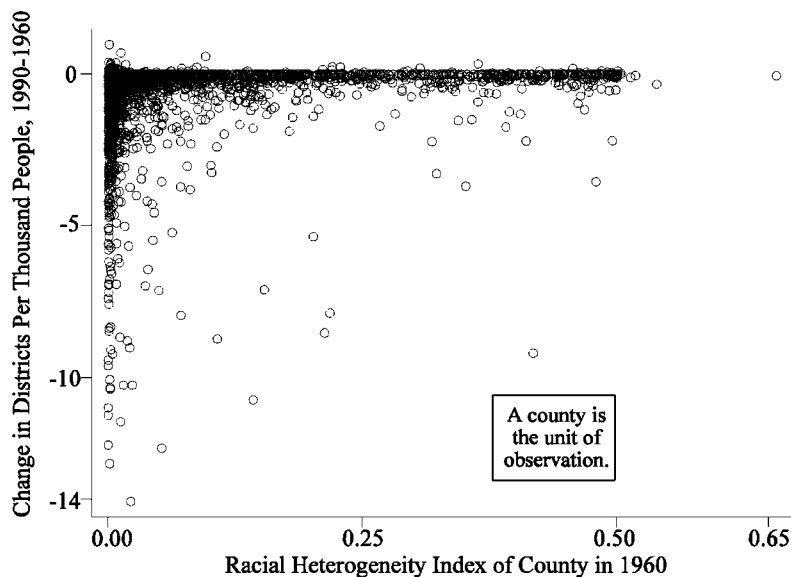


FIG. 2.—Does racial heterogeneity prevent districts from consolidating?

1990. The x -axis records a county's number of districts per person in 1960. The y -axis records the change in a county's racial heterogeneity index between 1960 and 1990. If having an amplitude of districts attracts racial heterogeneity, then we expect an upward-sloping scatter of points.

Figure 1 shows that the relationship appears to be flat. That is, counties with an amplitude of districts are no more likely to attract a racially heterogeneous population than counties with a scarcity of districts. (The observant reader may notice that many counties have fewer than 0.5 district per thousand people. If we focus on just these counties and look at the scatter of points, the relationship is quite flat but slightly *downward* sloping. This is the wrong slope if we are looking for indications that counties with an amplitude of districts attract heterogeneous populations.)

Figure 2 attempts to show whether having a racially heterogeneous population in 1960 made a county less likely to engage in district consolidation between 1960 and 1990. The x -axis records a county's racial heterogeneity index in 1960. The y -axis records the change in a county's number of districts per 1,000 people between 1960 and 1990. Only consolidation, not secession, is occurring over this period, so the top of the y -axis is zero for "no change" and the rest of the y -axis is in negative territory. If having a racially heterogeneous population prevents consolidation, then we expect an upward-sloping scatter of points.

TABLE 5
EFFECT OF CHANGES IN POPULATION HETEROGENEITY ON CHANGES IN THE NUMBER OF
SCHOOL DISTRICTS IN A COUNTY BETWEEN 1990 AND 1960
Dependent Variable: Change in $\ln(\text{Number of School Districts in a County})$, 1990
Minus 1960

	REGRESSION			
	(1)	(2)	(3)	(4)
Change in racial heterogeneity	.931 (.166)	.892 (.164)	.908 (.167)	.880 (.165)
Change in white ethnic heterogeneity		.410 (.048)		.370 (.049)
Change in Hispanic ethnic heterogeneity		.172 (.072)		.101 (.071)
Change in Gini coefficient house- hold income	3.269 (.566)	3.499 (.563)	1.042 (.605)	1.401 (.609)
Change in religious heterogeneity	-.359 (.115)	-.137 (.115)	-.258 (.114)	-.066 (.114)
Change in $\ln(\text{mean householdincome})$	1.132 (.077)	.939 (.079)	1.346 (.087)	1.177 (.089)
Change in percentage of adults with at least high school	-.021 (.003)	-.021 (.003)	-.027 (.003)	-.026 (.003)
Change in percentage of popula- tion aged 65 or older	-.004 (.005)	-.001 (.005)	.004 (.005)	.007 (.005)
20 variables that describe change in population and pattern of population density	↙	↙	↙	↙
Change in industry share variables			↙	↙
Observations	2,718	2,670	2,718	2,670

NOTE.—An observation is a county in which the number of districts can change. All variables are first-differenced (1990 minus 1960). Data are taken from U.S. Department of Education (1994), block files of the 1960 Census of Population, and National Council of Churches of Christ in the United States (1956). See also the note to table 2.

Figure 2 shows that the relationship is, indeed, upward sloping. Racial heterogeneity at the beginning of the period does appear to prevent consolidation. (The observant reader may notice that many counties experienced little or no jurisdictional change over the period. This accounts for the dense cloud of observations at zero on the y -axis.)

2. Panel Data in Regressions

Having looked at some of the data, we show regression estimates of the effect of changing county heterogeneity on changes in the number of districts between 1960 and 1990. Specifically, table 5 shows estimates from a first-differenced version of the equation estimated in table 2. The first-differenced specification not only clarifies whether jurisdictions are actually changing but also ensures that many idiosyncratic features of counties are no longer omitted variables. Any county characteristic that is fixed over time drops out, so this specification is a good test for

coincidental correlation between heterogeneity and a large number of jurisdictions.²¹

The first row of table 5 suggests that counties that experienced increasing racial heterogeneity between 1960 and 1990 were more likely to resist district consolidation over that period. The coefficient on the change in racial heterogeneity is approximately one, which implies that a county that experienced an increase of two standard deviations in racial heterogeneity lost 36 percent *fewer* districts between 1960 and 1990 than a county that experienced no change in racial heterogeneity.

In addition, between 1960 and 1990, it appears that a county that experienced a two-standard-deviation increase in white ethnic heterogeneity lost 6 percent fewer districts and a county that experienced a two-standard-deviation increase in income heterogeneity lost 11 percent fewer districts. Changes in Hispanic ethnic and religious heterogeneity do not appear to affect a county's number of school districts.²²

Ultimately, one can show only the *likely* causality by looking at the panel data in the ways that we have done in this section. To provide more certainty about causality, we need to focus on changes in counties' population heterogeneity that are credibly exogenous.

²¹ Another way to use panel data from 1960 to 1990 is to determine whether counties that had "too many" districts in 1960 (on the basis of their characteristics then) experienced more consolidation between 1960 and 1990. The answer to this question is "yes." We ran the regressions shown in table 2 with 1960 cross-section data and obtained residuals. We then regressed each county's change in its number of districts between 1960 and 1990 on the residuals. We obtained coefficients that suggest that a county that had 10 percent "too many" districts in 1960 was likely to lose about seven of those 10 percent between 1960 and 1990.

²² The reader may be interested in seeing a version of table 5 in which we control for a county's initial levels of racial, income, religious, and ethnic heterogeneity. These results are presented in the online Appendix (table A2). They are nearly the same as those in table 5. The reader may also wish to know what happens if we allow each state to have its own time trend in district consolidation between 1960 and 1990. Given that we are estimating a first-differenced specification, this means allowing a fixed effect for each state. Such a specification is not unambiguously more revealing. If counties in a state are experiencing increased heterogeneity (or even if just a few influential counties in a state are experiencing increased heterogeneity) and they want to resist consolidation, their resistance will show up in state laws and administrative decisions that make resisting easier. We know, in fact, from our reading that influential counties *did* get state laws changed, so it is likely that state time trends will *be* the causal reaction, not an exogenous thing around which the reaction occurs. Put another way, if we control for a time trend for each state, we are probably overcontrolling or throwing out a good deal of the causal relationship in an attempt to eliminate spurious ones. Nevertheless, in the online Appendix (table A3), we show a version of table 5 in which we allow each state to have its own time trend (fixed effect in the difference) between 1960 and 1990. As we would expect, the coefficients on racial and income heterogeneity are smaller (about 60 percent as large), but they have the same signs and are still highly statistically significant. Dixon and Kerstetter (1959) and U.S. Department of Urban Studies (1966) are "manuals" for jurisdictions on how to use and change their state's laws to achieve their consolidation/secession goals. They contain examples of successful campaigns for changes in state laws.

E. Exogenous Shocks to Racial Heterogeneity: Two World Wars

In an attempt to fulfill this need, we turn to historical events that generated credibly exogenous changes in certain counties' racial heterogeneity: the South-North migration of blacks to feed labor-needy war industries during World War I and World War II. The Great Black Migration was driven by the needs of the war industry. In fact, firms that produced war-related products sent trains to the South that they filled with workers whom they quickly recruited (for histories, see Harrison [1991] and Lemann [1992]). The migrations of blacks during the world wars are good natural experiments in racial heterogeneity because their timing and extent were driven largely by European political events. Also, they dramatically affected a small and easily identified group of northern counties that happened to have pre-existing concentrations of industries that produced goods particularly in demand during war. In each war, there were equally industrial counties that were much less affected, simply because their industries were not focused on war goods.

The affected counties were a peculiar subset, even among largely urban northern counties with substantial manufacturing. For instance, in Ohio, three counties experienced substantial black migration during World War I: Mahoning (Youngstown area), Cuyahoga (Cleveland area), and Franklin (Columbus area). Several similarly urban, manufacturing-based counties did not experience nearly as much migration: Hamilton (Cincinnati area), Summit (Akron area), Lucas (Toledo area), Montgomery (Dayton area), and Stark (Canton area). Using county-level data from the decennial Censuses of Population, we selected a group of "affected" counties that were urban and experienced an increase of at least two percentage points in the black share of their population between 1910 and 1920 (for World War I) or between 1940 and 1950 (for World War II). An increase of two percentage points corresponds to the ninetieth percentile of the change in urban counties' percentage black, for both the 1910–20 and 1940–50 decades. We verified that each county had a major war industry (each did). Although two percentage points may appear to be a small change, blacks picked residential locations in the industrial core of each county. Thus the urbanized areas of each affected county experienced substantial increases in racial heterogeneity. The 13 counties affected by World War I migration were concentrated in Ohio, Indiana, Pennsylvania, Michigan, Missouri, Kansas, Nebraska, and West Virginia. The 32 counties affected by World War II migration were more dispersed and were also located in California, Illinois, Maryland, and New Jersey. See the online Data Appendix for a more detailed description of the affected counties.

For each war, we selected a group of comparison counties by matching each affected county to another county in its state that was very similar

ex ante in population, urbanization, and initial share of its population who were black. Our procedure can be described as either matching or blocking on the propensity score (see below). We chose a matched group of comparison counties because putting all the counties into a regression and then controlling for covariates would work poorly, owing to the fact that many “control” counties would be so dissimilar to the affected counties that our estimates would depend strongly on functional form restrictions. Also, we needed to focus on a limited number of comparison counties for practical reasons. Getting school district data on 1910–50 is difficult (see the online Data Appendix).

The goal of matching is to find comparison counties that are very similar to the affected counties on ex ante characteristics that are correlated with the probability of selection into the affected group. Thus, within the matched group of counties, there will be a balance between the affected and comparison observations, at least on the observable variables that affect selection into treatment. Within the matched group, the functional form restrictions imposed by regression are therefore unlikely to be troubling.

We used ex ante characteristics of counties to estimate a probit regression for the probability of selection into the affected group of counties. The probit regression contained state fixed effects, the log of population, population density, the percentage of the population that was urban, the percentage of the population employed in manufacturing, and the initial percentage black. We then followed the recommended propensity score algorithm, adding powers and interactions of covariates until we had a block around each treatment observation that was balanced on the covariates. (See Dehejia and Wahba [2002] for more on propensity score-based estimators.) By balanced, we mean that the observations within a block did not display statistically significant treatment-control differences on any of the covariates, on the basis of two-sided *t*-tests at the .05 level of confidence. Then, for each treatment county, we chose the control county within its block that had the most similar propensity score. The procedure may sound elaborate, but it amounts to a simple match. We could have, in fact, easily picked out the comparison counties by visual inspection of a few key ex ante characteristics: the log of population, percentage urban, percentage employed in manufacturing, and initial percentage black. New York City is so unique that we could not find comparison counties for it (no comparison county was in its block).

1. The World War Experiments in Figures

Our world war data are small in quantity, so although we do regression analysis, we first display the data in a few figures. Figures 3 and 4, which

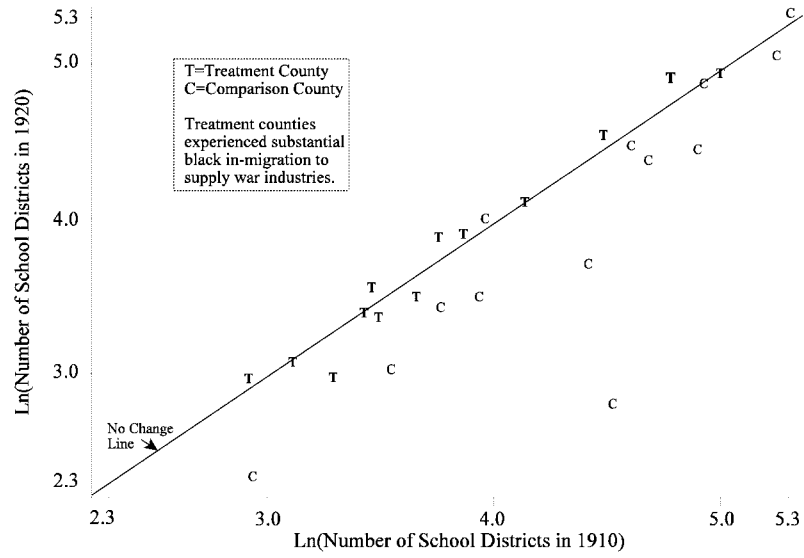


FIG. 3.—Black World War I in-migration and the number of school districts per county

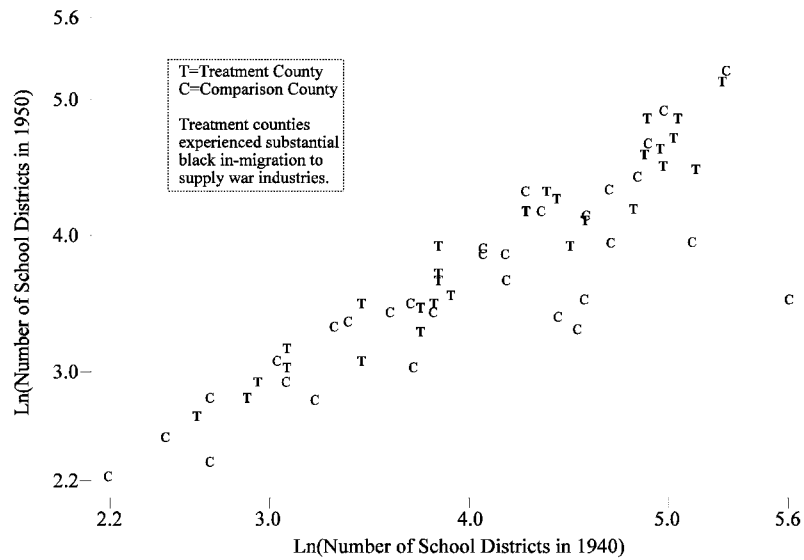


FIG. 4.—Black World War II in-migration and the number of school districts per county

represent World Wars I and II, respectively, work as follows. The x -axis shows the log(number of school districts) at the beginning of the war decade, for treatment and control counties; the y -axis shows the log(number of school districts) at the end of the war decade. A county that maintains its district structure will have its observation sit on a notional 45-degree line, a county with a shrinking number of districts has its observation below the 45-degree line, and so on. Looking at the figures, one can see that the treatment counties tend to lie very close to the notional 45-degree line (or above it, in some cases, showing that their number of districts *rose*). The control counties tend to lie further south, showing that their number of districts was more likely to shrink. In other words, the counties affected by the Great Black Migration resisted district consolidation more than similar counties that were unaffected.

We can make this demonstration clearer by looking at the difference between what the treatment counties actually did and what they would predictably have done if they had not experienced substantial black immigration. For this demonstration, we ran a regression, *using just the control counties*, of the change in the log of a county's number of districts on the variables in our regression: changes in racial heterogeneity, ethnic heterogeneity, religious heterogeneity, and so on. We then used the estimated coefficients to predict what the treatment counties would have done if they had behaved like the control counties and if they had experienced no change in racial heterogeneity (i.e., we zeroed out their change in racial heterogeneity).

Figures 5 and 6, which represent World Wars I and II, respectively, put the predicted change in the number of school districts on the x -axis. The actual change is on the y -axis. If the affected districts behaved as they would have been predicted to, without their real change in racial heterogeneity, the points would be an upward-sloping scatter around a notional 45-degree line. If the scatter is mainly above the notional 45-degree line, then the affected counties are consolidating less than we would have predicted, had there been no change in racial heterogeneity.

Figures 5 and 6 show that the affected counties reduced their number of districts far less than we would have predicted had there been no change in racial heterogeneity. In fact, the World War II treatment counties are lined up almost horizontally at "zero actual change," showing their stiff resistance to consolidation during a period in which aggressive district consolidation was the norm.

We would like to be able to show, cartographically, how school districts in our affected and comparison counties changed. Unfortunately, this is generally impossible because there is so little information about vanished districts (see the online Data Appendix). However, we were fortunate enough to be able to construct school district maps for a pair

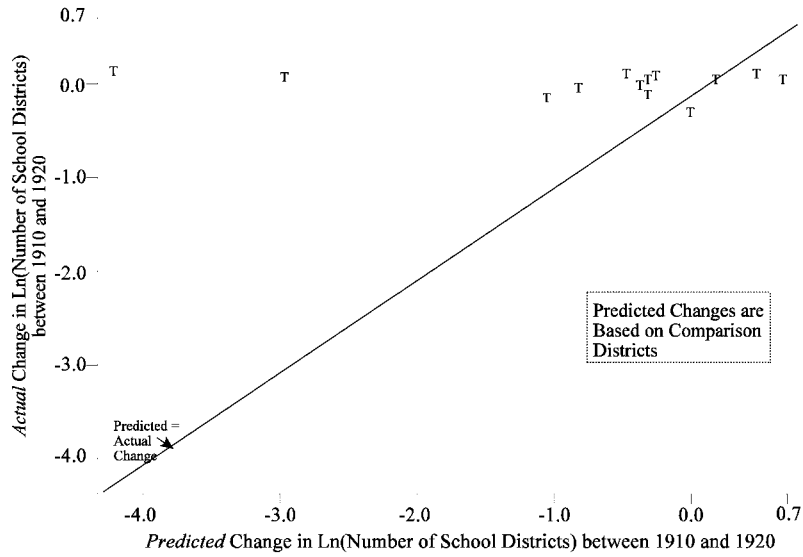


FIG. 5.—Counties with substantial black World War I in-migration do not have the predicted shrinkage in districts between 1910 and 1920.

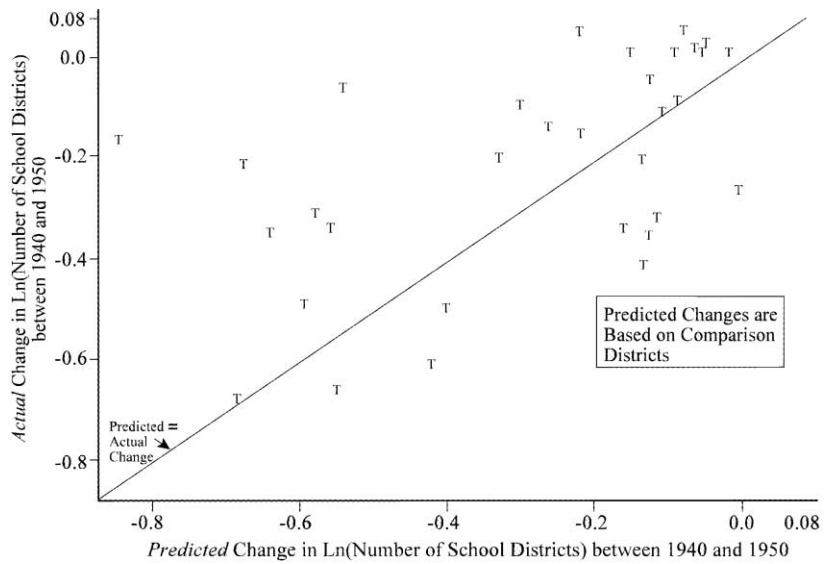


FIG. 6.—Counties with substantial black World War II in-migration do not have the predicted shrinkage in districts between 1940 and 1950.

of counties before and after World War I: Cuyahoga (Cleveland area) and Hamilton (Cincinnati area) counties in Ohio. In World War I, Cuyahoga was a typical treatment county; Hamilton was a typical control county. Figures 7–10 show district maps for Cuyahoga and Hamilton counties in 1910 and 1926. In both counties, there was a good deal of consolidation of rural and small township districts into larger districts. However, the crucial difference is in the area around the industrial city at the core of each county. The Cleveland City School District actually *shrank* between 1910 and 1926, as white residential areas seceded under a process called “village exemption.” Also, in Cuyahoga County, no area that was outside the Cleveland district in 1910 allowed itself to be consolidated with Cleveland. Conversely, the Cincinnati City School District *expanded* between 1910 and 1926, absorbing nearly all the exempted villages in its territory and consolidating smaller districts and schools around its margins.

2. The World War Experiments in Regressions

For comparability with our previous results, we perform a regression analysis parallel to the panel analysis of the last subsection. Specifically, we regress the change in each county’s log(number of districts) on the changes in its racial heterogeneity, ethnic heterogeneity, religious heterogeneity, percentage employed in manufacturing, and the population and density terms shown in equation (7).²³

Table 6 reports our regression results. Column 1 shows the results of the World War I decade regression. The number of school districts in a county is not statistically significantly affected by the change in white ethnic heterogeneity, Hispanic ethnic heterogeneity, religious heterogeneity, or the controls for population, population density, and the percentage of employment in manufacturing. The change in the racial heterogeneity index does, however, have a statistically significant effect on the number of school districts. Affected counties experienced average increases of 0.047 and 0.062 in their racial heterogeneity indices in World Wars I and II, respectively. Multiplying these numbers times the coefficients, we find that the affected counties had 5 percent more districts after World War I and 4 percent more districts after World War II because they had experienced the Great Black Migration. If Cuyahoga

²³ There are no income data in the 1910–40 Censuses of Population, so we cannot include a measure of income heterogeneity. We do not have enough observations to include numerous industry variables, so we include one key one: the percentage of employment in manufacturing. There are no available subcounty data on population density, so we cannot divide counties into subareas with different population densities. We therefore estimate eq. (7) with only one group. With one group, the last term in the equation is best proxied by the log of population plus the log of population density.

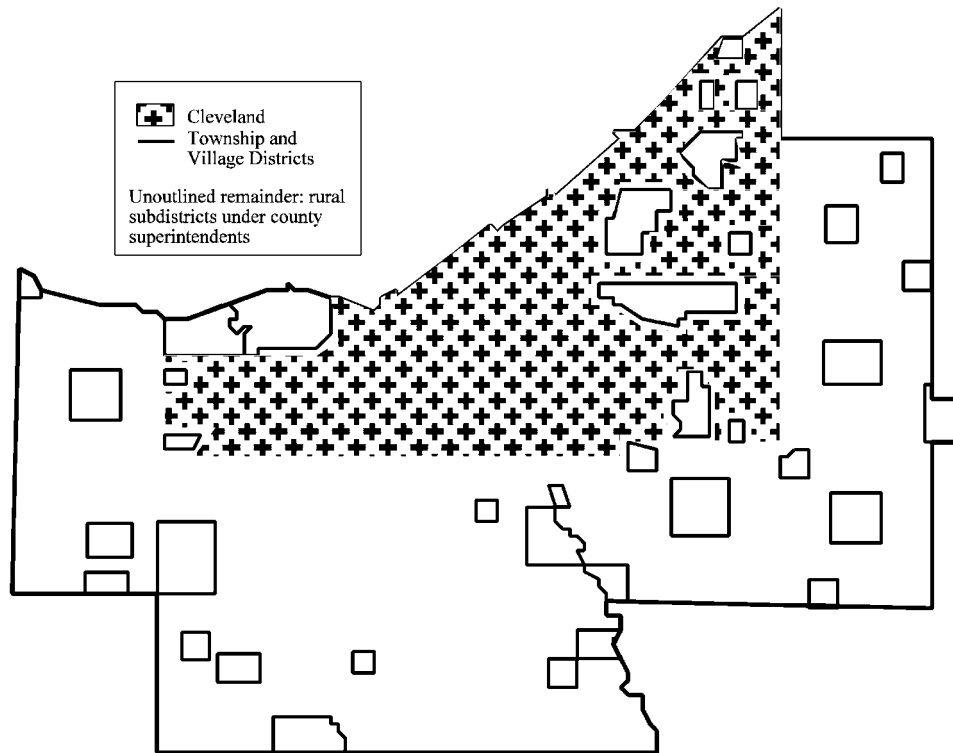


FIG. 7.—Cuyahoga County, Ohio, school districts in 1910

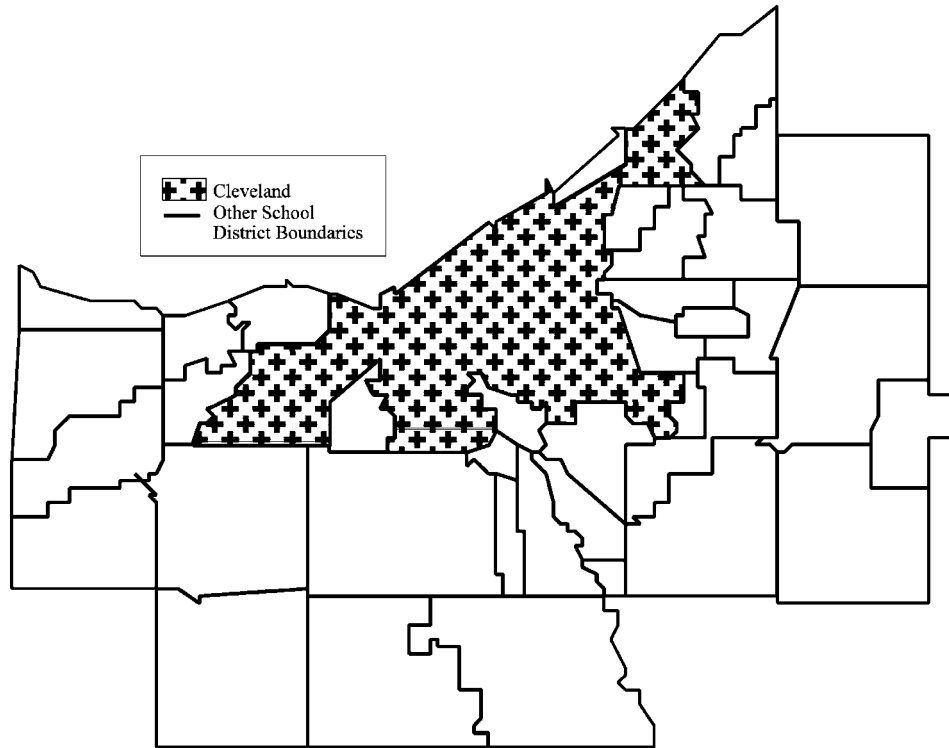


FIG. 8.—Cuyahoga County, Ohio, school districts in 1926

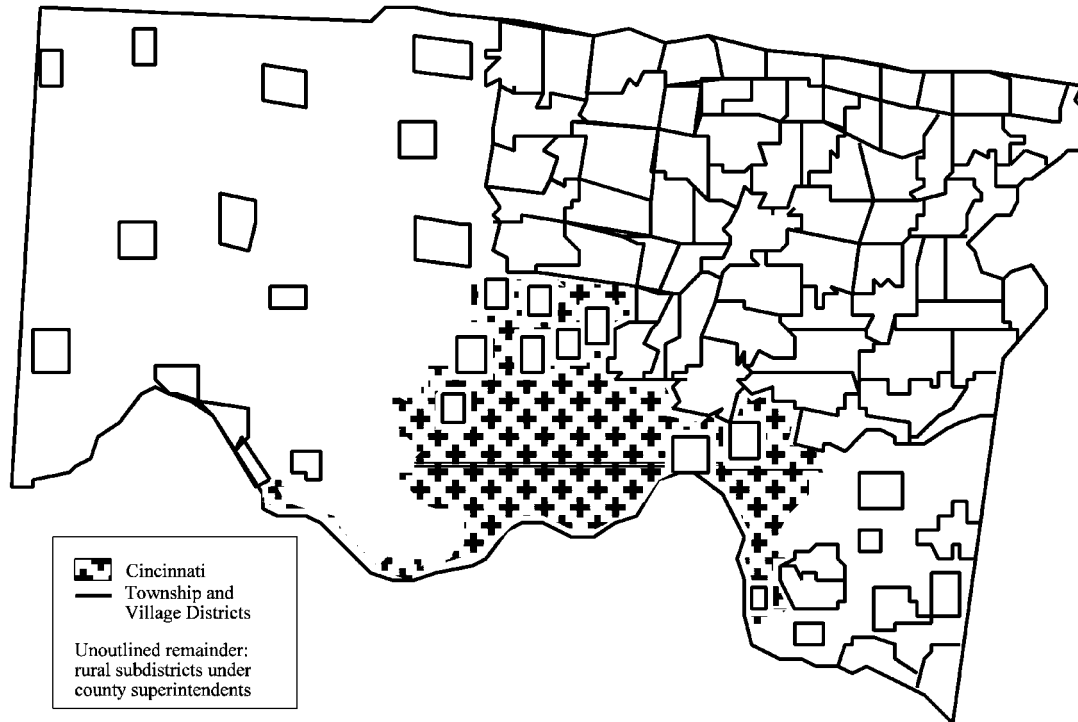


FIG. 9.—Hamilton County, Ohio, school districts in 1910

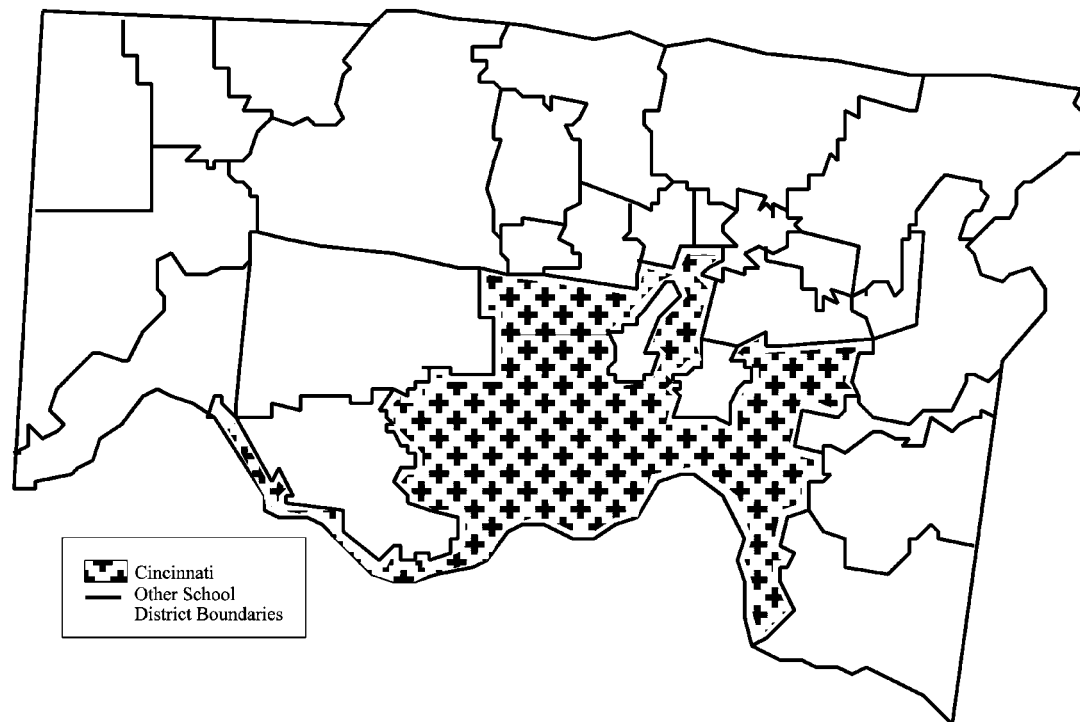


FIG. 10.—Hamilton County, Ohio, school districts in 1926

TABLE 6
EFFECTS OF RACIAL HETEROGENEITY SHOCKS GENERATED BY WORLD WAR INDUSTRY DEMANDS
Dependent Variable: Change in $\ln(\text{Number of School Districts in a County})$ over the Decade

	World War I "Experiment" (1910–20) (1)	World War II "Experiment" (1940–50) (2)
Change in racial heterogeneity	1.049 (.420)	.677 (.317)
Change in white ethnic heterogeneity	–.038 (.107)	.028 (.099)
Change in Hispanic ethnic heterogeneity	.041 (.185)	.259 (.186)
Change in religious heterogeneity	–.868 (.971)	–.687 (.810)
Change in $\log(\text{population})$ and $\log(\text{square root of population density})$	↗	↗
Change in percentage manufacturing	↗	↗
Observations	26	64

NOTE.—Least-squares estimates of a first-differenced version of eq. (7). Standard errors are in parentheses. An observation is either a county that experienced a shock to racial heterogeneity because of war industry demands or a county that was similar to a "shocked" county at the beginning of the decade. The comparison counties were chosen on the basis of their beginning-of-decade match with the affected counties (see text). Data are taken from the Censuses of Population, Censuses of Religious Bodies, Censuses of Governments, Government Units in the United States, and many reports of state superintendents of education. See the online Data Appendix.

and Hamilton counties are typical, the "extra" districts were concentrated around the industrial cities.

V. Results for Municipalities and Special Districts

We now turn briefly to municipalities and special districts because they are interesting in their own right and because we want to see whether the school district results are general. Also, special districts usually involve minimal personal interaction, so they may help us understand whether people value homogeneity because they dislike interaction with people from other groups.

A. *The Cross-Sectional Relationship between Heterogeneity and Jurisdictions*

Table 7 reports our basic regressions on municipalities and special districts. The specification is identical to the one used for the school district regressions in table 2, and the table is organized in the same way as table 2.

The first row shows that racial heterogeneity in a county is a statistically significant predictor of both the number of municipalities and the num-

TABLE 7
EFFECT OF POPULATION HETEROGENEITY ON THE NUMBER OF MUNICIPALITIES AND SPECIAL DISTRICTS IN A COUNTY

	DEPENDENT VARIABLE							
	ln(Number of Municipalities in a County)				ln(Number of Special Districts in a County)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Racial heterogeneity	.144 (.062)	.117 (.061)	.201 (.070)	.181 (.076)	.556 (.115)	.531 (.116)	.485 (.120)	.478 (.122)
White ethnic heterogeneity		-.618 (.178)		-.333 (.178)		.001 (.193)		-.039 (.194)
Hispanic heterogeneity		.070 (.067)		.092 (.068)		.087 (.074)		.007 (.074)
Gini coefficient household income	1.169 (.666)	1.829 (.676)	1.520 (.662)	1.703 (.674)	1.251 (.721)	1.375 (.734)	.901 (.716)	.958 (.733)
Religious heterogeneity	.020 (.095)	-.017 (.097)	-.001 (.094)	-.009 (.096)	.548 (.103)	.521 (.106)	.433 (.102)	.454 (.106)

ln(mean household income)	.625 (.114)	.614 (.116)	.382 (.143)	.377 (.145)	.382 (.124)	.368 (.126)	.344 (.155)	.332 (.157)
Percentage of adults with at least high school	.001 (.003)	.004 (.003)	.005 (.003)	.007 (.003)	-.011 (.003)	-.010 (.003)	-.013 (.003)	-.011 (.003)
Percentage of adults with at least college	-.018 (.004)	-.019 (.004)	-.022 (.005)	-.022 (.005)	.010 (.004)	.009 (.004)	.005 (.005)	.005 (.005)
Percentage of population aged 65 or older	.027 (.004)	.026 (.004)	.032 (.004)	.031 (.004)	.017 (.004)	.015 (.004)	.021 (.004)	.021 (.004)
ln(number of streams)	.033 (.012)	.033 (.012)	.027 (.013)	.023 (.013)	.083 (.014)	.080 (.014)	.069 (.014)	.070 (.015)
20 variables that describe population and pattern of population density	✓	✓	✓	✓	✓	✓	✓	✓
State fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Industry share variables			✓	✓			✓	✓
Observations	2,949	2,901	2,949	2,901	2,908	2,863	2,908	2,863

NOTE.—Least-squares estimates of 1990 cross-sectional data. Standard errors are in parentheses. The numbers of municipalities and special districts are taken from the 1992 Census of Governments. See also the note to table 2.

ber of special districts. For the number of municipalities, the effect of racial heterogeneity becomes stronger in magnitude when the industry employment shares are included. As before, we emphasize the coefficients in the most inclusive specifications (cols. 4 and 8). These coefficients imply that an increase of two standard deviations in the racial heterogeneity index raises the number of municipalities by 7 percent and raises the number of special districts by 18 percent.

Income heterogeneity has a statistically significant effect on the number of municipalities, but not on the number of special districts. An increase of two standard deviations in the Gini coefficient raises the number of municipalities by about 11 percent.

White ethnic heterogeneity has, if anything, a negative impact on the number of municipalities and special districts, but its effect varies with the specification and is at most marginally statistically significant at conventional levels. Religious heterogeneity has no effect on the number of municipalities but increases the number of special districts by a statistically significant amount: an increase of two standard deviations in the religious heterogeneity index raises the number of special districts by about 16 percent. The other covariates have effects much like those they have for school districts (compare to table 2).²⁴

Because adding industry employment shares substantially *raises* the coefficient on racial heterogeneity in the municipality regressions, we explored these variables further and found that they may be picking up the importance of private, local businesses. In fact, if we include the log of the number of private businesses in the county, the industry employment shares lose a great deal of statistical significance. Perhaps, as Burns (1994) argues, the demand for governments is more likely to be implemented when there are businesspeople who have a large enough stake in jurisdictional structure to organize secession campaigns or campaigns of resistance to consolidation.

²⁴ When we compare school districts and counties, it is obvious that provision of public schooling is delegated to the districts. We cannot make such a neat division of responsibility for the public services that municipalities and special districts provide. Some counties provide public services that municipalities or special districts provide in other counties. Although state fixed effects are likely to pick up much of the variation in the division of responsibility, we wanted to see whether the results in table 7 were sensitive to the division of responsibility. The Census of Governments puts public services into 17 categories. For each category, we constructed indicator variables for whether the county (i) owned and operated the public service, (ii) owned the service but contracted out for the operation, (iii) neither owned nor operated but contracted out for the provision of the service, or (iv) did not provide the service. We then ran regressions like those in table 7, except that we included these indicator variables. The coefficients on the measures of heterogeneity were robust to this change in the specification.

B. Panel Results

A necessary condition for causality is that the number of jurisdictions actually change. As before, we determine whether this necessary condition is satisfied using a first-differenced version of the estimating equation. The results are shown in table 8. (For results with controls for initial heterogeneity levels and state fixed effects, see the online Appendix tables A4 and A5.) For municipalities, the estimated coefficient on the racial heterogeneity index is statistically significant and implies that an increase of two standard deviations in racial heterogeneity raises the number of municipalities by about 6 percent. This effect is reassuringly close to the cross-sectional estimates of table 7. In addition, an increase of two standard deviations in income heterogeneity raises the number of municipalities by about 2 percent. Increases in white ethnic heterogeneity are also associated with increased numbers of municipalities. Overall, the municipality results are consistent across tables 7 and 8 (cross-section and first-differences).

The same cannot be said for the special districts results. Recall that the cross-section results in table 7 suggested that special districts were increasing in racial heterogeneity and religious heterogeneity but unaffected by other types of heterogeneity. The first-differenced results in table 8 show that changes in special districts do not appear to be systematically related to changes in any form of heterogeneity. In short, the first-difference results imply that the cross-section results are largely coincidental: arrangements for special districts are idiosyncratic, and special districts are coincidentally associated with some forms of heterogeneity in the cross section. We conclude that special districts are not affected by population heterogeneity, and we surmise that the lack of effect may be due to the low-interaction nature of special districts. Public services such as fire control, water supply, drainage, garbage collection, and flood control do not require interaction among the service recipients. We might also argue that people do not differ much in their demand for such services. However, it is clear that interaction is lacking, but we know that preferences must vary somewhat. If the lack of interaction is the reason that special districts do not change with population heterogeneity, it suggests that an unwillingness to mix with different people is part of the reason for the link between heterogeneity and school districts, school attendance areas, and municipalities.

VI. How Do Jurisdictions Change?

Finally, we are interested in getting inside the black box: *how* does jurisdictional structure respond to population heterogeneity? We argued at the outset that jurisdictions are formed as the trade-off between econ-

TABLE 8
EFFECT OF CHANGES IN POPULATION HETEROGENEITY ON CHANGES IN THE NUMBER OF MUNICIPALITIES AND SPECIAL DISTRICTS, 1960–90

CHANGE IN:	DEPENDENT VARIABLE							
	Log Change in Number of Municipalities				Log Change in Number of Special Districts			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Racial heterogeneity	.237 (.056)	.202 (.057)	.175 (.057)	.168 (.058)	.012 (.165)	.028 (.168)	-.160 (.166)	-.158 (.169)
White ethnic heterogeneity		.072 (.016)		.065 (.017)		-.018 (.049)		-.020 (.049)
Hispanic heterogeneity		-.084 (.025)		-.087 (.025)		.044 (.073)		.069 (.073)
Gini coefficient household income	.598 (.190)	.669 (.194)	.221 (.207)	.297 (.212)	1.316 (.562)	1.320 (.574)	.732 (.605)	.757 (.621)
Religious heterogeneity	.130 (.039)	.159 (.040)	.145 (.039)	.169 (.040)	.276 (.113)	.284 (.117)	.214 (.130)	.215 (.136)
ln(mean household income)	.156 (.026)	.134 (.027)	.194 (.030)	.176 (.031)	-.175 (.076)	-.183 (.081)	-.191 (.087)	-.209 (.091)
Percentage of adults with at least high school	.001 (.001)	.001 (.001)	.001 (.001)	.001 (.001)	.010 (.003)	.010 (.003)	.010 (.003)	-.010 (.003)
Percentage of population aged 65 or older	-.001 (.002)	.001 (.002)	.002 (.002)	.003 (.002)	-.013 (.005)	-.016 (.005)	-.005 (.005)	-.008 (.005)
20 variables that describe population and population density	✓	✓	✓	✓	✓	✓	✓	✓
Industry share variables			✓	✓			✓	✓
Observations	2,844	2,799	2,843	2,798	2,624	2,584	2,623	2,583

NOTE.—Least-squares estimates. Standard errors are in parentheses. An observation is a county. For the first four regressions, the dependent variable is ln(number of municipalities in a county in 1990) minus ln(number of municipalities in the county in 1960), and analogously for special districts for the last four regressions. All independent variables are in changes as well (1990 value minus 1960 value). See also the note to table 2.

omies of scale and population heterogeneity operates on the circumstances created by historical contingency. It is hard to observe directly many historical contingencies and many of the consolidation and secession decisions that generate our results. However, we can observe a key historical contingency: the way in which land was initially distributed in a state. We can also observe a key factor in consolidation and secession decisions: state laws governing such changes. In this section, we see whether we can explain the state fixed effects we estimated using these factors.

In most states, a supermajority of the voters in *each* (potential) jurisdiction must vote for a consolidation or secession (“mutual consolidation”). Getting a supermajority is costly if a minority of voters in either (potential) jurisdiction opposes the change. There is not much variation among states’ secession laws, but there is variation among states’ consolidation laws. States with “strong annexation” laws allow one jurisdiction to annex another, as long as the annexing jurisdiction gets support from the majority of *its* voters. “Weak” annexation laws provide for a similar process but require the annexing jurisdiction to provide substantial evidence that the target jurisdiction is dependent on its (the annexing jurisdiction’s) businesses or infrastructure. “Mutual annexation” is a very weak form of annexation in which annexation can be initiated by only one jurisdiction but must be completed by majority voting in both jurisdictions. We classified states as strong annexation, weak annexation, mutual annexation, or mutual consolidation states using U.S. Department of Urban Studies (1966).

In addition, state effects may be explained by laws that governed how land was sold and how local jurisdictions were set up when the state was still a territory or colony. In particular, if land was sold so that individual landowners tended to acquire diffuse acreage, then large jurisdictions were set up because they allowed a landowner with diffuse acreage to deal with only one local jurisdiction. Conversely, if land was sold so that individual landowners tended to acquire compact acreage, then small jurisdictions were created. Using Halverson (1939), we classify counties according to the legal system under which they were formed: the proprietor system, the direct purchase system, the Louisiana Purchase state system, and the Homestead Act. Under the proprietor system, proprietors (developers) bought large pieces of land, which they repackaged and sold as compact landholdings to individuals. Under the direct purchase system, an individual could buy acreage directly from the (colonial) government. This system encouraged individuals to find the best land available and produced landholdings that straggled over large areas. Most Louisiana Purchase territories had a system similar to the direct purchase system, although some land was repackaged by developers. Under the Homestead Act, individuals could satisfy the home-

stead requirement by buying adjacent parcels, thus creating compact holdings.

Our law variables explain about 30 percent of the variation in the state fixed effects estimated in table 2. (Fixed effects from the municipal regressions in table 7 work similarly.) For instance, we found that counties located in states with strong, weak, and mutual annexation laws have, respectively, 2.1 fewer districts, 1.7 fewer districts, and 0.2 fewer districts than counties in states with mutual consolidation laws. We found that proprietor system and Homestead Act states had, respectively, 10.7 and 7.7 more districts per county than Louisiana Purchase states.²⁵ The effect of the direct purchase laws was not statistically significantly different from that of Louisiana purchase laws.

In short, we have some idea of *how* local jurisdictional structure responds to heterogeneity. Counties begin with different “basic” numbers of local jurisdictions because of different geography and different initial laws; heterogeneity creates pressure for secession or (more often) resistance to consolidation; and the incidence of actual secessions and consolidations varies with the political cost of these activities.

VII. Conclusions

The trade-off between economies of scale and heterogeneity is an important determinant of the number and size of local political jurisdictions in the United States. Racial heterogeneity consistently has a significant positive effect on the number of local jurisdictions. That is, there is strong evidence that people are willing to sacrifice economies of scale in order to avoid racial heterogeneity in their local jurisdiction. The evidence also supports the existence of a trade-off between income heterogeneity and economies of scale, but the trade-off is smaller in magnitude and less robust to altering the specification than the trade-off with racial heterogeneity. We find little evidence of a trade-off between economies of scale and religious heterogeneity, white ethnic heterogeneity, or Hispanic ethnic heterogeneity, except *within* school districts in which forming additional attendance areas is very inexpensive.

We provide both cross-section and panel evidence. Since 1960, there has been strong pressure on jurisdictions to consolidate, but we find that less consolidation took place in counties that were more diverse racially. We offer suggestive evidence on causality using panel data from 1960 to 1990, but our most convincing evidence on causality comes from counties that received positive shocks to racial heterogeneity in

²⁵ All the effects described thus far in this paragraph are statistically significantly different from zero at the 5 percent level.

the two world wars. We find that such counties were significantly less prone to consolidate their jurisdictions than counties that appeared similar at the beginning of each war decade. Both the timing and location of war-related shocks to racial heterogeneity are credibly exogenous, so the war estimates strongly suggest that spurious correlation and endogenous migration do not account for our results.

We find evidence that people avoid heterogeneity because they do not want to interact with different people, but we also find evidence that people avoid heterogeneity because different people prefer different public goods. Our evidence is not decisive enough to allow us to apportion the blame neatly, but we conclude that both explanations are operative.

Our most striking result is the importance of racial heterogeneity relative to income heterogeneity. It is conventional to assert that households sort themselves among jurisdictions on the basis of income, if for no other reason than that they wish to avoid redistribution through the financing of local public goods. However, heterogeneity of preferences and avoidance of interaction receive very little attention from analysts of local public goods. Indeed, most models of local jurisdictions assume that households care *exclusively* about the income of other residents in their jurisdiction. Our work suggests that diverse preferences and avoidance of interaction play at least as important a role as income, perhaps even a more important role. Moreover, our results suggest that race and ethnicity are important determinants of these preferences.

References

- Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Q.J.E.* 114 (November): 1243–84.
- Alesina, Alberto, Edward L. Glaeser, and Bruce Sacerdote. 2001. "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Papers Econ. Activity*, no. 2, pp. 187–278.
- Alesina, Alberto, and Enrico Spolaore. 1997. "On the Number and Size of Nations." *Q.J.E.* 112 (November): 1027–56.
- Bolton, Patrick, and Gerard Roland. 1997. "The Breakup of Nations: A Political Economy Analysis." *Q.J.E.* 112 (November): 1057–90.
- Burns, Nancy. 1994. *The Formation of American Local Governments: Private Values in Public Institutions*. Oxford: Oxford Univ. Press.
- Central Intelligence Agency. 2003. *The CIA World Fact Book 2002*. Washington, D.C.: U.S. Government Printing Office.
- Cutler, David M., Douglas W. Elmendorf, and Richard J. Zeckhauser. 1993. "Demographic Characteristics and the Public Bundle." *Public Finance/Finances Publiques* 48 (suppl.): 178–98.
- Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. 1999. "The Rise and Decline of the American Ghetto." *J.P.E.* 107 (June): 455–506.
- Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score-Matching Meth-

- ods for Nonexperimental Causal Studies." *Rev. Econ. and Statis.* 84 (February): 151–61.
- Dixon, Robert G., Jr., and John R. Kerstetter. 1959. *Adjusting Municipal Boundaries: The Law and Practice in 48 States*. Washington, D.C.: American Municipal Assoc.
- Epple, Dennis, and Glenn J. Platt. 1998. "Equilibrium and Local Redistribution in an Urban Economy When Households Differ in Both Preferences and Incomes." *J. Urban Econ.* 43 (January): 23–51.
- Glaeser, Edward L., José A. Scheinkman, and Andrei Shleifer. 1995. "Economic Growth in a Cross-Section of Cities." *J. Monetary Econ.* 36 (December): 117–43.
- Goldin, Claudia, and Lawrence F. Katz. 1999. "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910–1940." *J. Interdisciplinary Hist.* 29 (March): 683–723.
- Hacker, Andrew. 1992. *Two Nations: Black and White, Separate, Hostile, Unequal*. New York: Scribner's.
- Halverson, F. Douglas. 1939. *County Histories of the United States, Giving Present Name, Date Formed, Parent County, and County Seat*. Salt Lake City: n.p.
- Harrison, Alferdteen, ed. 1991. *Black Exodus: The Great Migration from the American South*. Jackson: Univ. Press Mississippi.
- Huckfeldt, R. Robert, and Carol Weitzel Kohfeld. 1989. *Race and the Decline of Class in American Politics*. Urbana: Univ. Illinois Press.
- Hoxby, Caroline M. 1998. "How Much Does School Spending Depend on Family Income? The Historical Origins of the Current School Finance Dilemma." *A.E.R. Papers and Proc.* 88 (May): 309–14.
- . 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *A.E.R.* 90 (December): 1209–38.
- Lemann, Nicholas. 1992. *The Promised Land: The Great Black Migration and How It Changed America*. New York: Vintage Books.
- Luttmer, Erzo F. P. 2001. "Group Loyalty and the Taste for Redistribution." *J.P.E.* 109 (June): 500–528.
- National Council of Churches of Christ of the United States. 1956. *Survey of Churches and Church Membership, 1952*. Washington, D.C.: Nat. Council Churches of Christ United States.
- Poterba, James M. 1997. "Demographic Structure and the Political Economy of Public Education." *J. Policy Analysis and Management* 16 (Winter): 48–66.
- Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *J.P.E.* 64 (October): 416–24.
- U.S. Department of Education. National Center for Education Statistics. 1994. *School District Data Book: 1990 Census School District Special Tabulation*. Arlington, Va.: MESA Group.
- . 2003. *Common Core of Data 2001–02*. Washington, D.C.: Nat. Center Education Statis.
- U.S. Department of Urban Studies. National League of Cities. 1966. *Adjusting Municipal Boundaries: Law and Practice*. Washington, D.C.: Nat. League Cities.
- Wilson, William Julius. 1996. *When Work Disappears: The World of the New Urban Poor*. New York: Knopf.

Copyright of Journal of Political Economy is the property of University of Chicago Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.