DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 02-15

Do Households Vote With Their Feet?

Valeriy D. Gauzshtein
Department of Economics, University of Colorado at Boulder
Boulder, Colorado

Jeffrey S. Zax
Department of Economics, University of Colorado at Boulder
Boulder, Colorado

October 2002

Center for Economic Analysis
Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

October 2002

DO HOUSEHOLDS VOTE WITH THEIR FEET?

Valeriy D. Gauzshtein
gauzshte@ucsub.colorado.edu

Jeffrey S. Zax
jeffrey.zax@colorado.edu

University of Colorado, Boulder

ABSTRACT

We investigate the relationship between within-community heterogeneity in observed household characteristics and the number of local governments in a metropolitan area. One implication of the Tiebout hypothesis is that, *ceteris paribus*, households can more effectively sort among communities in metropolitan areas where communities are more plentiful. Therefore, communities in such metropolitan areas should be less heterogeneous with respect to the demand for local public goods. The contribution is to note that there are factors, such as geographic variations in housing prices, that cause households to stratify geographically and are unrelated to local public service levels. This stratification, referred to as "statistical sorting", implies that communities should again be more homogeneous if they contain a smaller proportion of the metropolitan population. Separate controls for the number of governments and the number of governments per capita distinguish between Tiebout and statistical sorting. They demonstrate that both types occur.

Keywords: household, heterogeneity, community, sorting.

JEL Classification Numbers: H31, H41, R22, R23.

1. INTRODUCTION

Forty five years ago Charles M. Tiebout (1956) proposed his well known hypothesis that "voting with one's feet" will reveal citizens' preferences for local public services and lead to the efficient provision of local public goods. In response to Samuelson's (1954) argument that the market can not correctly identify demand for collective goods, Tiebout proposed a model under which a market analog could lead to optimal expenditure on local public goods.

Since the initial work of Tiebout, economists have widely adopted the idea that a system of small local governments, competing with each other to attract residents, can result in the efficient provision of local public goods. This view requires that the number of local governments and the diversity of offered fiscal bundles be sufficient to satisfy the diversity of residential preferences present in the area. If this condition holds, then voters will group themselves with others having similar tastes so that public goods can be supplied efficiently. As a result of such sorting behavior, local communities will be populated by households, which are homogeneous with respect to the demand for local public goods.

In spite of the fact that the Tiebout's article has become one of the most cited articles in Economics, it is surprising how little has been done to test whether households actually sort themselves into homogeneous communities. If sorting (voting with feet) does not exist the Tiebout mechanism can not be relied upon.

Economists generally act as if the Tiebout effects are present and households sort themselves among local communities that are more or less homogeneous with respect to

the demand for local public services[1]. Household income is believed to be one of the most important determinants of the local demand. A reliance on income as the only measure of individual preferences for community services (Eberts and Gronberg, 1981; Grubb, 1982; Schmidt, 1992; Aaronson, 1999) is, however, questionable.

Income defines a household's budget constraint for public services. However, the level of income tells us very little about the specific content of a household's preferences. For example, two families with the same income, but with different backgrounds and numbers of school-aged children will probably demand different bundles of local public services. Moreover, empirical investigations of the degree of homogeneity of municipalities with respect to different determinants of public-service demand have found that local communities are rather income-heterogeneous.

Pack and Pack (1977), analyzing data from the metropolitan areas of Pennsylvania, show that only 11 per cent of all suburban towns can be rated homogeneous by household income. They also find that there is substantially more homogeneity with respect to occupation, education, and household type. Stein (1987) looks at data from municipalities in MSA's from 41 states. Confirming Pack and Pack (1977), Stein's findings show that suburban localities exhibit very heterogeneous income distributions. His results also indicate that the residential composition of municipalities within the same metropolitan area is highly heterogeneous with respect to age, housing, and occupation. The diversity scores for only two of Stein's six measures (education and race) show a homogeneous sorting of residential populations.

[1] See for instance papers by Goldstein and Pauly (1981), Rubinfield, D. L., P. Shapiro, and J. Roberts (1987), and Reid (1990) where authors consider the econometric consequences of the Tiebout sorting behavior on the parameter estimates in public service demand studies.

Although income is definitely not a sufficient statistic for representing individual demands for local public goods, it is hardly doubtful that sorting occurs. Grubb's (1982) study of Boston area communities between 1960 and 1970 discovers evidence of increasing segregation by income and (much more weakly) segregation by age. Ottensmann (1982), using data from 253 census tracks in Indianapolis SMSA, finds homogeneity by education. He also discovers that population density, race, and home ownership are the determinants of heterogeneity. Gramlich and Rubinfeld (1982) compare the variance of local spending demands within a community with those throughout the state. The authors argue that if there is Tiebout sorting, the within community variance will be significantly smaller than the entire statewide sample variance. Their data from Michigan indeed show that this is the case.

Although it is important to investigate the degree of homogeneity, the true test of the Tiebout hypothesis should involve testing the assumptions or/and the implications of the model.

In the Tiebout world households weight up the value of local services and the burden of local taxes and cross jurisdictional boundaries to get the most desirable package of local taxes and services. Tiebout (1956; p. 418) asserts:

> *Given these revenue and expenditure patterns, the consumer-voter moves to that community whose local government best satisfies his set of preferences. The greater the number of communities and the greater the variance among them, the closer the consumer will come to fully realizing his preference position.*

Therefore, if the Tiebout model catches some truth about actual residential choice behavior of households, we should expect to find more sorting in the areas where there are more local governments and where local governments supply more diverse set of bundles of local public goods.

The purpose of this paper is to tests the following two sorting implications of the Tiebout model:

1. The larger is the number of competing jurisdictions the more homogeneous each jurisdiction will be.
2. The presence of exogenous factors that reduce the ability of the local government to differentiate its public service bundle from the service bundles of competing governments decreases the homogeneity of local jurisdictions.

Theoretical models of the Tiebout equilibrium (Epple, Filimon, and Romer (1984), Epple and Romer (1991), Epple and Platt (1998)) also imply homogeneous sorting. In these models a necessary condition for equilibrium is stratification: each community is formed of families with incomes in a single interval. Thus, increasing the number of jurisdictions should reduce the length of each interval and, therefore, heterogeneity within each community.

Previous studies of the relationship between the homogeneity of local jurisdictions and their number have produced mixed results (Schmidt (1992), Stein (1987), Munley (1982), Eberts and Gronberg (1981)). Most of them tested the relationship between the number of school districts (single purpose local governments) in the area and income heterogeneity of each school district.

Eberts and Gronberg (1981), using the sample of school districts within 33 Metropolitan Statistical Areas (MSA's), discover that an increase in the number of school districts promotes within community income homogeneity. Examining data from 129 SMSAs, Schmidt (1992) confirms the result of Eberts and Gronberg. She finds that an increase in the number of school districts leads to less income heterogeneity in the typical district.

Assuming linear and log-linear functional forms of the demand for community services and normal and log-normal distribution of the set of optimal output levels in each school district, Munley (1982) measures the relative dispersion of the frequency distribution of voter demand. He discovers that greater number of school districts reduces the relative dispersion of the frequency distribution of voter demand characterizing any particular district.

Only Stein (1987) uses multiple purpose governments (municipalities and townships) to study the relationship between heterogeneity and the number of local governments. He is also the only one who does not rely on household income as the only proxy of household preferences for local public goods. Stein investigates the heterogeneity of municipalities with respect to personal income, race, housing, occupation, education, and age. He finds that an increase in the number of municipalities significantly reduces the heterogeneity of local communities only with respect to education. For other measures of heterogeneity the results are insignificant.

The interpretation of the coefficient on the number of municipalities in each MSA is, however, marred by the fact that Stein entered this variable in the model together with the population of the metropolitan area. In that case, the number of municipalities may serve

as a proxy for the average population of the jurisdiction. Holding MSA's population constant, changing the number of local municipalities effectively changes the average population.

The presence of factors, such as grants from upper levels of governments, tax and expenditure limitations (TELs), etc., that can reduce the ability of local government to provide a bundle of public services that is differentiated from the bundles provided by competing governments should decrease the degree of homogeneous sorting. This implication of Tiebout's model has received some empirical support. Both Eberts and Gronberg (1981) and Stein (1987) include state aid as a determinant of population heterogeneity in their models. Eberts and Gronberg find that an increase in the percentage of school revenue per student received from the state increases within district income inequality. Stein, however, does not discover statistically significant effect of compensating state aid on any of his six measures of within community heterogeneity.

Aaronson (1999), analyzing the effect of school finance equalization reform, obtains results indicating that in states where the reform took place there is less income sorting among poor and among rich in low property value jurisdictions. Aaronson also finds that the presence of TELs decreases high-income sorting in high property value districts.

Our purpose is to expand the evidence concerning the idea that the number of local governments in U.S. metropolitan areas influences the degree of homogeneous sorting among households. Consistent findings in this and previous studies would provide a body of preliminary evidence in support of the theory that has taken a central place in local public finance.

In the next section we describe the proposed empirical test and the departure of the present study from previous papers. Data issues related to the empirical test are discussed in Section 3. The results of the test are in Section 4. Section 5 contains a brief summary of our study.

2. THE PROPOSED EMPIRICAL TEST

This study tests two implications of the Tiebout hypothesis. **1.** Tiebout's notion that residential mobility and choice can lead to an efficient level of government service implies that local jurisdictions should be less heterogeneous with respect to the demand for local public services when there are more of them. When there are more jurisdictions to choose from, there will be better sorting. **2.** Decreased discretion of local governments over their tax and expenditure decisions reduces their ability and incentives to provide a fiscal package that is targeted at the particular group of households. Therefore, within-community heterogeneity should rise when there is an exogenous decrease in the ability of the local government to set its tax and expenditure levels.

The present study differs from the existing literature in the following respects.

First, the issue of "statistical" sorting is addressed. Two different forces can cause the increase in within-community homogeneity, as the number of local communities rises. One is Tiebout sorting: when there are more different jurisdictions in the area households can more fully realize their preferences for public services and jurisdictions become more homogeneous.

Another is "statistical" sorting. Even without the Tiebout mechanism we would expect that the populations of local communities would be more homogeneous when

there are more communities in the area. Take the population of some region, say MSA. Other mechanisms, unrelated to the demand for local public goods (zoning, racial and class antipathies, available housing stock, etc.)[2], imply that within MSA households do not locate randomly. Instead they form groupings that to some extent share similar characteristics: income, race, education, occupation. These are also the traits that we use to proxy for the demand for public services.

Given this nonrandom location of households, the division of MSA into smaller communities is likely to result in jurisdictions with populations that do not replicate the distribution of the population of the entire MSA. Such communities will be more homogeneous. The more jurisdictions we divide the MSA into the more homogeneous each jurisdiction is likely to be. Thus, the same relationship between the number of jurisdictions in the area and the degree of homogeneity of each jurisdiction may hold even if the Tiebout mechanism does not work. Therefore, disentangling the Tiebout sorting effects from 'statistical sorting' effects becomes a necessary part of the empirical tests of the Tiebout hypothesis.

None of the researchers that have previously investigated the relationship between homogeneity and the number of local communities have taken account of statistically generated sorting. To disentangle Tiebout sorting from 'statistical sorting' we propose the following. Instead of using only the number of local governments as an independent variable that controls for sorting, we introduce the number of governments, the area of

[2] In the literature on urban residential location it is a common finding that in a competitive land market bidders arange themselves acoording to the slope of their bid rent gradient (Beckmann (1974), Henderson (1985)). If the slope of the bid rent gradient is a function of income, than these models predict sorting by income. Kern (1981) shows that complete segregation by race can exist when racial preferences are incorporated into location decisions.

MSA, and the number of governments per capita, to separate these two effects. Ideally, we would like also to control for the population of MSA, but it is not possible[3].

According to the logic given in the several preceding paragraphs, communities are expected to be more homogeneous in the areas where there are more of them for two reasons. First, the choice is better and therefore households sort themselves more effectively. Second, in such areas there are more communities per capita, smaller average population per jurisdiction. Smaller populations are likely to encompass smaller ranges of variables that we use as proxies for household preferences for local public services. Therefore, these communities may be more homogeneous with respect to these variables.

Thus, two metropolitan areas with the same number of jurisdictions per capita should have the same degree of statistical sorting. Then MSA with more jurisdictions (and therefore larger population, to keep the number of jurisdictions per capita constant) will have more homogeneous communities only if there is active Tiebout sorting.

On the other hand, two metro areas with the same number of jurisdictions should have the same degree of Tiebout sorting. In this case, MSA with more jurisdictions per capita (and therefore smaller population, to keep the number of jurisdictions constant) will have more homogeneous communities only if statistical sorting effects are present.

Unfortunately, we can not control for the population size of MSA. However, this circumstance should, if anything, work against our hypotheses. Logically it seems to be more difficult to detect Tiebout sorting when metro area that has more communities is also more populous, than in the case when MSA with greater number of municipalities

[3] For instance, we need to control for the number of jurisdictions per capita. If we keep the population of MSA constant then we can not change the number of jurisdictions, because it will also change the number of jurisdictions per capita. Therefore, whenever we increase the number of governments we at the same time increase the population of the MSA, to keep the number of governments per capita constant. On the

has the same population. It also seems to be more difficult to find statistical sorting if metro area, that has more jurisdictions per capita, has smaller and, therefore, more disperse population, than in the case when MSA with more communities per capita has the same and, therefore, more dense population.

Roughly speaking, to find Tiebout sorting and statistical sorting we compare metropolitan areas with different populations. Everything else being equal, if statistical sorting is present, then MSA with larger population will have fewer jurisdictions per capita, larger average population and, therefore, more heterogeneous communities. Everything else being equal, if Tiebout sorting is present, local governments should be less heterogeneous in more populous metropolitan area that has more of them.

The geographical size of MSA may also play important role here. Given the number of municipalities, larger MSA will probably have higher transportation costs. If these costs are important determinants of residential choice, bigger metropolitan area will have more heterogeneous communities. Also, since we keep the number of local governments constant, in larger metropolitan area jurisdictions are farther apart from each other. If households consider as substitutes only jurisdictions that are within a certain distance, then the presence of municipalities that fall outside this distance does affect households' choice set. If this is true, then households living in two different metro areas, that have the same number of governments but differ in size, will have different degree of choice.

Second, the Tiebout mechanism implies that in areas with richer local fiscal bundles choice communities are more homogeneous with respect to demand for public

---

other hand, when we increase the number of municipalities per capita we at the same time reduce the population of MSA, to keep the number of municipalities constant.

services. Households living in such jurisdictions demand similar levels of public goods. Of course, the inability to measure the households' demand directly forced researchers to choose observable proxies. The majority of previous studies of the relationship between homogeneity and the Tiebout choice relied on income as the only measure of households' preferences for community services (Aaronson (1999), Schmidt (1992), Eberts and Gronberg (1981)). For this study, the measurement of residential heterogeneity includes using multiple indicators of individual preferences within each metropolitan community. In addition to household income four other population traits will be examined (education, occupation, household type, and race) as proxies of preferences for local public services.

Third, there is no general agreement in the literature on how to measure homogeneity and even on how to define it. In most previous studies income inequality is measured rather than income heterogeneity. As was demonstrated by Pack and Pack (1977) for the Gini coefficient, communities with extremely unequal income distributions can still be quite income homogeneous. Consider a community where everybody has the same income. For that community the Gini coefficient is equal to zero. This community has equal income distribution, according to Gini, and it is income homogeneous. Now, suppose in a community 90% of the households (with the same income) have only 10% of community's income and 10% of the households have 90% of community's income. The Gini coefficient for this community approaches to 1 - great income inequality. However this community is quite homogeneous with respect to income - 90% of households have the same income. So, it is possible that community with great income inequality, as measured by Gini, is homogeneous with respect to income.

Ideally we would like to have a measure which assumes a value of zero, absolute homogeneity, if all households fall into one category and a value of 1 if no two households are alike.

Lieberson's (1968) diversity-in-population measure possesses such a property and also has a very appealing interpretation. This measure computes the probability that randomly paired members of a jurisdiction will be different on a specified characteristic. In the simplest case of only one population trait, such as, for example, education, this measure, A_w, is equal to one minus the sum of squares of the proportions of the total population affiliated with each category, such as high school diploma, college degree, etc.:

$$A_w = 1 - \sum_i^n (s_i^j)^2$$

where s_i^j is the share of the population in jurisdiction **j** who identify themselves as a given education category **i , n** is the number of categories, and $\sum_i^n (s_i^j)^2$ is the Herfindahl index. If everyone has the same education, A_w would be zero. If every resident has a different education, then the index would be one.

Lieberson's measure can also be computed for two or more characteristics together, such as education, occupation, household type, etc. Therefore with its help we can describe the heterogeneity of city populations with respect to several characteristics simultaneously. The computation formula for A_w in this case is: $A_w = 1 - \left(\sum_{k=1}^{K} Y_k^2 \Big/ V \right)$, where Y_k is the proportion of the population falling within a given category within each of the variables, V is the number of variables, and K is the total number of categories within

all of the variables. Now we interpret A_w in the following way: if all households in a city are randomly paired, Lieberson's measure indicates the average proportion of mismatch between pairs on the characteristics under study[4].

Although Lieberson's index is a good measure of homogeneity for some qualitative variables, such as race and household type, it is less appropriate for variables that have a natural ranking of categories, such as education and income. This measure does not take into account the ranking of categories. It is just a probability that two randomly pared households have different traits.

For variables which have categories that can be ranked we seek a measure of heterogeneity that incorporates the ranking. It is impossible to meaningfully define cardinal distances between variable classifications, for qualitative as well as quantitative variables. Therefore we need an index that takes into account the distance between categories only in terms of rankings, but does not quantify distances between categories.

For example, suppose a variable has five categories that can be ranked. Then community with its population equally divided between the first and the second categories should have a higher homogeneity index than community with population equally divided between the first and the third categories.

An index proposed by Leik (1966) has the required properties. Leik's index takes a value of 0 when all individuals fall in one class. Its maximum value, 1, occurs when the community is evenly divided among its extreme classes. Since distance between categories is defined only with respect to ranking, not absolute difference, the index is

[4] For example, if A_w is equal to 0.72, it means that if all households of a particular city are randomly paired, on average they will differ with respect to 72% of characteristics specified.

defined on the cumulative frequency distribution of the population across variable categories: denoting the index value by C, it is defined as:

$$C = \left[\left(2\sum_i d_i \right) \Big/ (n-1) \right]$$

where $d_i = CF_i$ if $CF_i \leq 1/2$ and $d_i = 1 - CF_i$ if $CF_i > 1/2$; $CF_i = (1/N)\sum_{j \leq i} f_j$ (cumulative relative frequency); N = the number of observations; n = the number of categories. When all observations fall into one category, CF_i equals 0 until that category and 1 afterwards, making every d_i equal to 0 and C equal to 0. On the other hand, when all observations are divided between extreme categories $\sum_i d_i = \frac{1}{2}(n-1)$ and C = 1. When population is equally divided across all categories the value of index depends on the number of categories. For example, C is equal to 2/3 in case of three and four categories, and it is equal to 3/5 when there are 5 categories. The value of the index is a function not only of the number of intervals into which the population is divided but also of relation of the intervals to each other.

To avoid the confusion of heterogeneity with inequality that plagued some previous papers (Aaronson (1999), Schmidt (1992), Grubb (1982), Eberts and Gronberg (1981)), we use Leik's (1966) index of ordinal dispersion and Lieberson's (1968) diversity-in-population measure.

Fourth, the issue of the endogeneity of the number of jurisdictions is addressed. Among previous studies only Schmidt (1992) and Eberts and Gronberg (1981) recognized this problem. Others (Stein (1987), Munley (1982)) assumed the number of local governments to be exogenous.

It has been shown that the factors that affect the degree of within-community heterogeneity, such as metropolitan area wide heterogeneity, also affect the number of local governments in the area. Fisher and Wassmer (1998) and Alesina, Baqir and Hoxby (2000) find that after controlling for political, historical, and institutional factors, variations in the characteristics that affect demand for local government services do influence the number of local governments.

To deal with endogeneity of the number of local jurisdictions Eberts and Gronberg (1981) applied the instrumental variables technique. As instruments they used state dummies. State dummies, however, do not satisfy the property of good instruments, because they also affect the within-community homogeneity.

In order to address the endogeneity problem instruments that are unlikely to correlate with the homogeneity of jurisdictions should be employed. Legal and physical barriers to the creation of local governments represent such instruments. Legal institutions that allow the number of local communities to change represent the primary means by which local government structure can respond to changes in economic and other factors.

Annexation is the main instrument by which existing jurisdictions expand their boundaries. Incorporation is the procedure that produces new jurisdictions. Fisher and Wassmer (1998) have shown that the differences in annexation and incorporation laws do affect the number of local governments. Since these laws are most probably not correlated with the sorting behavior of households, they represent the instruments we need.

Fifth, instead of using single purpose governments (school districts), as a unit of analysis we use multiple purpose governments (municipalities and townships)[5]. Although most of the previous empirical tests have investigated school districts, in terms of local spending municipalities are equally important. For example, the 1990 issue of "Significant Features of Fiscal Federalism" (ACIR 1990) demonstrates that, for all tabulated years (1955, 1960, 1965, 1970-1988), total direct expenditures are roughly the same for school districts and for multiple purpose local governments.

We believe that multiple purpose governments are more appropriate for our analysis. Ideally we would like to analyze the territories within MSA that provide their residents with the same bundle of services from the same array of local governments (school districts, special districts, and municipalities). Since such data are not available, we think that the choice of multiple purpose governments is the closest to this ideal. According to Tiebout, households locate in the jurisdiction that supply the most desired bundle of public services, not one particular service (schooling).

Only Stein (1987) uses municipalities as a unit of analysis. His results are, however, mostly negative. He did not find any relationship between the number of local governments and within-community heterogeneity measured with respect to personal income, race, housing, occupation, and age. Dowding, John, and Biggs (1994) conclude: "Stein's work is the most comprehensive and therefore his negative results must represent the current state of knowledge." To the contrary, in this paper we present evidence which is truly corroborative with Tiebout.

[5] Census of Governments, 1987 defines municipalities and townships as: organized local governments authorized in state constitutions and statutes and established to provide general government for specific concentration of population in a defined area (municipality) or for areas defined without regard to population concentration (township).

Sixth, to see if the sorting is less effective in the areas where governments face exogenous constraints on their ability to set tax and expenditure limits, we include two measures of such constraints in our analysis: grants from upper levels of governments and tax and expenditure regulations.

To empirically test the premise that the greater is the number of local jurisdictions in the metropolitan area (*ceteris paribus*) the more homogeneous each jurisdiction will be with respect to demand for local public goods, we must control for regional, economic, demographic, and size characteristics. Our empirical model is:

$$
\begin{aligned}
H_i^k = H [&(\#\text{ of local governments})^k, (\text{geographical size of MSA})^k, \\
&(\#\text{ of local governments per capita})^k, (\text{tax and expenditure limitations})^k, \\
&(\text{grants from upper level governments})^k, \\
&(\text{metro-wide heterogeneity})^k, \\
&(\text{economic characteristics})^k, (\text{demographic characteristics})^k, \\
&(\text{state dummies})]
\end{aligned}
\qquad (1)
$$

where $\mathbf{H_i^k}$ is the degree of heterogeneity of local jurisdiction **i** in metropolitan area **k**.

3. DATA ISSUES

We consider multiple purpose local governments: municipalities and townships. Our dependent variables are the Leik's indices of within-jurisdiction heterogeneity for education, household income, and occupation, and Lieberson's diversity in population measures for race, household type, and all five characteristics combined. For the most part, independent variables are measured at the MSA level.

Table 1 lists variables and their sources.

Table 1

Descriptive Statistics

Variable	Mean	Std Dev
Lieberson's measure household type	0.708	0.046
Leik's index education	0.373	0.074
Leik's index occupation	0.502	0.068
Leik's index household income	0.449	0.058
Lieberson's measure race	0.126	0.154
Lieberson's measure combined	0.679	0.042
Lieberson's measure household type, MSA	0.741	0.015
Leik's index education, MSA	0.425	0.045
Leik's index occupation, MSA	0.526	0.031
Leik's index household income, MSA	0.491	0.028
Lieberson's measure race, MSA	0.279	0.116
Lieberson's measure combined, MSA	0.731	0.035
# of local governments	42.16	70.54
MSA area, m^2	2054.6	2550.8
# of local governments per capita	0.0001	0.0001
grants	0.272	0.127
Herfindahl index of land concentration, MSA	0.020	0.032
percentage of black MSA population	0.099	0.099
percentage of MSA population below the age of 18	0.258	0.031
percentage of MSA population age 65 and above	0.125	0.034
mean income, MSA	13247.1	2165.9
percentage of MSA population with income below poverty level	0.136	0.049
Herfindahl index of house value concentration	0.174	0.092

Notes: With the exception of the grants variable, which is from the Census of Governments, 1987, the variables are from the Census of Population and Housing, 1990.

The Tiebout model assumes that the residential choice of a household depends only on the demand of that household for local public goods. Of course, in reality the demand for local public goods is neither the only nor probably the major determinant of the residential choice. Perhaps the most important determinant of that choice is employment. By analyzing only municipalities located within MSAs, we believe that we effectively hold employment opportunities of households constant. We, thus, assume that

entire MSA represents one labor market and the residential choice of any household within the MSA does not affect employment opportunities of that household. This assumption rests upon the fact that the Bureau of Census defines MSA as "a large population nucleus, together with adjacent communities that have a high degree of economic and social integration with that nucleus", and economic integration is measured, in part, by the level of commuting.

To separate Tiebout sorting effects from "statistical sorting" effects we include the number of jurisdictions, MSA area, and the number of jurisdictions per capita among regressors. . Holding the number of local governments per capita and the area of MSA constant, the changes in the number of communities represent the changes in the Tiebout choice. Holding the number of communities and the area constant, the changes in the number of communities per capita represent the effects of "statistical" sorting.

The presence of grants from upper government levels equalizes the spending opportunities of jurisdictions, impairing their ability to provide differentiated public service bundles. Therefore, we expect less sorting in metropolitan areas where governments have greater reliance on grants. We measure grants as the share of the sum of grants received by all municipal and township governments in the sum of all total general revenues collected by all multiple-purpose local governments within SMSA.

Another variable that we use to test the implication that less discretion over tax and expenditure decisions reduces sorting, is tax and expenditure limitations (TEL). State governments impose many different types of TELs on local governments. For example, states may place limits on local (1) property tax rates, (2) property tax levies, (3) revenues, and (4) expenditures. Some of these TELs are more restrictive than others.

Joyce and Mullins (1991) separate state-imposed TELs into two categories: "nonbinding" and " potentially binding", based on the probability a TEL will restrict overall local taxing and spending power.

Given the diversity of TEL laws and the variation in their restrictive power, we classify TEL states in three ways. First, following Poterba and Rueben (1995) and Shadbegian (1998) this study treats all potentially binding limits uniformly. Our TEL variable is equal to one for all states that had potentially binding TELs in 1987, it equals zero otherwise[6].

Second, we follow Shadbegian (1998) by classifying TELs as stringent and nonstringent according to how fast they allow property taxes to grow. The state is classified as having stringent TEL (STRT=1) if it allows property tax levies to grow by less than 5 percent in 1987.

Third, to estimate the relative importance of the different types of TELs this study, following Preston and Ichniowski (1991), introduces two dummy variables. LEVLIM indicates states with property levy growth rate limits. REVLIM measures the presence of the various types of 'total revenue' growth rate limits.

Holding everything else constant, increase in the overall heterogeneity of the MSA's population should increase within-community heterogeneity. We assume that metropolitan area wide heterogeneity is not endogenous in this model. This assumption rests upon the belief that intermetropolitan migration is not primarily based on an individuals' search for their ideal bundle of local public goods. We believe that Tiebout forces have to work at the local level. That is why it is more appropriate to investigate

Tiebout sorting effects at the level of municipalities rather than at higher levels of governments.

The effectiveness of sorting depends not only on the degree of choice households have, but also on the household ability to exercise that choice. MSAs with higher costs of exercising the choice will have more heterogeneous communities. To control for some of these costs we introduce the following variables: land concentration, the share of the MSA population that is black , and the share of the MSA population with income below poverty level.

Hoxby (2000) suggests that MSAs where there are a lot of municipalities within a few minutes of most jobs the cost of exercising choice is low. On the other hand, in MSAs where one municipality encompasses the vast majority of residences and jobs and the commute to the nearest alternative municipality is long, the cost of being able to exercise the residential choice is high. Thus, we expect to find a positive coefficient on the land concentration variable. This variable is computed as the Herfindahl index, H, of municipal land shares within SMSA:

$$H^j = \sum_i^n (s_i^j)^2$$

where s_i^j is the share of jurisdiction **i**'s land and **n** is the number of jurisdictions in metropolitan area **j**.

In the presence of racial prejudice households that belong to racial minority groups can find it more difficult to choose the community with the level of public services they desire, because even if such communities are present the psychological

[6] TEL=1 for the following states: Arizona, Colorado, Oregon, Utah, Kansas, Minnesota, Washington, Delaware, Indiana, New Jersey, Ohio, California, Louisiana, Michigan, Idaho, Kentucky, New Mexico,

costs of living there may be too high. Therefore MSAs with greater share of racial minorities may have less homogeneous communities.

To exercise residential choice households must be mobile. In MSAs with less mobile population there should be less sorting. We include for each MSA the share of population with income below the poverty level as a proxy for the proportion of MSA households that could face limited mobility due to financial constraints. The greater is that share the more heterogeneous municipalities should be.

Demographic variables – the share of MSA population aged 65 or older, the share of MSA population under 18 years, and mean household income – proxy for the benefit from local public goods. Since households that value public goods more have greater incentives to sort, we expect these variables to affect the within-community heterogeneity negatively.

Municipalities with more homogeneous housing stocks are more likely to be populated by more homogeneous households. To control for that we use Herfindahl index of house value concentration.

The specific characteristics of the central city may affect the residential choice of its residents. To control for these effects we include dummy variable identifying central cities in all MSAs. To capture the influence state-specific fixed effects on sorting behavior of households we include state dummies (Wyoming being the excluded state) in our empirical model.

A Hausman specification error test for endogeneity confirms that variables the number of local jurisdictions and the number of jurisdictions per capita are endogenous to the factors influencing the heterogeneity of municipalities. We use two-stage least

---

Massachusetts, Mississippi, Missouri, Arkansas, North Dakota, Nevada, Rhode Island.

squares procedure to address the endogeneity problem. As instruments for the first stage we employ dummy variables reflecting the differences in annexation, consolidation, and incorporation laws[7].

To capture the significant variation among states in the laws and procedures governing local government structure, we divide the states into four groups reflecting the difficulty of annexation, into four groups reflecting the ease of incorporation, and into four groups reflecting the difficulty of consolidation. Dummy variables identifying each group are used in the instrumental equations. Most of these variables are significant in both instrumental regressions.

Our data are from the Census of Population and Housing, 1990, from the Census of Governments, 1987, and from State Laws Governing Local Government Structure and Administration, 1978. 11671 observations are used in our analysis.

4. ESTIMATION RESULTS

Table 2 contains the regression results for the heterogeneity measures for local communities in the 282 U.S. metropolitan areas in 1990. Two MSAs (one from Alaska and one form the District of Columbia) are excluded from the analysis, for lack of data on laws governing local government structure. Our dependent variables are the measures of heterogeneity with respect to different socio-economic characteristics. To measure the heterogeneity of a jurisdiction with respect to household type and race we use Lieberson's measure. For education, occupation, and income we employ Leik's index.

[7] These data are from State Laws Governing Local Government Structure and Administration, 1978.

Table 2

Estimates of the Sorting Equation using Two-Stage Least Squares

Dependent variables: Lieberson's heterogeneity measure (household type, race, and combined five) and Leik's heterogeneity indices (education, occupation, and household income)

	Household type	Education	Occupation	Household income	Race	Combined
	1	2	3	4	5	6
Constant	27.68*** (6.993)	4.714 (3.533)	52.92*** (3.329)	21.20*** (3.502)	12.50** (5.268)	41.22*** (2.834)
# of governments[1]	-0.025*** (0.006)	0.049*** (0.009)	-0.084*** (0.009)	-0.067*** (0.008)	-0.068*** (0.019)	-0.026*** (0.005)
The area of SMSA[2]	-0.007 (0.016)	-0.032 (0.024)	0.098*** (0.023)	0.021 (0.021)	0.084** (0.042)	0.032*** (0.012)
# of governments per capita[3]	-0.382 (1.517)	-7.717*** (2.127)	-9.856*** (1.956)	-6.364*** (1.978)	-1.513 (3.581)	-4.644*** (1.061)
Tax and expenditure limitations (TEL)	0.326 (1.725)	0.726 (2.682)	-0.327 (2.633)	-0.829 (2.227)	-6.628 (4.759)	-1.690 (1.352)
Stringent TEL	0.871 (1.838)	-0.992 (2.861)	-0.856 (2.808)	1.470 (2.370)	2.761 (5.075)	0.908 (1.442)
Property levy growth rate limits	-0.134 (1.804)	1.623 (2.808)	2.117 (2.756)	3.034 (2.327)	2.755 (4.981)	1.592 (1.415)
Total revenue growth rate limits	-0.616 (4.337)	15.28** (6.751)	-6.956 (6.628)	-4.198 (5.596)	3.297 (11.98)	1.683 (3.403)
Grants from upper levels of government	-0.567 (0.700)	1.175 (1.089)	-1.402 (1.069)	-1.506* (0.907)	-3.632* (1.943)	-1.027* (0.552)
Central city	2.136*** (0.224)	6.052*** (0.349)	0.025 (0.343)	3.021*** (0.289)	16.77*** (0.619)	4.827*** (0.176)
Metro-wide heterogeneity	0.762*** (0.094)	0.411*** (0.036)	0.335*** (0.036)	0.589*** (0.054)	0.485*** (0.028)	0.479*** (0.037)
% of SMSA population with income below poverty level	-5.331** (2.313)	7.643* (4.039)	-13.92*** (3.599)	-19.70*** (2.983)	0.245 (6.571)	-6.493*** (1.832)
Land concentration	-7.264*** (2.699)	20.74*** (4.293)	-7.859* (4.148)	5.963* (3.474)	-4.158 (7.599)	0.384 (2.138)
Percentage of black SMSA population	0.969 (1.333)	-2.978 (1.877)	0.430 (1.834)	-0.873 (1.526)	-6.532 (4.155)	-0.987 (1.156)
Mean household income[4]	-0.251*** (0.064)	0.384*** (0.097)	-0.357*** (0.094)	-0.216*** (0.082)	-0.498*** (0.168)	-0.199*** (0.047)
Percentage of SMSA population under 17 years	-18.42*** (3.447)	17.76 (0.976)	-21.87*** (4.981)	2.105 (3.994)	2.699 (8.456)	-4.099* (2.397)
Percentage of SMSA population age 65 or older	-12.56*** (2.538)	26.01*** (4.681)	-10.74** (3.804)	6.398** (3.062)	-6.441 (6.613)	-3.695** (1.888)
House value concentration index	-13.88*** (0.454)	-4.508*** (0.707)	-15.24*** (0.695)	-12.24*** (0.586)	-12.12*** (1.257)	-11.84*** (0.357)
Adjusted R^2	0.21	0.28	0.16	0.17	0.46	0.44

[1]The number of governments is measured by ten.

[2]MSA area is in thousands of square miles.

[3]The number of governments per one thousand residents.

[4]In thousands of dollars.

Standard errors are in parentheses. 11671 observations. Significance levels are indicated by asterisks as follows: * < 0.10; ** < 0.05; *** < 0.01

Lieberson's measure is also used for combined heterogeneity with respect to all five characteristics.

The results reported in table 2 provide persuasive evidence that within community heterogeneity with respect to the demand for local government services partly reflects the variation in the number of communities in the area.

As expected, the heterogeneity of a jurisdiction with respect to household type, occupation, income, race, and five characteristics combined relates negatively and significantly to the number of communities in the MSA. This finding supports the hypothesis that households sort themselves among local communities and sorting is more efficient in the areas with better residential choice.

Quantitatively this result indicates the following. In metro areas with ten more municipalities, on average, in each community the probability that two randomly paired residents will belong to different household types, will hold different occupations, will earn different incomes, and will belong to different race is lower by 0.03%, 0.08%, 0.07%, and 0.07% respectively. The coefficient in the second row of column 6 suggests that if all residents of a community are randomly paired, then in MSA that has ten more jurisdictions such a community will, on average, have pairs that are different on 0.03% less of specified characteristics.

However, our results suggest that in metro areas with greater residential choice communities are more heterogeneous with respect to education. Theoretically, there are forces that could lead households to sort themselves into communities that are heterogeneous with respect to the very same characteristics that we use as demand proxies. For example, peer group effects, job and skill complementarity are among those

forces. People with different education levels might find it beneficial to live together. This result is consistent with the predictions of the model developed by de Bartolome (1990), where peer group effects may cause communities to become more heterogeneous.

As can be seen from the fourth row of table 2, "statistical sorting" effects are present. The coefficient on the number of local governments per capita variable is negative and significant in all but two specification. Even if households do not care about local public goods, metropolitan areas with greater number of jurisdictions will have more homogeneous communities.

MSA area has the expected influence in three regressions. From the third row of table 2, we can see that bigger metropolitan areas have less sorting by occupation, race, and combined measure. This may be an indication that transportation costs are an important factor affecting residential choice. Also, this finding may suggest that substitutability of municipalities in the residents' choice set diminishes with the distance. Jurisdictions that farther away may be considered as being effectively out of the choice set.

Our measures of constraints on local government taxing and spending authority have produced mixed results. Out of four variables controlling for differences in tax and expenditure limitations among states only the variable measuring limits on total revenue growth rate is significant and in only one specification (education equation). In this case it has the expected positive effect: in states that impose limit on total revenue growth rate there is less sorting by education. The reason for the poor performance of TEL variables may lie in the fact that we use state dummies to control for state fixed effects. As a result, high correlation between these variables reduces significance by increasing the

variances[8]. Of course, another explanation may be that state imposed limits on local tax and spending decision do not reduce the variation set of local fiscal bundles offered by local governments. These limits may not reduce the ability of a municipality to supply the package of public services that is differentiated from the packages supplied by competing governments.

The presence of grants from upper government levels significantly affects heterogeneity measures in three regressions. However, in all three cases the effect is opposite to what we have expected. Local communities are, on average, less heterogeneous with respect to income, race and combined index in metro areas with larger share of grants from upper levels of government. This finding may indicate that, contrary to our initial assumption, grants are not equalizing. They may not reduce the variation in public service bundles offered by communities, but rather encourage it.

Metropolitan area-wide heterogeneity positively and significantly affects heterogeneity of communities within MSA in all six equations.

Our measures designed to control for the differences between MSAs in the costs of exercising the residential choice do not perform very well. Contrary to our expectations, the share of MSA population with income below poverty level negatively affects the heterogeneity by household type, occupation, household income and combined index. One possible explanation may be that, on the one hand, poor households may not have much flexibility of choice, on the other hand, not poor households have greater incentives to sort in metropolitan areas with larger poor population. This finding is consistent with the predictions of the models presented in Benabou (1993) and Durlauf

[8] When we used regional dummies to control for fixed effects most TEL variables had significant effect in most specifications, though not always of correct sign.

(1994), where complementarities in human capital investment induce occupational segregation and stratification.

The index of land concentration significantly influences heterogeneity by household type, education, occupation, and income. Only in the case of education and income it has the expected influence on sorting. MSAs with greater land concentration presumably have larger costs of exercising residential choice. Therefore, they exhibit less educational and income sorting.

Demographic variables – the share of SMSA population age 65 or older, the share of SMSA population under 18 years, and mean household income – we use as indicators of how much people desire the local public goods. Mean household income is negatively associated with within-community heterogeneity in five regressions. Our interpretation is that higher income households have greater demand for public goods and have more incentives to gather information on fiscal bundles offered by different jurisdictions. Also, for these households it is less costly to gather such information and to exercise residential choice. However, contrary to our expectations, municipalities in SMSAs with high mean income exhibit less sorting by education.

The share of SMSA population aged 17 or younger negatively affects the heterogeneity of communities by household type, occupation, and combined measure. Variable, the share of population aged 65 or older has significant influence on heterogeneity in five regressions. It has negative effect in household type, occupation, and combined equations and positive effect in education and income equation.

5. SUMMARY AND CONCLUSION

The number of jurisdictions in a metropolitan area is an important determinant of within-community heterogeneity. Two different forces can cause the increase in within-community homogeneity, as the number of local communities rises. One is Tiebout sorting. Another is "statistical" sorting. We found that, after controlling for statistical sorting, there is evidence that households actively sort themselves into communities that are relatively homogeneous with respect to the demand for local public goods. This finding has important practical implications. Since households sort themselves among local communities, state and federal programs that encourage the variation of fiscal bundles offered by local governments should be designed. Such programs will improve sorting and efficiency.

All researchers that investigated the relationship between homogeneity and the number of local communities have failed to control for statistically generated sorting. This study shows that statistical sorting is present. Even if households do not care about local public goods, metropolitan areas with greater number of jurisdictions will generally have more homogeneous communities. Thus, our results suggest that it is important to control for "statistical sorting" effects while one is investigating the degree of Tiebout sorting.

Our findings also indicate that grants from upper levels of government and tax and expenditure limitations may not be the factors that reduce the ability of the local government to differentiate its public service bundle from the service bundles of competing governments.

Thus, the results of this paper are corroborative with the Tiebout sorting hypothesis. This interpretation of the results is, however, based on the assumption that

socio-economic characteristics of residents can be used as proxies for their demand for local public services. Future research that will use more direct measure of the demand for local public goods may provide additional evidence on the sorting behavior of households.

REFERENCES

1. Advisory Commission on Intergovernmental Relations "Significant Features of Fiscal Federalism" Washington D.C.: ACIR, 1990.

2. Aaronson, Daniel. "The Effect of School Finance Reform on Population Heterogeneity" *National Tax Journal* **52**, no. 1 (1999): 5-29.

3. Alesina, Alberto, Reza Baqir and Caroline Hoxby. "Political Jurisdictions in Heterogeneous Communities" *NBER Working Paper No.7859, 2000.*

4. Beckmann, M. "Spatial equilibrium in the housing market", *Journal of Urban Economics* **1** (1974): 99-107.

5. Benabou, Roland. "Workings of a City: Location, Education, and Production." *Quarterly Journal of Economics* **108**, (1993): 619-652.

6. de Bartolome, Charles A. "Equilibrium and Inefficiency in a community Model with Peer Group Effects." *Journal of Political Economy* **98**, (1990): 110-133.

7. Dowding, Keith, Peter John and Stephen Biggs. " Tiebout: A Survey of the Empirical *Literature." Urban Studies* **31**, no.4/5 (1994): 767-797.

8. Durlauf, Steven N. "Spillovers, Stratification, and Inequality." *European Economic Review* **38**, (1994): 839-845.

9. Eberts, Randall W. and Timothy J. Gronberg. "Jurisdictional Homogeneity and the Tiebout Hypothesis." *Journal of Urban Economics* **10**, (1981): 227-239.

10. Epple, Dennis, Radu Filimon, and Thomas Romer. "Equilibrium among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice." *Journal of Public Economics* **24**, (1984): 281-304.

11. Epple, Dennis and Thomas Romer. "Mobility and Redistribution." *Journal of Political Economy* **99**, no. 4 (1991): 828-858.

12. Epple, Dennis and Glenn Platt. "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Preferences and Incomes." *Journal of Urban Economics* **43**, no. 1 (1998): 23-51.

13. Fisher, Ronald C. and Robert W. Wassmer. "Economic Influences on the Structure of Local Government in U.S. Metropolitan Areas." *Journal of Urban Economics* **43**, (1998): 444-471.

14. Gramlich, Edward M. and Daniel L. Rubinfeld. "Micro Estimates of Public Spending Demand Functions and Tests of Tiebout and Median-Voter Hypotheses." *Journal of Political Economy* **90**, no. 3 (1982): 536-560.

15. Grubb, Norton W. "The Dynamic Implications of the Tiebout Model: The Changing Composition of Boston Communities,1960-1970." *Public Finance Quarterly* **10**, no. 1 (1982): 17-38.

16. Henderson, J.V. "Economic Theory and Cities." *New York: Academic Press*, (1985)

17. Hoxby, Caroline. "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review* **90**, no. 5 (2000):

18. Joyce, Phillip and Daniel Mullins. "The Changing Fiscal Structure of the State and Local Public Sector: The Impact of Tax and Expenditure Limitations." *Public Administration Review* **51**, no.3 (1991): 240-253.

19. Kern, C. "Racial Prejudice and Residential Segregation: the Yinger model revisited" *Journal of Urban Economics* **10**, (1981): 164-173.

20. Leik, R.F. "A Measure of Ordinal Consensus." *Pacific Sociological Review* **9**, (1966): 85-90.

21. Lieberson, Stanley. "Measuring Population Diversity." *American Sociological Review* **34**, no. 6 (1968): 850-862.

22. Munley, Vincent G. "An Alternative Test of the Tiebout Hypothesis." *Public Choice* **38**, (1982): 211-217.

23. Ottensmann, John R. "Neighborhood Heterogeneity Within an Urban Area." *Urban Studies* **19**, (1982): 391-395.

24. Pack, Howard and Janet R. Pack. "Metropolitan Fragmentation and Suburban Homogeneity." *Urban Studies* **14**, (1977): 191-201.

25. Poterba, James and Kim Rueben. "The Effect of the Property-Tax Limits on Wages and employment in the Local Public Sector." *American Economic Review* **85**, no. 2 (1995): 384-89.

26. Preston, Anne and Casey Ichniowski. "A National Perspective on the Nature and Effects of the Local Property Tax Revolt, 1976-1986*." National Tax Journal* **44**, no. 2 (1991): 123-145.

27. Samuelson, Paul A. “The Pure Theory of Public Expenditure.” *Review of Economics and Statistics*,**36**, (1954): 387-389.

28. Schmidt, Amy B. “Private School Enrollment in Metropolitan Areas.” *Public Finance Quarterly* **20**, no. 3 (1992): 298-320.

29. Shadbegian, Ronald. “Do Tax and Expenditure Limitations Affect Local government Budget? Evidence from Panel Data.” *Public Finance Review* **26**, no. 2 (1998): 218-236.

30. Stein, Robert M. “Tiebout’s Sorting Hypothesis.” *Urban Affairs Quarterly* **23**, no.1 (1987): 140-160.

31. Tiebout, Charles M. “A Pure Theory of Local Expenditures.” *Journal of Political Economy* **64**, no. 5 (1956): 416- 424.