TIEBOUT BIAS AND THE DEMAND FOR LOCAL PUBLIC SCHOOLING

Daniel L. Rubinfeld, Perry Shapiro and Judith Roberts*

Abstract—Until recently estimates of demand functions for public goods were obtained (either with aggregate or micro survey data) using single equation estimation techniques. However, demand estimates may be biased when individuals' choices of communities are dependent upon the quantity and quality of public good provided. This paper spells out the nature of this bias (called Tiebout bias), and suggests an improved maximum-likelihood estimation technique. The technique is applied to a data set involving local public education in Michigan.

I. Introduction

THE estimation of household demands for public goods has long been a concern of public finance economists. Until relatively recently, demand estimates were obtained primarily using aggregate data and single equation estimation techniques. However, Goldstein and Pauly (1981) have argued that these estimates may be biased if individuals sort themselves into communities in part on the basis of local public sector activity. They illustrate the possibility of such bias —called Tiebout bias—using a model of demand in which income is the single explanatory variable. The model suggests that, under reasonable conditions, previous estimates of income elasticities obtained from aggregate data are likely to be biased upward.

A new approach to the estimation of demand for public goods, suggested by Bergstrom, Rubinfeld and Shapiro (1982), uses micro data collected from a survey of individuals. Like its predecessors based on aggregate data, Bergstrom, Rubinfeld and Shapiro's approach employs a single equation technique to estimate demand parameters. In comparison to most aggregate models, however, this approach produces income and price elasticities which are relatively low.

This paper presents an argument similar to that made by Goldstein and Pauly in the context of aggregate models, raising the possibility that micro-based estimates might also be subject to a kind of Tiebout bias. Simply put, the bias arises because single equation models of demand fail to account for the sorting of individuals among communities. We make the argument in the context of Bergstrom, Rubinfeld and Shapiro's model.

The micro approach has substantial promise because the data are rich enough to allow one to test for Tiebout bias and, to the extent that it is present, devise new and better estimation techniques. This paper provides a start in that direction by suggesting the conditions under which Tiebout bias might occur and by attempting to estimate the extent to which this bias may be present. The paper goes further by suggesting a model structure in which more detailed questions concerning the specification and estimation of the demand for local public goods can be answered. The structure is sufficiently broad to allow for the inclusion of political as well as economic determinations of the demand and supply of local public goods.

Section II contains a brief heuristic discussion of the Tiebout bias problem in the context of both micro and aggregate data. The argument is meant to illustrate both the source of the potential bias due to the community selection process as well as the direction of bias. In section III the theoretical econometric discussion of Tiebout bias in the micro model is presented. The theory both suggests the source of bias, and allows us to develop a consistent and efficient estimator of the demand function parameters. The estimator is similar in spirit to the two-stage estimators proposed by Heckman (1979), but substantially different in practice: our estimator involves a least-squares first-stage and an ordinal multinomial probit second stage and the solution is obtained using an

Received for publication March 6, 1986. Revision accepted for publication November 10, 1986.

*University of California, Berkeley; University of California, Santa Barbara; and California State University, Long Beach, respectively.

The authors wish to thank Ted Bergstrom, Valerie Bencivenga, George Borjas, Stephen Bronars, Bob Deacon, Robert Inman, Bill Parke, Paul Ruud, and two anonymous referees for their advice and discussion of this paper. We especially thank George Olmsted who raised the issue of selectivity bias in the context of our work on the estimation of the demand for public goods. This research was supported by grants from the National Science Foundation and the Center for Real Estate and Urban Economics, University of California, Berkeley.

iterative procedure. Section IV contains the empirical analysis of the demand for local public schooling with and without Tiebout bias corrections. The results make it clear that Tiebout bias can be an important problem, but that estimates of its importance are sensitive to the model of community choice and public goods choice. Some brief concluding comments appear in section V.

II. The Tiebout Bias Problem—A Heuristic Description

Demand functions for public goods such as education have historically been estimated using aggregate data and single-equation estimation procedures. A typical approach involves relating the aggregate outcome in terms of dollars per pupil of school spending to the indicators of the demographic and economic composition of the relevant populations. In order for demand functions to be inferred from such data, a political theory, typically the "median-voter" model, is invoked to relate the expenditures of a jurisdiction to the characteristics of its population. The "median-voter" is usually taken to be the individual with the median income, residing in a house with median house value in a community with no renter population.

Goldstein-Pauly (1981) have raised serious questions concerning the application of the median voter concept. Their argument builds upon the notion that the population of each jurisdiction should not be taken as exogenous; instead individuals will sort themselves among jurisdictions through migration. To the extent that the sorting is based on the level of public expenditures rather than exogenous factors, the usual estimates of price and income elasticities will be biased upward.

The argument applies directly to median voter models which utilize aggregate data. Goldstein and Pauly suggest, however, that the use of individual (micro) data could eliminate the Tiebout bias problem. Thus, it would be natural to believe that the micro-based estimates of the demand for public goods of Gramlich and Rubinfeld (1982) and Bergstrom, Rubinfeld and Shapiro (1982) are not Tiebout-biased.[1] However, the Tiebout bias problem in the micro context has nothing to do with the median voter model or with the use of median income as a variable in the demand equation. Rather, it involves a direct application of the argument that when selection bias is a possibility (because of community sorting), the error term in the demand equation (conditional on the explanatory variables) will no longer have zero expected value.

The nature of the Tiebout bias problem can be seen with reference to a simple example. Assume that income is the sole determinant of spending demand and that half of all individuals have income $x_1 = 2$ and half have income $x_2 = 4$. The desired level of spending of individual i is $E_i = .5x_i + \epsilon_i$, where ϵ_i is assumed to have a uniform distribution on the interval $[-1, +1]$. Thus, the demand distribution of income group $x_1 = 2$ is given by a uniform density between 0 and 2, while the demand distribution for income group $x_2 = 4$ is a uniform density between 1 and 3.

There are two communities. Community 1 supplies 1 unit of the public good ($A_1 = 1$) and community 2 supplies 2 units ($A_2 = 2$). Individuals are assumed to reside in the community which supplies the level of public expenditures closest to their individual demands. Therefore, one-quarter of individuals with income x_1 will choose to live in community 2 and one-quarter of individuals with income x_2 will live in community 1 ($E(\epsilon|A_1, x_1) = E(\epsilon|A_1, x_2) = -0.25$, and $E(\epsilon|A_2, x_1) = E(\epsilon|A_2, x_2) = 0.75$). By construction, this distribution of individuals is both a Tiebout equilibrium, in which no one will migrate from one community to the other, and a median voter equilibrium, in which as many people in each community want more spending as want less spending.

With the two levels of community expenditures and the distribution of residents, the relation between actual and desired expenditures is $A_i = E_i + v_i$, where v_i is the difference between actual and desired expenditure. Employing the demand specification (with the slope parameter $b = 1/2$ to be estimated), this equation can be rewritten $A_i = bx_i + v_i + e_i$. By construction, however, x and v are correlated—in particular, $E(v|x_1) > 0$ and $E(v|x_2) < 0$. Therefore a regression of A_i on x_i would result in a biased estimator of b. In this case, since $\mathrm{Cov}(x_i, v_i) = -1/4$, the plim of the least squares estimator of the income elasticity of

[1] Ladd and Christopherson (1983) have argued to the contrary in the context of the Gramlich-Rubinfeld paper, while Olmsted (undated) has done the same with respect to the Bergstrom-Rubinfeld-Shapiro paper.

demand will be 1/4, an underestimate of the true slope parameter.

Of course, this story is not complete. For one thing, a similar story could be told about the price elasticity of demand as well as other demand parameters. For another, a more general model would allow for the level of public spending to be endogenous. We consider the more general model in section III, while at the same time illustrating how consistent parameter estimates can be obtained.

The Tiebout bias problem, therefore, can arise when either micro or macro data are being utilized. To illustrate the nature of the problem in more detail we have chosen to focus on the micro approach to the estimation of education demand functions as applied by Bergstrom-Rubinfeld-Shapiro (hereafter BRS). We stress, however, that the general approach to the specification of demand models has broader application—to the estimation of the demand for all locally provided public goods (including air pollution and other neighborhood attributes).

III. Theoretical Analysis

Assume that the demand for public school expenditures is given by

$$E_i = \beta_0 + x_i'\beta + \epsilon_i \tag{1}$$

where E_i is the logarithm of individual i's desired per pupil expenditure on public schools; x_i is a $k \times 1$ vector including socio-economic and demographic characteristics of the individual (including income, tax-price, children in school and race, for example) and school district characteristics; β_0 is a constant; and β is a $k \times 1$ vector of demand parameters. The random variable ϵ is distributed $N(0, \sigma_\epsilon^2)$ and is assumed to be uncorrelated with all personal characteristics.

Not all individuals within a jurisdiction will get to consume the level of school expenditures that is desired, for a host of reasons relating to the fact that the pure model of Tiebout sorting is not descriptive of the real world. We represent the difference between the actual provision of per pupil spending and desired spending by

$$A_i - E_i = \gamma_0 + x_i'\gamma + u_i \equiv v_i \tag{2}$$

where A_i is the logarithm of the jurisdiction's per pupil spending (actual spending) on education, γ is a $k \times 1$ vector of sorting parameters, and u_i is a random disturbance term, with u distributed $N(0, \sigma_u^2)$.[2] v_i is defined as the difference (mismatch) between actual and desired expenditures.

Note that the variables that explain sorting can include some of the demand-determining variables —e.g., income can affect demand and mobility as well. However, not all demand variables need appear in (2), since some of the x_i's can be constrained to have zero coefficients. Furthermore, there may be variables in (2) that do not affect demand. In this case the relevant demand parameters, β's, are zero. We assume in our theoretical discussion that education is a pure public good so that A does not vary among individuals within a given jurisdiction. This assumption can be relaxed in the empirical work without much difficulty. To keep the analysis as clear as possible we have chosen to omit jurisdiction subscripts.

The relationship between v_i and x_i is a complex one that might be related to the process by which individuals locate themselves among communities and the process by which A is politically determined. For example, we might expect v_i to be low in absolute value for recent movers who had some selection among public service bundles in making their move. It is the fact that γ may be nonzero as well as the fact that ϵ and v might be correlated that increases the Tiebout bias problem.

Of course, v_i is not directly observable. However, BRS utilized the responses to a survey of 2001 Michigan voters conducted by Courant, Gramlich and Rubinfeld (1980) to obtain information about v_i. Each survey respondent was asked whether he or she wanted more, about the same, or less expenditures on public education, as well as his or her individual characteristics. A response of "more" was assumed to be made if the level of educational expenditure was sufficiently smaller than the desired level.[3] "Same" was the response if

[2] To the extent that A_i was determined in general equilibrium by the public choice of individuals within jurisdictions, the x vector would have to reflect the public choice process as well. As long as individuals take public spending levels as given when making their micro choices (the typical assumption in demand studies), the estimation procedures which follow remain unchanged.

[3] A "more" response was recorded if and only if the respondent said more and then answered "yes" to a follow-up question: Your taxes will go up if there are larger expenditures; do you still want more?

the actual level was sufficiently close to the desired amount, while "less" was the answer if expenditures were substantially larger than the desired level. These qualitative ranges can be specified in terms of the random variable v and a threshold level δ by the following three equations:

$$\text{More if } v < -\delta \tag{3}$$
$$\text{Same if } -\delta < v < \delta \tag{4}$$
$$\text{Less if } v > \delta. \tag{5}$$

Substituting from (1) and (2) the conditions then become:

More if $E_i > A_i + \delta$ or

$$\epsilon_i > A_i - \beta_0 - x_i'\beta + \delta \tag{6}$$

Same if $A_i - \delta \leq E_i \leq A_i + \delta$ or

$$A_i - \beta_0 - x_i'\beta - \delta \leq \epsilon_i \leq A_i - \beta_0 - x_i'\beta + \delta \tag{7}$$

Less if $E_i < A_i - \delta$ or

$$\epsilon_i < A_i - \beta_0 - x_i'\beta - \delta. \tag{8}$$

Maximum-likelihood estimators of the parameters of the demand function can now be obtained if we make some additional assumptions about the distributions of the relevant variables. To allow for Tiebout sorting and its effects, we continue to assume that ϵ and x are uncorrelated. Likewise we assume that u is uncorrelated with x. However, we require x to be normally distributed and we allow for the possibility that A and ϵ (and therefore v and ϵ) are correlated. The complete set of assumptions is given as follows, with x a $k \times 1$ vector.

(x, A, ϵ) is distributed $N((\bar{x}, \bar{A}, 0), \Sigma)$, where:[4]

$$\Sigma = \begin{pmatrix} \Sigma_x & \Sigma_{xA} & 0 \\ \Sigma_{xA}' & \sigma_A^2 & \sigma_{A\epsilon} \\ 0' & \sigma_{A\epsilon} & \sigma_\epsilon^2 \end{pmatrix}.$$

Here Σ_x is a $k \times k$ matrix, Σ_{xA} is $k \times 1$ and 0 is $k \times 1$.

[4] Some of the variables in the vector x are discrete, rather than continuous so that the normality assumption is not appropriate. It would not be difficult, however, to expand the model to incorporate a vector z of discrete demand determinants. The individual z's would then be interpreted as mean-shifting variables, so that x would become $x(z)$, a vector of demand variables whose mean was conditional upon the values taken by the elements of the z vector. We have not included this expanded model in the text solely to simplify the presentation.

Using the normality of ϵ it follows directly that

$$\text{Prob (Less)} = \int_{-\infty}^{L} f(\eta)\, d\eta$$
$$\text{Prob (Same)} = \int_{L}^{M} f(\eta)\, d\eta$$
$$\text{Prob (More)} = \int_{M}^{\infty} f(\eta)\, d\eta,$$

where $f(\cdot)$ is a standard normal density function

$$L = [A_i - \beta_0 - x_i'\beta - \delta - E(\epsilon|x, A)]/\sigma_e \tag{9}$$

$$M = [A_i - \beta_0 - x_i'\beta + \delta - E(\epsilon|x, A)]/\sigma_e. \tag{10}$$

$E(\epsilon|x, A)$ represents the expectation and σ_e the standard deviation of ϵ conditional on x and A. The conditional variance is assumed to be constant.

The maximum-likelihood procedure in the BRS paper implicitly assumed that ϵ was independent of A as well as x. In this special case the likelihood function to be maximized is given below (F is the cumulative normal distribution function):

$$\begin{aligned}\mathscr{L} = &\prod_{i \in \text{Less}} F(\theta_0^L + \theta_1 A_i + \theta_2 x_i) \\ &\times \prod_{i \in \text{Same}} [F(\theta_0^M + \theta_1 A_i + \theta_2 x_i) \\ &\qquad - F(\theta_0^L + \theta_1 A_i + \theta_2 x_i)] \\ &\times \prod_{i \in \text{More}} [1 - F(\theta_0^M + \theta_1 A_i + \theta_2 x_i)].\end{aligned}$$

The θ's are parameters, with θ_0^L and θ_0^M reflecting the fact that the constant will vary by response category. Since

$$\theta_0^L = -(\beta_0 + \delta)/\sigma_\epsilon \qquad \theta_0^M = -(\beta_0 - \delta)/\sigma_\epsilon$$
$$\theta_1 = 1/\sigma_\epsilon \qquad \theta_2 = -\beta/\sigma_\epsilon,$$

consistent estimators of the demand parameters are given by

$$\hat{\beta}_0 = -(\hat{\theta}_0^L + \hat{\theta}_0^M)/2\hat{\theta}_1 \qquad \hat{\beta} = -\hat{\theta}_2/\hat{\theta}_1$$
$$\hat{\delta} = -(\hat{\theta}_0^L - \hat{\theta}_0^M)/2\hat{\theta}_1.$$

In the more general case in which Tiebout bias is a possibility, $E(\epsilon|x, A)$ is no longer identically zero and, in particular, no longer independent of x and A. This possibility is illustrated by the example in section II. In the example income is the only demand determining variable. For income $x = x_1 = 2$ and expenditures $A = A_1 = 1$, $E(\epsilon|A_1, x_1) < 0$ and for the same income and $A = A_2 = 2$,

$E(\epsilon|A_2, x_1) > 0$. Similarly $E(\epsilon|A_1, x_2) < 0$ and $E(\epsilon|A_2, x_2) > 0$. In this case the maximum-likelihood estimation procedure becomes substantially more complex. The correct procedure can be developed if we first evaluate $E(\epsilon|x, A)$ in the general case, as is done in the lemmas which follow.

LEMMA 1:

$$E(\epsilon|x, A) = \lambda\left[(A - \bar{A}) - \Sigma'_{xA}\Sigma_x^{-1}(x - \bar{x})\right]$$

where

$$\lambda = \sigma_{A\epsilon}\frac{\det(\Sigma_x)}{\det\begin{pmatrix} \Sigma_x & \Sigma_{xA} \\ \Sigma'_{xA} & \sigma_A^2 \end{pmatrix}}.$$

Proof: The proof follows from a standard result from multivariate statistics for the conditional expectation of a normal variable.

LEMMA 2:

$$E(\epsilon|x, A) = \lambda\left[(A - \bar{A}) - (x' - \bar{x}')(\beta + \gamma)\right].$$

Proof: From (1) and (2), $A = \beta_0 + \gamma_0 + x'(\beta + \gamma) + (u + \epsilon)$. Therefore, $\Sigma'_{xA}\Sigma_x^{-1} = (\beta + \gamma)$ and the result follows (recall that x is uncorrelated with both u and ϵ).

In this case, the bounds of the integrals which go into the likelihood function (equations (9) and (10)) change in the following manner (we have substituted for $E(\epsilon|x, A)$ using lemma 2):

$$L^* = \{A_i(1 - \lambda) - \beta_0(1 - \lambda) - \delta - x_i'[\beta(1 - \lambda) - \lambda\gamma] + \lambda(\bar{A} - \beta_0 - (\beta + \gamma)\bar{x})\}/\sigma_e \tag{11}$$

$$M^* = \{A_i(1 - \lambda) - \beta_0(1 - \lambda) + \delta - x_i'[\beta(1 - \lambda) - \lambda\gamma] + \lambda(\bar{A} - \beta_0 - (\beta + \gamma)\bar{x})\}/\sigma_e. \tag{12}$$

Because $\text{plim}(\bar{A} - \beta_0 - (\beta + \gamma)\bar{x}) = \gamma_0$, it follows that the probability limits of the BRS estimated demand parameters ($\hat{\beta}_0$, $\hat{\delta}$ and $\hat{\beta}$) are

$$\text{plim}\,\hat{\beta}_0 = \beta_0 - \frac{\lambda}{1 - \lambda}\gamma_0 \tag{13}$$

$$\text{plim}\,\hat{\delta} = \delta/(1 - \lambda) \tag{14}$$

$$\text{plim}\,\hat{\beta} = \beta - \frac{\lambda}{1 - \lambda}\gamma. \tag{15}$$

Equations (13)–(15) identify two possible sources of bias in the BRS estimators. One possibility is that the covariance between A (or v) and ϵ is non-zero (and thus λ is non-zero). This might be labelled Tiebout bias, and occurs because the demand for public goods affects the choice of residential communities, causing a non-zero correlation between A and ϵ. The second possibility is that the mismatch between actual and desired expenditures (v) is correlated with the demand determining variables, i.e., that some $\gamma_i \neq 0$ (when the corresponding $\beta_i \neq 0$). The first source of bias causes difficulties for the estimation of all the relevant parameters. However, in the special case in which $\gamma = 0$, $\hat{\beta}$ is a consistent estimator of β, but the constant β_0 and threshold parameter δ are inconsistently estimated by $\hat{\beta}_0$ and $\hat{\delta}$. If there is no Tiebout bias, $\lambda = 0$, and all the estimators are consistent.

On the basis of the more general formulation the following important theorems can be summarized:

THEOREM 1. When Tiebout bias is present and demand-determining variables are related to community sorting, $\hat{\beta}_0$, $\hat{\delta}$ and $\hat{\beta}$ as estimated using the BRS maximum-likelihood procedure are inconsistent.

THEOREM 2. When Tiebout bias is present and the set of sorting variables and demand determining variables are mutually exclusive ($\gamma_i = 0$ for $\beta_i \neq 0$) then the BRS technique will yield a consistent estimator of β.[5]

THEOREM 3. When Tiebout bias is not present, $\hat{\beta}_0$, $\hat{\beta}$ and $\hat{\delta}$ are consistent estimators.

The theorems suggest that the BRS demand parameters are consistently estimated when $\text{Cov}(A, \epsilon) = 0$, or when the actual level of politically determined per-pupil spending on education is uncorrelated with any omitted demand determining characteristics. One possible assumption under which $\text{Cov}(A, \epsilon) = 0$ might occur is if individuals are randomly assigned to communities and migration is not possible.[6] Consistency will also

[5] This is a special case of a more general result concerning the consistency of slope estimators in qualitative choice models. See Ruud (1983).

[6] We should note that in the case of perfect Tiebout sorting, $\delta = \tau = 0$, and there is no Tiebout bias. However, the BRS technique is not applicable in this situation.

occur when $\gamma_i = 0$ for $\beta_i \neq 0$, so that sorting and demand will not be confused. This assumption holds when the extent to which individual's demands deviate from the actual public service provision is uncorrelated with demand characteristics.

This technical discussion of the problem caused by endogenous community choice can be understood in a non-technical way. Consider the coefficient of A in the BRS maximum-likelihood procedure. The larger is θ_1, the greater the increase in the probability of a *less* response resulting from an increase in A (as we survey in communities with larger A's), holding x constant. However, if people select communities according to their tastes for public goods, the change in the probability of a *less* response to a change in A is smaller than if A's and x's were matched at random (that is, if community choice were not influenced by public good preferences). In fact, if there is sorting by public good preferences, there is a tendency towards a *same* response. This means that Tiebout sorting will systematically bias downward the absolute value of θ_1. In other words, without correcting for changes in the conditional expectation of ϵ, the effect of actual expenditure levels on response probabilities is systematically underestimated.

A similar argument implies that the absolute value of the estimated coefficients of x, the θ_2 vector, will be downward biased. However, since the maximum-likelihood estimates of the demand parameters involve the ratio of θ_2 to θ_1 we cannot say a priori whether the BRS demand parameters would be biased upward or downward.

Consistent Maximum-Likelihood Estimation

The Tiebout bias problem can be solved by recognizing that the demand and community matching functions represent a pair of simultaneous equations. In terms of equations (1) and (2),

$$E_i = \beta_0 + x_i'\beta + \epsilon_i$$

$$A_i = (\beta_0 + \gamma_0) + x_i'(\beta + \gamma) + \omega_i$$

where $\omega_i = u_i + \epsilon_i$. If u_i is normally distributed, the random variables ϵ_i and ω_i are bivariate normal with correlation r.[7] However, identification of the parameters in the demand equation is only possible if at least one variable in the model affects actual expenditures, A, but not demand, E. In effect this requires at least one $\beta_i = 0$ and the corresponding $\gamma_i \neq 0$. We will designate the set of variables excluded from the demand equation as the vector $\hat{x}$ and the corresponding parameters by the vector τ. With this notation, $A_i = (\beta_0 + \gamma_0) + x_i'(\beta + \gamma) + \hat{x}'\tau + \omega_i$. In this case, the log likelihood function for the observed pattern of survey responses is

$$\begin{aligned}\mathscr{L} = &\sum_{i \in \text{Less}} \log F(L^*) \\ &+ \sum_{i \in \text{Same}} \log[F(M^*) - F(L^*)] \\ &+ \sum_{i \in \text{More}} \log[1 - F(M^*)] - \tfrac{1}{2}\sum \log 2\pi\sigma_\omega^2 \\ &- \frac{1}{2\sigma_\omega^2}\sum\{A_i - [(\beta_0 + \gamma_0) + x_i'(\beta + \gamma) \\ &\qquad + \hat{x}_i'\tau]\}^2 \end{aligned} \tag{16}$$

where L^* and M^* are defined in equations (11) and (12). In this case $\lambda = r(\sigma_\epsilon/\sigma_\omega)$. Maximization of (16) with respect to the parameters of the demand and sorting functions yields consistent and efficient estimators.[8]

IV. Empirical Analysis

As in the original BRS paper, the data involved a subsample of 945 homeowners who responded to the question of whether they would like more, the same, or less spending on public education. The definitions of all variables utilized in the estimation procedure are given in table 1. (BRS contains a more detailed description of the data.)

The Tiebout bias question revolves around two central issues. The first is whether community choice is systematically affected by preferences for public goods. In terms of the model presented here the test for this bias is simple—whether or not the parameter λ is significantly different from zero. The second is the extent to which the variables that explain the matching of individuals to

[7] Thus, $\sigma_\epsilon^2 = \sigma_\omega^2(1 - r^2)$.

[8] In order to find the maximum likelihood estimates we used an iterative, two-stage procedure which is the sequential solution to a probit and regression problem. This procedure, which is of some inherent interest because of its broader applicability, is described in the appendix so as not to divert the reader's attention from the substantive issues at hand.

TABLE 1.—DEFINITION OF VARIABLES

LNP	Natural logarithm of tax price—defined as (reported house value)/(1977/78 State Equalized Value per pupil)
LNEXP	Natural logarithm of general fund school expenditure per pupil in respondent's school district (1977/78)
LNY	Logarithm of respondent's 1977 household income
BLACK	dummy = 1 if black; 0 otherwise
*K*05	number of children age 0 to 5 in household
*K*611	number of children age 6 to 11
KNPUB	dummy = 1 if respondent has 1 or more children in private school
COLGRD	dummy = 1 if respondent is a college student
NONHS	dummy = 1 if respondent did not graduate from high school
FEMALE	dummy = 1 if respondent is female
RTRDI	dummy = 1 if respondent is retired or permanently disabled
*AGE*65	dummy = 1 if respondent is 65 or over
UNEMP	dummy = 1 if respondent is unemployed
TRNSF	dummy = 1 if respondent receives either AFDC or food stamps
LNENRL	logarithm of enrollment in school district (1977–78)
LNPPB	logarithm of 1977–78 enrollment/number of school buildings
DETRT	dummy = 1 if respondent lives in Detroit
LNCTCH	logarithm of county average teacher's salary (1977–78)
LNCY	logarithm of county median income as reported in 1970 census × county per capita income growth (1969–77)
LNCW	logarithm of weighted average of county wages by occupation (1970) ×county per capita income growth (1969–77)
PCEXP	percentage change in school expenditures (1976–77 to 1977–78)
SMSA	dummy = 1 if respondent lives in SMSA
CCITY	dummy = 1 if respondent lives in central city

their preferred public expenditures are the same variables that explain their demand. In order to test for these two sources of bias there must be some measurable set of variables, $\hat{x}$'s, that explain the matching of preferences with communities, but do not affect demand.

A proper empirical analysis of the Tiebout bias question would involve a complete theoretical specification of a model of community choice and public choice which would suggest the proper identifying restrictions. This specification, which is beyond the scope of this paper, would have to incorporate variables which arose from the modelling of the politics of local public goods supply as well as the socioeconomics of migration. We offer instead a set of $\hat{x}$ variables which could reasonably be expected to affect the degree of preference-community mismatch, v, but not to affect underlying preferences.

In general we would expect that individuals who are most likely to have low values of v in absolute value are those (a) who have recently moved, (b) who live in a metropolitan area in which there is substantial choice among public sector bundles, and/or (c) who have tastes that are reasonably similar to others with equivalent incomes. If we were explaining the absolute value of v, therefore, a number of explanatory variables would come immediately to mind. These would include (a) a dummy reflecting a recent move (not available in the Courant, Gramlich, Rubinfeld survey), and (b) an indicator of the extent to which community choices are available (e.g., a dummy indicating presence in a suburban district in a metropolitan area). However, our concern is somewhat different. We are looking for instruments that might be correlated with v, not its absolute value, and which do not appear in the demand equation. Whether the same variables are relevant is a question for which we do not have a confident response at this time.

In any case, we have tried three variables as instruments in our tests for Tiebout bias. One such variable, *PCEXP*, measures the percentage change in per pupil expenditures in the school district over the previous year (1976–77 to 1977–78). Because moving is costly, households will often choose to remain in a community even though local public spending (or average community demand) changes at a different rate than household demand. A relatively large value of *PCEXP* would, if unexpected, reflect a greater likelihood that individuals are consuming more than their desired levels of public spending.[9] Sec-

[9] See Roberts (1985) for a complete discussion of the conditions under which *PCEXP* will be correlated with v, but not with demand.

ond, was the variable *SMSA*, an SMSA dummy meant to reflect the availability of community choice in the metropolitan area. Finally, the third variable was *CCITY*, a dummy variable equal to one when the school district was located in the central city. The central city location was assumed to reflect a limited public education selection despite the fact that the individual resided in a metropolitan area in which there was substantial choice.

The results of the Tiebout bias analysis are given in tables 2 and 3. The first results in column (1) of table 2 are the single equation probit coefficients, equivalent to the ones computed in BRS. If there were no Tiebout bias, these would be estimates of the demand parameters divided by the conditional standard deviation of ϵ. The resulting implied demand parameters are given in column (2) of table 3. These results are discussed, in detail, in the BRS paper. The striking result is that both the price and income elasticities of demand are smaller than those found in previous median voter studies. However, if there is residential sorting by preferences for education, these original estimates could be inconsistent.

Column (2) of table 2 gives the equivalent results if the possibility of Tiebout sorting is allowed. (Column (3) represents the sum of the demand and sorting coefficients.) The straightforward test for Tiebout bias is the test of the hypothesis that the parameter λ is significantly different from zero. The results from table 2 are convincing—an asymptotic t-ratio of 2.36 (on the coefficient λ/σ_e) suggests that the hypothesis that $\lambda = 0$ can be rejected at the 5% significance level. In fact, $\hat{\lambda}$ turns out to be very close to 1 for the results described in table 2. From a practical perspective, this suggests strongly that individuals may sort themselves based upon public goods preferences. We have chosen to present this particular specification, less because of the attraction of the underlying theoretical model, but more because it illustrates the real possibility of Tiebout bias.

As suggested before, failure to correct for Tiebout bias will cause the effects of changes in actual level of expenditures, A, on the response probabilities to be understated. This is exactly what we observe. The coefficient on A (*LNEXP*), the estimated value of $1/\sigma_e$, is 0.43 without the sorting correction. When the effect of Tiebout

TABLE 2.—MAXIMUM LIKELIHOOD COEFFICIENTS

		(1) Single Equation Probit	(2) Full Information Maximum Likelihood	(3)
θ_0^L			4.00	4.38
			(0.94)	(7.13)
θ_0^M			5.74	
			(1.35)	
λ/σ_e			1.74	
			(2.36)	
$1/\sigma_e$	*LNEXP*	0.43	1.74	
		(1.54)	(2.60)	
		β/σ_e	β/σ_e	$(\beta + \gamma)$
x:				
	LNP	−0.14	−0.19	−0.04
		(2.80)	(−3.24)	(−8.32)
	LNY	0.14	0.17	0.02
		(2.33)	(2.71)	(2.90)
	BLACK	1.11	1.13	0.02
		(6.17)	(6.20)	(1.24)
	K05	0.20	0.18	−0.02
		(3.33)	(3.10)	(−2.64)
	K611	0.15	0.15	0.00
		(3.00)	(2.79)	(0.78)
	KNPUB	−0.28	−0.27	0.00
		(2.00)	(−1.91)	(0.31)
	COLGRD	0.16	0.20	0.03
		(1.60)	(1.91)	(3.08)
	NONHS	−0.18	−0.19	−0.01
		(1.80)	(−1.96)	(−1.00)
	FEMALE	0.13	0.15	0.01
		(1.86)	(2.04)	(1.35)
	RTRDI	−0.47	−0.47	−0.01
		(3.62)	(−3.56)	(−0.40)
	AGE65	0.16	0.19	0.02
		(1.14)	(1.32)	(1.49)
	UNEMP	−0.33	−0.28	0.03
		(1.43)	(−1.25)	(1.39)
	TRANSF	−0.19	−0.22	−0.03
		(0.79)	(−0.91)	(−1.05)
	LNENRL	−0.09	0.00	0.04
		(1.80)	(0.03)	(7.01)
	LNPUB	0.17	0.01	−0.11
		(1.31)	(0.04)	(−7.97)
	DETRT	0.27	−0.22	−0.21
		(1.23)	(−0.92)	(−9.12)
	LNCTCH	0.75	1.33	0.35
		(1.23)	(1.80)	(4.37)
	LNCY	−0.76	0.16	0.74
		(1.04)	(0.17)	(9.35)
	LNCW	1.25	0.15	−0.79
		(1.40)	(0.13)	(−8.20)
				τ
$\hat{x}$:				
	PCEXP			1.11
				(12.63)
	CCITY			0.08
				(5.71)
	SMSA			0.07
				(4.22)

TABLE 3.—DEMAND (β'S) AND COMMUNITY SORTING (γ'S) PARAMETERS

	(1) Single Equation Regression Demand Parameters	(2) Single Equation Probit Demand Parameters	(3) Full Information Demand Parameters	(4) Maximum Likelihood Community Matching Parameters
Threshold Constant	2.84	7.49	1.65	1.58
	(6.62)		(2.80)	
x:				
LNP	−0.04	−0.32	−0.11	0.07
	(−7.83)			
LNY	0.02	0.32	0.10	−0.08
	(3.32)			
BLACK	0.01	2.57	0.65	−0.63
	(0.57)			
*K*05	−0.02	0.46	0.10	−0.12
	(−2.76)			
*K*611	0.01	0.34	0.09	−0.09
	(1.22)			
KNPUB	−0.00	−0.65	−0.15	0.15
	(−0.15)			
COLGRD	0.04	0.37	0.12	−0.09
	(3.49)			
NONHS	−0.01	−0.42	−0.11	0.10
	(−0.83)			
FEMALE	0.01	0.29	0.09	−0.08
	(1.62)			
RTRDI	0.00	−1.08	−0.27	0.27
	(0.07)			
*AGE*65	0.02	0.36	0.11	−0.09
	(1.39)			
UNEMP	0.06	−0.76	−0.16	0.19
	(2.23)			
TRANSF	−0.03	−0.44	−0.13	0.10
	(−1.07)			
LNENRL	0.06	−0.22	0.00	−0.11
	(10.99)			
LNPUB	−0.12	0.40	0.00	−0.04
	(−8.56)			
DETRT	−0.14	0.62	−0.13	−0.08
	(−6.15)			
LNCTCH	0.53	1.73	0.77	−0.42
	(8.22)			
LNCY	0.90	−1.75	0.09	0.65
	(11.88)			
LNCW	−0.97	2.90	0.08	−0.87
	(−10.45)			
$\hat{x}$:				
PCEXP				1.11
CCITY				0.08
SMSA				0.06

sorting is accounted for, the coefficient increases by a factor of 4 to 1.74. All the remaining coefficients relating to personal characteristics are relatively unchanged under the Tiebout bias correction. Only the coefficient on the variables that relate to the school district—*LNENRL*, *LNPUB*, *DETRT*, *LNCTCH*, *LNCY*, *LNCW*—change substantially after the correction, and most of these coefficients have low t-ratios.

Each of the variables that is used to explain community matching, but not demand, appears to be significant. The coefficient on *PCEXP* is large

and over 12 times its standard error. Since these variables explain the size of the matching error, v_i, it appears that the larger is the percentage change in expenditures the greater the probability that actual expenditures exceed desired expenditures.

Table 3 reports the demand parameters estimated in three different ways. The first is a regression of actual expenditures on various independent variables. The second is a single equation probit, equivalent to the one reported in BRS. The third uses a full information maximum-likelihood estimator which accounts for the Tiebout bias. The t-ratios are reported for the regression estimates. In the case of the bias-corrected procedure, the community matching parameters, γ and τ, are reported as well (column 4). These are calculated as the difference between the values of $\hat{\beta} + \hat{\gamma}$ calculated from column 3 of table 2 and the estimated demand parameters, $\hat{\beta}$'s, reported in column 3 of table 3.

A number of conclusions seem quite striking. First, a comparison of columns (2) and (3) shows that the correction for Tiebout bias leads to lower price and income elasticities of demand. This is especially interesting since the micro price and income elasticities of BRS were substantially lower than the macro elasticities obtained by most other demand studies. It confirms our view that the income and price elasticities of demand for education are quite low. Second, the correction for Tiebout bias substantially lowers the values of a number of other demand coefficients such as *BLACK* which we found to be unusually high in the BRS paper. Finally, the estimated threshold parameter (table 3) is substantially lower than in the BRS paper. Its value of 1.65 suggests that people do not discern differences in per pupil expenditures that are within 65% of their ideal level. This value is quite high, but considerably more reasonable than the 649% value suggested by the uncorrected single equation probit.

Column (4) includes the consistent estimates of the γ's and the τ's, which allow us to get a sense of the magnitude of the omitted variable bias. The relevant comparison involves the magnitude of each of the γ's to the corresponding β's that were obtained from the probit equation. In this comparison the $\hat{\gamma}$'s tell us (roughly) the extent to which demand parameters would be biased were the demand functions to be estimated without taking simultaneity into account. For example, the price coefficient of 0.07 should be compared to the price elasticity of demand of -0.11. This suggests that failure to account for simultaneity in the BRS estimation caused the price elasticity of demand to be overestimated by more than 50%. Repeated another way, the Tiebout bias correction leads us to the conclusion that the price elasticity of demand is even lower than the BRS results suggest. A similar conclusion is reached about the size of the income elasticity of demand, since the relevant $\hat{\gamma}$ is equal to -0.08, while the $\hat{\beta}$ from table 3 is equal to 0.10.

Another demand variable of particular interest is the *BLACK* variable. The BRS results suggest a very large differential between black and nonblack demands for public school. Our analysis suggests that the correct black coefficient is roughly one-quarter the size of the BRS coefficient which also incorporated the effect of race-related differences in mobility. A similar comparison would apply to many of the coefficients of the demand variables. To the extent that one believes that the demand variables also determine the migration-community choice decision, then the BRS estimates are likely to overstate the magnitude of the demand parameters.

The regression derived demand parameters in column (1) of table 3 are consistent with our understanding of the effect of omitted variable bias. As we suggested earlier, if the degree of community matching is correlated with the demand variables x, then a regression of A_i on x_i will produce biased parameter estimates. It is easy to show that this bias is equal to the value of the community matching parameters, γ, reported in column (4) of table 3. This suggests that the differences between the x parameters in column (1) and the equivalent parameters in column (4) should produce unbiased estimates of the true demand parameters.

V. Conclusion and Additional Comment

This paper represents an assessment of our own earlier work in the light of potential selectivity biases. Bergstrom, Rubinfeld and Shapiro used survey data to estimate the demand function for public education. The price and income elasticities of demand for public education were found to be considerably smaller than those found using aggregate data to estimate median voter models.

Further thought about the econometric issues regarding the estimation problem convinced us that the estimates reported might be biased because of selectivity effects similar to those studied by Heckman and others.

The possibility of selectivity induced (Tiebout) biases arises if people select communities on the basis of their individual preferences for public goods. The practical problem in the BRS framework is that strong assumptions are made about the distributions of random variables. To the extent that this error specification is incorrect, the estimated demand parameters will be biased. A complete formulation yields a simultaneous equation model and a more complicated likelihood function than originally specified by BRS.

The mere possibility of bias is not sufficient in itself to justify a reassessment of previous results. However, estimating the model with the complete specification suggests that the Tiebout bias can be important. How important depends heavily on the choice of instruments used to correct for selectivity bias. Thus, a final resolution of this issue awaits more elaborate and thorough models, as well as empirical testing. Our results might be different, for example, if one were to argue that a number of discrete variables appearing in the demand equation ought to be removed and placed in the community choice equation instead.

In any case, we remain quite confident about the relative magnitudes of our original estimates of price and income elasticities of demand for education. We believe that the price and income elasticities are quite small, substantially smaller than has been suggested by most studies of aggregated data. We look forward to further discussion and analysis of this result since it has important policy implications.

We feel that an additional comment about the low income elasticity of demand is called for because our estimated value is much smaller than values found with community median income estimators. Although we do not have a fully developed model to explain the low value, the finding led us to consider possible explanations.

Suppose education were considered an investment in human capital, the output of which is measured as changes in the wealth, or permanent income, of the student. The marginal product of that investment is the change in permanent income due to a small change in education expenditures. If the marginal product is independent of family income, the optimal investment should not vary substantially between income groups.

One might suspect that both the taste for and productivity of education varies directly with the education of the parents. Our results indicate that the demand for education is between 12% (with the selectivity corrected estimate) and 34% (with the uncorrected estimate) higher for college graduates than for nongraduates.

One finds that community expenditures on education increase substantially with community income. Our results suggest that this may be due to factors which are highly collinear with income, such as education, rather than to income itself.

APPENDIX

Due to the difficulty of programming a full maximum-likelihood procedure, we estimated the parameters by means of an iterative two-stage procedure. The procedure follows from the first-order necessary conditions for maximizing the likelihood function (13). It is useful for this case to rewrite the log likelihood function

$$\begin{aligned}\mathscr{L} = &\sum_{i\in\text{Less}} \log F\left[\theta_0^L + \frac{1-\lambda}{\sigma_e}A_i - \frac{\beta-\lambda\theta_1}{\sigma_e}x_i - \frac{\lambda\tau}{\sigma_e}\hat{x}_i\right] \\ &+ \sum_{i\in\text{Same}} \log\left\{F\left[\theta_0^M + \frac{1-\lambda}{\sigma_e}A_i - \frac{\beta-\lambda\theta_1}{\sigma_e}x_i - \frac{\lambda\tau}{\sigma_e}\hat{x}_i\right] - F\left[\theta_0^L + \frac{1-\lambda}{\sigma_e}A_i - \frac{\beta-\lambda\theta_1}{\sigma_e}x_i - \frac{\lambda\tau}{\sigma_e}\hat{x}_i\right]\right\} \\ &+ \sum_{i\in\text{More}} \log\left\{1 - F\left[\theta_0^M + \frac{1-\lambda}{\sigma_e}A_i - \frac{\beta-\lambda\theta_1}{\sigma_e}x_i - \frac{\lambda\tau}{\sigma_e}\hat{x}_i\right]\right\} \\ &- \frac{1}{2}\sum \log 2\pi\sigma_\omega^2 - \frac{1}{2\sigma_\omega^2}\sum (A - \theta_1 x - \tau\hat{x})^2.\end{aligned}$$

The parameter $\theta_1 = (\beta + \gamma)$.

The necessary conditions for maximizing this likelihood function are as follows:

$$\frac{\partial\mathscr{L}}{\partial\hat{\theta}_0^L} = \sum_{i\in\text{Less}} \frac{f(L)}{F(L)} - \sum_{i\in\text{Same}} \frac{f(L)}{F(L)} = 0, \tag{A1}$$

where $f(L)$ and $F(L)$ are the values of the density and cumulative density functions with values θ_0^L;

$$\frac{\partial\mathscr{L}}{\partial\hat{\theta}_0^M} = \sum_{i\in\text{Same}} \frac{f(M)}{F(M)} - \sum_{i\in\text{More}} \frac{f(M)}{F(M)} = 0. \tag{A2}$$

The presentation of the remaining first order conditions is facilitated by defining a vector Λ, the elements of which, Λ_i, take on the values

$$\Lambda_i = \begin{cases} \dfrac{f(L)}{F(L)} & i \in \text{Less} \\ \dfrac{f(M) - f(L)}{F(M) - F(L)} & i \in \text{Same} \\ \dfrac{f(M)}{1 - F(M)} & i \in \text{More.} \end{cases}$$

Then

$$\frac{\partial \mathscr{L}}{\partial (1/\sigma_e)} = \Lambda' A = 0 \tag{A3}$$

$$\frac{\partial \mathscr{L}}{\partial (\hat{\beta}/\hat{\sigma}_e)} = \Lambda' x = 0 \tag{A4}$$

$$\frac{\partial \mathscr{L}}{\partial (\hat{\lambda}/\hat{\sigma}_e)} = -\Lambda'\left(A - \hat{\theta}_1 x - \hat{\tau}\hat{x}\right) = 0 \tag{A5}$$

$$\frac{\partial \mathscr{L}}{\partial \hat{\theta}_1} = \Lambda' \frac{\hat{\lambda}}{\hat{\sigma}_e} x + \frac{1}{\hat{\sigma}_\omega^2} \sum \left(A - \hat{\theta}_1 x - \hat{\tau}\hat{x}\right) x = 0 \tag{A6}$$

$$\frac{\partial \mathscr{L}}{\partial \hat{\tau}} = \Lambda' \frac{\hat{\lambda}}{\hat{\sigma}_e} \hat{x} + \frac{1}{\hat{\sigma}_\omega^2} \sum \left(A - \hat{\theta}_1 x - \hat{\tau}\hat{x}\right) x = 0 \tag{A7}$$

$$\frac{\partial \mathscr{L}}{\partial \hat{\sigma}_\omega^2} = -\frac{1}{\hat{\sigma}_\omega^2} + \frac{1}{\hat{\sigma}_\omega^4} \sum \left(A - \hat{\theta}_1 x - \hat{\tau}\hat{x}\right)^2 = 0. \tag{A8}$$

The final equation (A8) gives the estimate of σ_ω^2 as the usual sum of squares of the residuals from a regression of A on x and $\hat{x}$.

The two-stage iterative procedure starts with the observation that equations (A6) and (A7) would be satisfied in a least squares regression of

$$A + \frac{\hat{\lambda}}{\hat{\sigma}_e} \Lambda \hat{\sigma}_\omega$$

on x and $\hat{x}$. A beginning value (in our case $\Lambda = 0$) is chosen and a regression of A on x and $\hat{x}$ is run to compute initial values of θ_1, $\hat{\tau}$ and $\hat{\sigma}_\omega$. The estimates of θ_1 and τ are used to find estimates of θ_0^L, θ_0^M, $1/\sigma_e$, β/σ_e and λ/σ_e that satisfy A1–A5 conditional on the estimated values of θ_1 and τ. With these estimates, a value of

$$\frac{\hat{\lambda}}{\sigma_e} \Lambda \hat{\sigma}_\omega$$

is computed and a regression is run to obtain new estimates of σ_1, τ and σ_ω.

The procedure converges to a set of parameter values that satisfy all the first-order conditions.[1]

REFERENCES

Bergstrom, Ted C., Daniel L. Rubinfeld, and Perry Shapiro, "Micro-Based Estimates of Demand Functions for Local Schools Expenditures," *Econometrica* 50 (Sept., 1982), 1183–1205.

Courant, Paul N., Edward M. Gramlich, and Daniel L. Rubinfeld, "Why People Support Tax Limitation Amendments: The Michigan Case," *National Tax Journal* 33 (Mar., 1980), 1–20.

Goldstein, Gerald S., and Mark V. Pauly, "Tiebout Bias on the Demand for Local Public Goods," *Journal of Public Economics* 16 (1981), 131–144.

Gramlich, Edward M., and Daniel L. Rubinfeld, "Micro Estimates of Public Spending Demand Functions and Tests of the Tiebout and Median-Voter Hypotheses," *Journal of Political Economy* 90 (1982), 536–560.

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica* 47 (1979), 153–161.

Ladd, Helen, and Claire Christopherson, "Demand for Local Public Services: Evidence from Survey Data," Harvard University (1983).

Olmsted, George, "Micro-Based Estimates of Demand for Local School Expenditures: Comment," Northwestern University, undated.

Roberts, Judith, "Housing Tenure and Demand for Local Public Goods," Ph.D. Dissertation, University of Michigan (1985).

Ruud, Paul, "Sufficient Conditions for the Consistency of Maximum Likelihood Estimation Despite Misspecification of Distribution in Multinomial Discrete Choice Models," *Econometrica* 53 (Jan. 1983), 225–228.

[1] The procedure just described does not yield independent estimators of σ_ϵ and r. We have derived maximum-likelihood estimators which do so, but have not presented them here because of the additional detail involved.