Tiebout Sorting and Discrete Choices: A New Explanation for Socioeconomic Differences in the Consumption of School Quality*

Patrick Bayer
Department of Economics
Yale University

April 2000

All Comments Welcome

---

* Earlier versions of this paper were circulated during the 1999 academic job market and included in my Stanford University dissertation. I wish to thank my advisors, Doug Bernheim, Tim Bresnahan, and Tom Nechyba, for their continual guidance, encouragement, and support. I have also benefited from many insightful discussions with Tracy Falba and Robert McMillan and wish to thank Pat Bajari, Steve Berry, Peter Davis, Tom MaCurdy, Mike Mazzeo, Patrick McEwan, Roger Noll, John Pencavel, Jon Rork, Anne Royalty, Chris Timmins, Frank Wolak and seminar participants at Columbia, Cornell, Duke, New Hampshire, North Carolina, Penn, Stanford, Texas, and Yale for their helpful comments, ideas, and suggestions. Financial support from the Alfred P. Sloan Foundation and the National Bureau of Economic Research in the form of doctoral dissertation fellowships is gratefully acknowledged. Correspondence may be sent to Patrick Bayer, 37 Hillhouse Avenue, Department of Economics, Yale University, New Haven, CT 06511, phone 203-432-6292, email patrick.bayer@yale.edu.

Abstract

This paper estimates a discrete choice model of the residential location and schooling (private vs. public) decisions of households with elementary school-aged children in California. The parameter estimates describe how households with different characteristics (including race, education, income, and employment location) value a broad set of community, school, and housing characteristics. Using these estimates, the central aim of the paper is to clarify the underlying basis for the substantial differences in the consumption of school quality associated with race, parental education, and income. In particular, the parameter estimates distinguish whether these observed differences in consumption are driven primarily by differences in preferences for schooling itself or by variation in other factors that affect the residential location, school, and housing choices of households with these different characteristics.

Because data matching households with their particular school and community choices are not available for a large study area, the estimation procedure developed in this paper is designed to make the best use of the available 1990 Census data. In particular, the parameters are chosen so that the model best predicts the individual decisions that are observed in the household-level microdata as well as the socioeconomic composition of the families attending public and private school in each community. These composition data by sector are drawn from a special tabulation of the 1990 Census by school district conducted by the National Center for Education Statistics (NCES). Taken together, these two organizations of the 1990 Census provide more than enough information to precisely identify the parameters of the underlying discrete choice model.

The results of the analysis indicate that the differences in the consumption of school quality associated with parental education are driven primarily by differences in preferences for schooling itself. The differences in consumption associated with race, however, are explained instead by differences in the geography of employment opportunities, the demand for housing (whether driven by differences in housing preferences or discrimination in the housing market), and preferences for the race of one's neighbors. This measured disconnect between the preferences and consumption patterns of minority households implies a shortage of communities that combine relatively high quality schooling with either poor quality housing or geographic proximity to employment centers for minority households. That this lack of availability persists despite the fact that minority households would be willing to forgo other forms of consumption to live in such communities implies a potentially serious imperfection in local education markets.

1 Introduction

A direct examination of residential location patterns reveals that low-income, less educated, and minority (black and Hispanic) households tend to live in relatively poor quality public school districts. Moreover, each of these characteristics (race, education, and income) continues to be correlated with school quality consumption even after controlling for the other two factors.[1] What explains these differences in the consumption of public school quality? A common perception is that the differences associated with income simply signify that school quality is a normal good, while the differences associated with race and parental education (holding the other factors constant) represent differences in preferences for school quality.

But this view ignores the complexity of the decision-making process through which a household determines its consumption of school quality, essentially treating school quality like an ordinary consumption good. A household decides how much public school quality to consume, however, through its residential location decision, which simultaneously sets its consumption of a full *bundle* of other local goods (including housing and public safety) as well as its commuting distance and other geographic arrangements. Thus, a household's school quality consumption decision is shaped not only by its preferences for school quality itself, but also by its preferences for other elements of this bundle of local goods, its geographic preferences, and, importantly, the nature and location of the discrete bundles of local goods available in the market. Consequently, the observed differences in the consumption of school quality associated with parental education and racial characteristics need not arise because of underlying differences in preferences for school quality, but might instead be explained by variation in other factors that affect the household's location decision.[2]

The primary contribution of this paper is to clarify the processes that drive the observed differences in school quality consumption across households. To this end, I develop a random-coefficients discrete-choice model of the household's residential location and public-private schooling decisions and estimate the model using data for households with school-aged children in a large study area within California. This approach explicitly accounts for the complex nature of the decision-making process that determines a household's consumption of school quality and controls for many of the other factors that influence this decision. The resulting

[1] I present empirical evidence in support of these statements for California households in Section 6. The measure of school quality referred to is based on an adjusted achievement score measure that I develop carefully later in the paper.

[2] The extent to which a household's consumption of an element of this bundle of local goods deviates from its preferences is directly related to the thinness of the market for its preferred bundle. If, for example, a household prefers to purchase a small house in a community with high quality public schools, its consumption of school quality may be lower than expected if small houses are not typically available in high quality school districts. This particular scenario would be implied by much of the literature that advances the benefit view of zoning. See, for example, Hamilton (1975a) and (1976).

parameter estimates describe how households with different characteristics, both observed (e.g., parental education and race, income, work status, and place of work) and unobserved (i.e., random coefficients), trade-off between the broad set of community, schooling, tenure, and housing characteristics determined by the residential location and schooling decisions. I use these estimates to decompose the observed differences in the consumption of school quality associated with race, parental education, and race into components related to preferences for each element of the bundle of local goods.

That no researcher has attempted to address this research question is most likely due to a fundamental data limitation that has made the estimation of household preferences for community, schooling, and housing characteristics difficult. Simply put, data matching individual households to their particular residential locations and schools have historically been unavailable for a large, representative sample of households. In order to determine the forces that drive the observed differences in the consumption of school quality, then, this paper also makes a methodological contribution by developing a new way to estimate preferences using the available data. In particular, I develop an empirical methodology that uses the information in both the household-level and community-level organizations of the 1990 Census to distinguish household preferences for a large set of community, schooling, and housing characteristics. In the next section of this paper, I describe the nature of this fundamental data limitation and provide intuition as to how the empirical methodology that I introduce permits the estimation of preferences despite this limitation.

The results of the analysis indicate that a substantial portion of the differences in consumption associated with differences in parental education levels is indeed driven by differences in preferences for school quality itself. That is, more highly educated parents are more willing to give up other forms of consumption in order to ensure that their children attend high quality schools, *ceteris paribus*. At the same time, the results indicate that households from different racial backgrounds have remarkably similar preferences for school quality. Consequently, the observed differences in the consumption of school quality associated with race cannot be understood as resulting primarily from differences in preferences for schooling itself. Instead, the analysis suggests that minority households choose to live in communities with relatively poor quality schools for a variety of other reasons associated with the residential location decision. These include differences in the geography of employment opportunities, differences in the demand for housing, (whether driven by differences in housing preferences or discrimination in the housing market), and differences in preferences for the race of one's neighbors.

The fact that minority households live in poorer quality school districts despite having preferences for schooling similar to white and Asian households suggests that minority households are particularly constrained by the bundling of schooling with the other local goods and geographic relationships jointly determined by the residential location decision. In fact, this *disconnect* between consumption patterns and preferences for minority households implies that certain bundles of local goods are not readily available in the market. In particular, my results imply that communities with relatively poor quality housing and relatively good public schools as well as communities that provide high quality public schooling in close geographic proximity to employment centers for low-income and minority households are not commonly found in the market. Consequently, in the case of the differences in the consumption of school quality associated with race, the answer to the key economic question of this paper is not that minority households place a substantially different value on their children's education than white households. Instead, the constrained nature of the choice set associated with the residential location decision in both physical geography and product space provides a substantial part of the underlying explanation for the observed differences in the consumption of school quality. These results have important implications for policy as they suggest that the economy is systematically under-providing school quality to certain segments of the population. I discuss these policy implications in a final section of this paper.

2 Estimating Preferences for Community, Schooling, and Housing

The main empirical exercise of this paper is the estimation of preferences for a broad set of community, schooling, and housing characteristics. The estimation of these preferences is complicated, however, by some unavoidable limitations of the available data. In particular, data matching individual households with their particular choices of school and residence are simply unavailable for a representative sample of households in a large study area. Moreover, because many important choice characteristics including school quality, crime rates, and transparently neighborhood demographic and socioeconomic characteristics depend, in part, on the characteristics of the households that make each choice, appropriate estimation of these preference parameters must account for the fact that these characteristics will certainly be correlated with any important unobserved community, schooling, and housing characteristics. In light of this important endogeneity issue and data limitation, this paper offers a new approach to estimating preferences for community, schooling, and housing characteristics. To see why this new approach is necessary, however, it is instructive to consider how researchers have handled these complications in previous studies.

The Previous Literature

In the past, researchers who have attempted to understand the determinants of the residential location decision or public-private school choice have always analyzed these two decision processes separately. In both strands of the literature, however, researchers have faced the same data limitation – that data matching individual households with their particular choices of school and residence are simply unavailable for a representative sample of households in a large study area. Researchers studying the residential location decision, therefore, have typically relied on data sets that have been limited in their representativeness, geographic scope, or both. The majority of these studies have focused on estimating how households trade off between housing characteristics and commuting modes and times.[3] The few studies that have considered local public goods or other broader community characteristics explicitly in the analysis have always relied on expenditure measures to proxy for the public goods provided in each community and have typically used data for only a small number of communities.[4] Moreover, none of these studies that have examined the location decision directly has attempted to control for the endogeneity of local public goods or other sorting-dependent choice characteristics.

Of course, estimating a discrete choice model of residential location is not the only imaginable way to estimate preferences for local public goods and other location-specific characteristics.[5] Recognizing, for example, that household preferences for local public goods impact local political processes, many researchers in the 1970s attempted to estimate models of demand for local public expenditures consistent with theoretical models of local voting processes.[6] Unfortunately, these studies did not control for the nonrandom sorting of households across communities, and a subsequent line of research demonstrated that the estimates of these models were subject to a potentially serious form of selection bias – called Tiebout bias in this setting.[7] In particular, Rubinfeld, Shapiro, and Roberts (1987) found that this bias may be quite substantial, but were unable to accurately measure the magnitude of the bias or the role of family characteristics in demand. By these authors'own admissions, it is very difficult to measure preferences for any local public good

---

[3] Many researchers in the urban economics literature, for example, have used this framework to estimate the importance of housing characteristics and commuting costs in the residential location decision. Ellickson (1981), Quigley (1985), Blackley and Ondrich (1988) focus on estimating the influence of housing characteristics on the location decision, while Anas (1982), Pollakowski (1982), Anas and Chu (1984) focus on estimating the influence of commuting costs.

[4] McFadden (1978) presents an econometric framework for estimating a discrete choice model of the residential location decision given individual-level data for a large study area. Quigley (1985) and Nechyba and Strauss (1998) include measures of public expenditures on education, policing, and all public goods in their analyses.

[5] Other literatures attempt to provide (at least indirect) evidence about preferences for local public goods. Starting with Oates (1969), for example, many researchers have attempted to estimate whether increases in the level of public goods provided in a community are capitalized into property values. Unfortunately, this literature has had difficulty testing for capitalization, and other papers have posed serious questions about this method of estimating preferences for local public good. See, for example, Hamilton (1975b).

[6] See Deacon (1977) and Inman (1979) for reviews of the literature that attempts to estimate models of demand consistent with the median voter model.

[7] A sequence of studies by Goldstein and Pauly (1981), Bergstrom, Rubinfeld, and Shapiro (1982), and Rubinfeld, Shapiro, and Roberts (1987) demonstrate the empirical importance of Tiebout bias regardless of whether aggregate- or micro-level data are used in the analysis.

without directly modeling the full community choice decision itself, which is precisely the exercise that I carry out in this paper.

Handling the Fundamental Data Limitation

The approach to estimating preferences for local public goods (and schooling and community characteristics more generally) put forth in this paper is most closely related to the discrete choice models of Quigley (1985) and Nechyba and Strauss (1998). While these studies used datasets that were limited in representativeness and geographic scope respectively, my approach makes use of the Public Use Microdata Survey (PUMS) of the 1990 Census for a large study area in California. These *household-level* data provide detailed information about the characteristics of each household in a very large study area as well as some information about its choices of residence and school (e.g., its public-private school choice). Unfortunately, however, these data do not identify each household's residential location precisely enough to accurately match the household to its particular community and school (or school district), rendering traditional methods of estimating discrete-choice models (which require data matching each household to its particular choice) unusable.

In order to estimate a discrete-choice model of the household residential location and schooling decisions, then, I supplement these household-level Census data with a partially aggregated version of these same data organized at the school district level.[8] These *school district-level* data provide information about the characteristics of the households that live in each school district, and, importantly, distinguish between households that send their children to private versus public school. In this way, while the household-level data do not contain information that identifies each household's particular residential location and schooling decisions, the district-level data provide a great deal of information about the characteristics of the households that send their children to public and private school in each community. Estimation of the household utility function proceeds, then, by selecting the preference parameters that best predict both the individual choices (e.g., private vs. public schooling) observed in the household-level data and the distributions of households and their characteristics observed in the district-level data. Taken together, these two organizations of the 1990 Census provide more than enough information to precisely identify the parameters of the underlying discrete choice model.

This approach draws on recent work in the empirical Industrial Organization literature, where researchers have faced a similar data limitation in a comparable economic context – that of differentiated products markets.[9] Using only aggregate

---

[8] These *district level* data are made available through a special tabulation of the 1990 Census conducted by the U.S. Department of Education.

[9] In fact, the analogy between the economics of differentiated products markets and local public economics is well established in both the theoretical literature (Epple and Romer

product-level data, Berry (1994) and Berry, Levinsohn, and Pakes (1995) develop an estimation procedure that allows the recovery of the parameters of the underlying household utility function.[10] In essence, the parameters of the utility function are chosen so that when the product choices of the individuals in the sample are aggregated, the resulting market shares predicted by the model match those observed in the data as closely as possible. In more recent work, Berry, Levinsohn, and Pakes (1998), develop an empirical approach that uses both aggregate- and individual-level data in order to estimate a specification of the household utility function that allows rich variation in preferences across households with different observable characteristics. It is this empirical framework of this latter paper that my approach most resembles.

The Endogeneity of Sorting-Dependent Choice Characteristics

Many important choice characteristics including school quality, crime rates, and transparently neighborhood demographic and socioeconomic characteristics depend, in part, on the characteristics of the households that make that particular choice. Consequently, these sorting-dependent characteristics are certainly correlated with any other choice characteristic (whether observed or unobserved) that affects how households sort across communities and between the public and private sector of the education market. This correlation between sorting-dependent characteristics and any unobserved characteristics of communities and schools leads to an important endogeneity problem that must be addressed in the estimation. None of the previous papers in the literature has attempted to deal with this issue in any way.

In order to estimate the full model, it is necessary to find a set of instruments that are correlated with the sorting-dependent characteristics of a particular alternative but uncorrelated with any unobserved characteristics of that choice that are valued by households. Fortunately, the nature of the choice process itself suggests a natural set of instruments that fits this requirement – the *exogenous* characteristics of nearby alternatives (in an exogenous dimension).[11] In this paper, I

(1978), Epple, Filimon, and Romer (1983, 1984, and 1993)) and in the hedonics literature (Epple (1987), Bartik (1987), Rosen and Fullerton (1974)).

[10] Beginning with the studies of Bresnahan (1981) and (1987), researchers in the Industrial Organization literature have estimated equilibrium models of differentiated products markets using only aggregate product-level data. The studies of Berry (1994) and Berry, Levinsohn, and Pakes (1995), however, were the first to estimate a demand system based on explicit aggregation of the underlying household utility functions. Following these influential papers, there has been an explosion of interest in problems that require the estimation of random utility models in differentiated products markets. For a particularly accessible description of these models as well as a very complete bibliography of recent applications see Nevo (1999).

[11] The important attribute that a good instrument must have in this case is that it must influence the set of households that sort into a particular community and/ or community-school combination. To the extent that much of the structure of the choice set is exogenously determined (e.g., through the physical geography of community locations relative to one another), many instruments can be found that should affect sorting

use the geographic dimension to specify nearby alternatives, using the exogenous characteristics of neighboring communities to instrument for sorting-dependent characteristics. These exogenous neighbor characteristics make appropriate instruments because they affect household sorting and consequently are correlated with the sorting-dependent characteristics of the particular choice in question in equilibrium. At the same time, because these characteristics are assumed to be uncorrelated with even their own alternative's unobservable characteristics, it seems natural to assume that they will be uncorrelated with the unobservable of nearby choices. Thus, the construction of an appropriate set of instruments follows directly from the nature of the choice process itself.[12]

The Wider Applicability of This Approach

The primary goals of this paper are to estimate household preferences and to use these estimates to distinguish the factors that drive differences in school quality consumption across households. The empirical model of location and school choice developed in this paper is applicable, however, for addressing a much broader set of questions involving the sorting of households across communities and between the public and private sectors of the education market. In particular, the methodology can be used to control for selection (Tiebout) bias in the estimation of the education production function. For example, in attempting to estimate the role of school inputs or peer effects in the education production function, researchers have often worried about the selection bias that might result if the same households that produce higher achieving children also choose better quality schools or peer groups for their children. Bias results because the observable school or peer characteristics are correlated with the unobservable household characteristics that contribute to the error term.[13] Researchers have had similar difficulties estimating the role of the family in the production of achievement, in this case due to the possible correlation of observable family characteristics with unobservable components of school and peer quality contained in the error term. Controlling for the nonrandom sorting of households between communities and the public-private sector of the education

but do not affect the utility that any household gets from choosing a particular alternative. Other potential instruments include measures of how isolated the particular choice is in geographic and/or product space.

[12] It is worth emphasizing that the construction of a set of instruments based on the exogenous characteristics of nearby choices (in an exogenous dimension) is always possible in discrete choice applications when the economic agent must choose from a wide range of choices. I discuss this IV approach in more detail in Section 5.

[13] See, for example, Boozer and Rouse (1999) for a discussion of Tiebout bias as it relates to estimating the role of school characteristics in the education production function. See Evans, Oates, and Schwab (1992) for a discussion related to estimating the role of peer characteristics.

market permits the unbiased estimation of the role of family, peer, and school characteristics in the education production function.[14]

Controlling for Tiebout sorting is also critical for the proper estimation of the determinants of the public-private school choice and, consequently, the evaluation of the relative performance of public and private schools.[15] Moreover, the specification of the household choice process developed in this paper forms, in essence, the demand side of the education market and consequently forms the backbone for studying the supply side of education and housing markets– including issues of private school entry and competition (both public-public and public-private).[16] Finally, the estimation of the household choice process is also useful for attempting to understand the likely response of households to major policy changes such as school finance reforms like education vouchers.[17] In this way, the potential applications of the methodology developed in this paper go well beyond the specific application developed herein.

3 The Model and Study Area

In this section, I discuss the empirical framework for the analysis undertaken in this paper. I begin by describing the choice of California as a study area, the utility specification used in the discrete choice model of residential location and school choice, and the importance of the stochastic components of this utility specification. In order to provide a coherent framework for the empirical analysis, I also describe additional features of the empirical framework that are important for estimation. These include:

(i) assumptions about the market clearing conditions in the housing market and the determination of housing prices,

(ii) the empirical specification of the education production process and the estimation of preferences for public school quality, and

(iii) the conditions that ensure a sorting equilibrium exists and the conditions under which this equilibrium is unique.

A full discussion of the identification of the model is left until Section 5.

The California Study Area

---

[14] In a companion paper (Bayer, 2000a), I estimate the education production function in a manner consistent with the nonrandom sorting of households across communities.

[15] Bayer (2000b) focuses on these issues.

[16] In Bayer and McMillan (in progress), we estimate a structural model of competition in the housing market.

[17] Bayer, McMillan, and Rueben (in progress) studies the impact of a series of school finance reforms on the levels of residential segregation and stratification by income.

The study area for my analysis consists of the 620 school districts that contain the sixth grade in 30 counties of California.[18] California's unique school financing system, which gives local communities little control over either the level of education expenditures or the property tax rate, makes it an excellent setting for this analysis because it simplifies the estimation of the household preference parameters that determine the residential location decision. Until the early 1970's, California's school districts were funded almost exclusively by local property taxes. Beginning in 1971, however, a series of landmark decisions by the Supreme Court of California and legislation by the California legislature transformed this locally financed system into one in which the majority of funding is provided by the state.[19]

Following these court decisions and the passage of Proposition 13 in a statewide referendum, the California State Legislature passed legislation that set in place the system school finance used in California today. In order to receive much needed revenue for education from the state, each district is required to levy a one-percent property tax rate and to allocate revenue among public expenditures in the same proportions as it did in 1978. When the district meets these two requirements, the state provides the additional revenue necessary to bring each district up to a revenue limit for education expenditures. In this way, increases in property tax revenues due to a growing tax base are directly offset by a decrease in the amount of revenue that the state provides. Consequently, California's school finance system does not allow a school district to control either the level or its share of the contribution to education expenditures.

California's school finance system has two immediate consequences for the estimation of household preferences. First, since local property taxes are fixed at an effective rate of one percent, variation in property tax rates does not affect the residential location decision of households, as it does in most states. Second, as Dynarski, Schwab, and Zampelli (1989) point out, since the preferences of the constituents of each district have no direct effect on either education expenditures or property tax rates, the primary way that household preferences for school quality are revealed is through the sorting of households across districts. Consequently, in developing a model of the local economy, I ignore the part of demand for education that would ordinarily operate through voting over local taxes and expenditures.

A Discrete Choice Model of the Household Location and Schooling Decisions

The discrete-choice model specified here describes the residential location and schooling (public vs. private) decisions of all households with school-aged

[18] Unfortunately, a missing data problem (which I describe in the Data Appendix) prevents the inclusion of the other 26 counties of California. This problem is not as bad as it may appear, however, as the missing counties are almost entirely rural counties with small populations. The study area includes all of the major counties of Southern California and the San Francisco Bay area. This region contains rich variation in climate, air quality, housing, economic opportunities, and local public goods and includes the large cities of Los Angeles, San Diego, San Francisco, and San Jose as well as many suburban and rural communities.

[19] More details on school finance in California can be found in Appendix 2.

children in the California study area. The particular model of utility maximization that I adopt is based on the random utility model developed in McFadden (1975, 1978) and the random coefficients specifications of Berry, Levinsohn, and Pakes (1995, 1998).

Each household chooses its residence and school type to maximize its utility, which depends on the observable and unobservable characteristics of its choice. The household's choice set can be characterized by the triple *(c, s, h)*, where *c* represents the household's community choice, *s* its school choice, and *h* its particular housing choice (including its tenure choice). Let X_c represent the observable characteristics that are determined by the household's community choice alone. These include characteristics of the climate, air quality, crime rates, urban density, and geographic location. Let X_{cs} represent the observable characteristics that vary with the household's choice of public versus private school, given that it has chosen community *c*. Note that the *cs* subscript here denotes the fact that these school characteristics vary with the household's community choice. In a similar manner, let X_{ch} represent the observable characteristics that vary with its community and particular housing choice including the annual housing cost, p_{ch}. Again, the characteristics of the housing units available to the household depend on the particular community choice that the household makes.

Each household's valuation of these various choice characteristics is allowed to vary with its own (observable and unobservable) characteristics, z_i, and endowments. The observable household characteristics include parental education levels, race, and household size and each household is initially endowed with a primary employment location, l_i, and income, y_i. Combining all of these elements, the household's optimization problem is given by:

$$(1)\qquad \underset{(c,s,h)}{Max} \quad V^i_{csth}(X_c, X_{cs}, X_{ch}, p_{ch}, z_i, y_i, l_i, e^i_{csh})$$

where (for now) e^i_{csh} represents all of the unobservable components of the utility function. The particular utility specification taken to the data is given by:

$$(2)\qquad \underset{(c,s,h)}{Max} \quad V^i_{csh} = a^i_c X_c + a^i_s X_{cs} + a_h X_{ch} - a_p p_{ch} - a^i_D D^i_c + x_c + x_{cs} + x_{ch} + e^i_{csh}$$

where each parameter a^i_j, j ,*{c, s, h, p, D}*, is allowed to vary with a household's own (observable and unobservable) characteristics:

$$(3)\qquad a^i_j = a_{0j} + \sum_{r=1}^{R} a_{rj} z^i_r + a_{yj} y^i + u^i_j \ .$$

and x_c, x_{cs}, and x_{ch} are fixed effects associated with each community *c*, (x_c), each school type *s* in community *c* (x_{cs}) and each housing choice *h* in community *c* (x_{ch}), respectively.

A couple of additional notes about this specification should help to clarify the particular form of the utility function developed here. Equation (3) describes household *i*'s preference for choice characteristic *j*. The first term of this equation captures the taste for the choice characteristic that is common to all households, the second and third terms capture observable variation in the valuation of this choice characteristics across households due to socioeconomic characteristics, and the final term n_i represents a stochastic taste component associated with each household *i*. The inclusion of this stochastic component of tastes gives rise to the familiar random-coefficients model. Let s_j^2 represent the variance of the stochastic taste component for each choice characteristic, *j*, which will be estimated along with the rest of the model's parameters.

Stochastic Structure

The specification of utility given in equations (2) and (3) contains three distinct stochastic components, which collectively allow the model great flexibility in explaining the observed data. The first of these stochastic components is the idiosyncratic term e^i_{csh}, which is assumed to be additively separable from the rest of the utility function and to be distributed according to the Type II extreme value distribution, giving rise to the traditional multinomial logit model. As an alternative to the multinomial logit framework, I also estimate a series of nested logit specifications, which produce more desirable substitution patterns under most conditions.[20] For expositional purposes, it is easier to work with the simpler multinomial logit framework and so I carry this assumption throughout the exposition of the model.

The second stochastic component of the utility function includes the choice-specific fixed effects (x_c, x_{cs}, and x_{ch}) associated with each community, school, and residence, respectively. These terms are assumed to enter each household's utility function in a like manner and capture those aspects of a particular choice that are observable to the households in the sample but unobservable to the econometrician. The inclusion of these terms in the utility function gives the model the flexibility to explain differences in demand (and prices) for what might appear to be roughly equivalent choices. Note that by construction x_{cs} and x_{ch} are each mean-zero within each community. So, for example, if a community's public and private schools are

---

[20] The nested logit framework allows the household's substitution pattern between two choices within any particular branch of a specified decision tree to differ from its substitution pattern between two choices in separate branches, thereby avoiding the general property of the independence of irrelevant alternatives (IIA) associated with the multinomial logit. As its description hopefully suggests, the IIA property – that the ratio of the probabilities that a household chooses any two districts depends only on the characteristics of those districts and not the nature of the remainder of the choice set – is well known to produce strange substitution patterns.

both unobservably better than those in other communities, the mean of this unobserved quality will be absorbed in x_c while x_{cs} will capture the unobserved difference between the quality of the private and public schools in the community.

The final stochastic component of utility includes the random coefficients terms (n^i_j) which allow an individual's preference for each choice characteristics to vary from that of other observationally equivalent individuals. Along with the differences in preferences associated with observable household characteristics, these terms allow the model to predict reasonable substitution patterns across choices. For example, households with a particularly strong taste for public school quality may substitute between communities that provide high levels of public school quality *ceteris paribus*, while households with a strong taste for clean air may substitute between communities in which such air quality is more readily available. It is also important to note that this random-coefficients specification explicitly captures the fact that some households may simply have a stronger taste for certain choice characteristics than others. These unobserved differences in tastes are precisely what gives rise to many of the manifestations of Tiebout bias that have concerned researchers in many areas of public and urban economics.[21]

Market Clearing Conditions in the Housing Market

Separating the idiosyncratic stochastic component (e^i_{csh}) from the rest of the log-utility function specified in equation (2), such that $V^i_{csh} = W^i_{csh} + e^i_{csh}$, the probability that household i chooses choice (c, s, h) can be written as:

$$(4) \qquad P^i_{csh} = \frac{\exp(W^i_{csh})}{\sum_{x=1}^{C}\sum_{y=1}^{S}\sum_{z=1}^{H}\exp(W^i_{xyz})}.$$

In this way, the probability that each household makes any particular choice depends not only on the characteristics of the choice itself (which enter this probability statement through the numerator), but also on the characteristics of every other choice (which enter the probability statement through the denominator).[22]

In characterizing the housing market, I restrict my attention to those housing units currently occupied by households with school-aged children. While it is straightforward to extend this characterization to include households without children, exposition and estimation are both made easier by restricting attention to only households with children. Assume also that the supply of housing available to

[21] See, for example, Bayer (2000a and 2000b).

[22] As should be clear from equation (4), it is necessary to normalize the utility of one choice, as utility is only meaningful as a relative concept. In practice, we will normalize the common component of utility of one of the choices to be zero.

these households is fixed.[23] The system of equations that defines the market clearing condition for the housing market can then be written:

$$(5) \qquad \sum_i P^i_{csh} = 1$$

where there are as many equations as houses available in the market. Simply put, the market clearing condition in this random utility framework is that one household resides in each housing unit. Given this market clearing condition, it is possible to show that a unique set of housing prices clears the housing market using an immediate application of the proof in Berry (1994).

The intuition behind this uniqueness result is easy to see if one imagines a house for which the sum of probabilities in (5) is greater than one. In this case, the demand for the house exceeds the supply and so, the price of that particular house would be bid up, thereby reducing the sum of the probabilities in (5). The random utility framework is particularly important here as it gives each household some small positive probability of choosing each house and allows this probability to move continuously with community and housing characteristics including price. This continuity is useful for proving that a unique set of prices clears the housing market and is also important for the establishing the existence and uniqueness of a sorting equilibrium in this model. I discuss these equilibrium properties later in this section. Finally, it is worth noting here that given the implicit determination of the vector of housing prices described here, each price is a function of the full distribution of choice characteristics (both observed and unobserved) as well as the full distribution of consumer characteristics and tastes.

The Education Production Process and Measuring Preferences for Public School Quality

Measuring household preferences for school characteristics is an important part of the analysis undertaken in this paper. In particular, in order to compare consumption patterns across households, it is useful to examine a single achievement-based measure of public school quality. The measure used in this paper is based on the contribution to achievement made by a child's peers in the school and the school itself. In order to make this measure of public school quality clear, it is helpful to specify the education production process. It is important to note, however, that I do not actually estimate the education production function as part of this paper.[24]

---

[23] It is straightforward to extend the empirical exercise in this paper to include a housing supply function that is not completely inelastic. Both the exposition of the model and estimation are made significantly easier with the assumption that the housing supply is fixed.

[24] In Bayer (2000a), I estimate the education production function – specifically controlling for the nonrandom sorting of households across communities and between the public and private sectors.

Assume that public schools transform inputs (including student characteristics) into academic achievement according to an education production function. The education production function describes the determination of a district's average test score, t_c, from the characteristics of the school, S_c – including average class and school size, average teacher experience, education, and salary, and the family backgrounds of its students, Z_c. To simplify the exposition that follows, assume that the production function is linear in family and school characteristics, as well as the stochastic term, w_c:[25]

(6) $$t_c = b_1 Z_c + b_2 S_c + w_c$$

Here w_c represents the stochastic part of the production function and includes both the unobservable characteristics of the families attending public school and the public school district itself in community c.

Given this production function, we can now discuss the construction of an appropriate measure of school quality. From the household's perspective, this measure would include the contribution to its child's achievement made by both the school itself as well as the child's peers in the school. The $b_1 Z_c$ term in equation (6), however, combines the *direct effect* of a family's characteristics on its own children with the effect that a child's family background has on other students – a *peer effect*.[26] Decomposing the family effect into direct and peer effects ($b_1 = b_{dir} + b_{peer}$), then, the appropriate measure of school quality from the family's perspective would simply remove the direct effect from the district's average test score:

(7) $$q_c \quad = \quad t_c - b_{dir} Z_c \quad = \quad b_{peer} Z_c + b_2 S_c + w_c \, .$$

Focusing on the first part of equation (7), notice that both t_c and Z_c are observed in the data. Therefore, by including these variables in the utility function, it would be possible to estimate preferences for school quality and b_{dir} as long as the school quality measure in equation (7) captured all of the aspects of a school about which households care. In particular, if the utility parameters that multiply the test score and peer characteristics are given by a_t and a_Z respectively, then a_t measures preferences for school quality and b_{dir} is given by the expression $-a_Z / a_t$. If, however, parents care about the characteristics of their child's peers above and beyond the effect these characteristics have on achievement, it would be impossible to separately identify a household's tastes for peer characteristics from the role that

---

[25] All that is really required here is that the production function be additively separable in terms of family and school characteristics, as well as the stochastic component, w_j.

[26] The $bZ_{c,pub}$ term in the production function also picks up any other contributions to school quality made by parents, (volunteering time, providing additional resources through private foundations, etc.). While a wide variety of papers consider the impact of peer effects on the behavior of children, their empirical importance is still not well measured. For a summary of this literature and a particularly insightful discussion of the issues involved in the estimation of peer effects see Nechyba, McEwan, and Older-Aguilar, (1999).

peer characteristics play in measure of school quality in equation (7). *Importantly, though, it remains possible to infer a household's preferences for school quality, which continues to be based on the utility parameter that multiplies the test score.* In a like manner, it is possible to include other school characteristics (such as average class size) in the utility function if one thought that the household might value these above and beyond their role in the production of achievement. Again, it is still possible to interpret the utility parameter associated with the average test score as a household's preference for school quality.

Existence and Uniqueness of the Sorting Equilibrium

In this paper, I estimate the parameters of the household utility function in a manner consistent with a sorting equilibrium in which:

(i) each household makes its optimal residential and school choice, given the choices made by all other households,

(ii) a unique vector of housing prices clears the housing market, and

(iii) the processes that determine public school quality, the crime rate, and the characteristics of a child's peers are allowed to depend (in part) on the characteristics of the households who make the relevant community or school choice.

I assume that households play a one-shot game in making their residential and school choices and I use a Nash equilibrium concept. While I do not estimate (or completely specify) the processes that determine the endogenous choice characteristics described in condition (iii), it remains important to describe the properties of the sorting equilibrium in this model.

An easy point of confusion relates to the conditions that are necessary to estimate the parameters of household utility function. To help make these conditions clear, consider the simple figures collectively depicted in Figure 1, which depict a series of correspondences from parameter and unobservable *(q, e)* space to data *(X)* space. In order to estimate the model using the Generalized Method of Moments (GMM), it is necessary that each point in *(X)* space map to a unique point in (*q*, *e*). This corresponds to the usual notion that the model be just- or over-identified. *Importantly, however, it is not necessary for estimation by GMM that each point in (q, e) space map to a unique point in (X) space.* Thus, estimation by GMM is possible under situations like those displayed in either of the top two panels of Figure 1, but not in the case depicted in the bottom panel. It is worth noting, though, that in order to use the resulting parameter estimates to carry out equilibrium policy experiments, it would be necessary that each point in parameter and unobservable space map to a unique point in data space. Such a unique mapping is also necessary for estimation via Maximum Likelihood (ML), as ML requires the researcher to specify the complete process through which the data is generated. Thus, for conducting policy experiments after estimation or estimation via ML, it is necessary that the correspondence between *(q, e)* and *(X)* space is one-to-one.

In Bayer and Timmins (2000), we develop the equilibrium properties of the model used in this paper. We first prove that a sorting equilibrium always exists with the random utility specification of household behavior used in this paper. The only additional requirement that we must impose is that the choice characteristics that depend in equilibrium on the characteristics of the households that make that particular choice be continuous in these household characteristics. In this way, it is always possible to estimate the parameters of the household utility function using GMM as long as we can specify an appropriate set of moment conditions that just- or over- identifies the model. In Bayer and Timmins (2000), we also specify conditions under which the correspondence between (q, e) and (X) space is one-to-one. Because I do not use the parameter estimates to perform equilibrium policy experiments in this paper, however, I do not focus on these conditions here.[27] I discuss the identification of the model in Section 5.

4 Data

The California Dataset

In order to understand the logic of the estimation procedure, it is critical to understand the structure of the dataset that I have assembled for this analysis. Consequently, the primary goal of this section is to provide a clear exposition of the underlying structure of the data. More detailed descriptions of the data sources and the construction of the variables used in the analysis are provided in the Data Appendix.

The primary source for the household data is the 1990 Census, from which I draw two distinct datasets. The unit of observation in the first dataset is the household itself. The information provided in this dataset includes the income, education, race, and employment location(s) of each household with children aged 10-13.[28] These data also provide information about each household's school (public or private) and tenure choice (rent or own), as well as the characteristics of its housing unit, (e.g., value or rent, size, and age). These household-level data are summarized in the first column of Table 1.

---

[27] In general, these conditions require restrictions on preferences that limit the contribution to utility made by preferences for endogenous choice characteristics (such as public school quality or crime rates) relative to the contribution to utility made by exogenous choice characteristics. These conditions also require restrictions on the differences in preferences across households with different characteristics. In general, these restrictions on differences in preferences are more complicated and depend in part on role of household characteristics in the production technology for public school quality, crime rates, and other endogenous choice characteristics.

[28] I choose this subset of the data (households with children aged 10-13) in order to be as consistent as possible with the Census data summarized by district that includes all children in grades 5-8.

Unfortunately, this household-level dataset does not match each household with the school or school district in which it lives. Instead these data match households only to a Public Use Microdata Areas (PUMA), which contain populations of at least 100,000 residents and are generally much larger than school districts.[29] If I were able to match each household to the school district in which it lives, I could estimate the household utility function in a fairly straightforward manner using a traditional approach to estimating discrete choice models. Such an approach would choose the parameters of the utility function in order to maximize the predicted probability that each household makes the correct choice, *(c, s, h)*. Because these household-level data do not match each household with its chosen district, however, the particular choice made by each household is not known precisely. Consequently, in order to estimate the model in the face of these limitations, I make use of another organization of these same Census data, summarized at the school district-level.

For each of the 620 school districts in the sample, I draw aggregate information about the households in the district from a special tabulation of the 1990 Census.[30] These district-level data identify the fraction of households in the district in each of four parental education categories, each of many income categories, and each of four racial categories and are summarized in the last four columns of Table 1.[31] In this way, while I do not observe the school district that each household chooses, this special tabulation of the Census provides a good amount of information about the distribution of household characteristics in each school district. Importantly, this special tabulation allows the full sample of households to be restricted in two important ways. First, it is possible to isolate the characteristics of only those households with school-aged children and more specifically school-aged children within specific grade ranges. Consequently, in constructing the education, income, and race variables defined above, I use the fraction of households with children in grades 5-8.[32] Importantly, it is also possible to distinguish the characteristics of the households that send their children to public school from those that send their children to private school.[33]

In addition to the Census data that provide information about the households in my sample, I have assembled a variety of additional characteristics that describe

---

[29] The employment location information provided by the census is also at the PUMA level.

[30] This special tabulation was conducted by the National Center for Education Statistics and published in the School District Data Book. See the Data Appendix for further details.

[31] As is apparent by comparing the first and last columns of Table 1, the characteristics of the household-level data match the characteristics of the district-level data quite well.

[32] While I would have liked to use only households with children in the sixth grade, the nature of this special tabulation of the Census forces me to include all households with children in grades 5-8. To be as consistent as possible, therefore, I attempt to distinguish the same households in constructing the rest of the data set.

[33] Due to data limitations, many previous studies have been forced to simply use the average characteristics of the full district's population as proxies for the characteristics of the students in the public school.

each community's geographic location, crime rates, climate, population density, and air quality. Summary statistics for these variables are provided in Table 2 and the construction of these variables is documented in the Data Appendix. In addition, I have collected a set of variables that describe the characteristics of a district's schools. These variables include the average sixth grade test score; the average class and school sizes; average teacher education, experience, tenure, and salaries; the fraction of employees in teaching, administration, and pupil services; and the percentage of expenditures dedicated to various components of the school budget. These variables are summarized in Table 3. Characteristics of the private schools in each public school district are displayed in Table 4.

The Structure of the Data and the Logic of the Estimation Strategy

The estimation strategy uses the information about household decisions contained in both organizations of the 1990 Census to estimate the parameters of the household utility function.[34] These Census data provide two distinct sources of information about how the households choose residences and schools. First, the household-level Census data provide direct information about each household's school, tenure, and housing choices. Moreover, these data provide some information about the household's location decision, although this information is not precise enough to determine the household's exact community choice. The second source of information about the choices made by the households in my sample is the community-level data derived from the special tabulation of the 1990 Census. These data provide a great deal of information about the demographic and socioeconomic composition as well as the school choices made by the households with school-aged children in each community.

The estimation procedure combines these sources of information in order to form two types of moment conditions and estimates the model by choosing the set of parameters that minimizes an optimal function of these moment conditions. First, the household level data can be used directly, as in a traditional discrete choice estimation procedure, to choose the parameters of the utility function in order to maximize the predicted probability that each household makes the appropriate school type, tenure, and housing (and, to some extent, location) choices. Because I do not observe the household's community choice directly, however, it is necessary to specify additional moment restrictions based on attempting to correctly predict the choice patterns observed in the school district-level data. The logic of this part of the estimation strategy is to choose the parameters of the utility function so that when the households in the household-level data sort across school districts, the predicted distribution of household characteristics in each community matches the actual

[34] It is worth noting that these two parts of the estimation procedure can in theory be combined for greater efficiency, but in practice because of the large set of parameters involved with the estimation of the underlying choice set, it is necessary to perform this stage of the estimation as a separate first stage.

distribution of these characteristics observed in the district-level data, as closely as possible. I now describe the details of the estimation procedure.

5 Estimation

I begin the description of the estimation procedure by introducing some notation that simplifies the exposition. Dividing the terms of the utility function into a *common-value* component, d_{csh}, an *interaction* component, m^i_{csh}, which includes any parts of the utility function that interact household and choice characteristics, and the idiosyncratic error term, e^i_{csh},, the utility function can be rewritten as:

$$V^i_{csh} = d_{csh} + m^i_{csh} + e^i_{csh}. \tag{8}$$

Here, the d_{csh} term captures those components of the utility function (including the unobservable characteristics) that are common to all households. This common-value component can be further decomposed into components attributable to the household's community choice, and *mean-zero* components attributable to the household's particular choice of school and residence within that community:

$$\begin{aligned} d_{csh} &= d_c + d_{cs} + d_{ch}, \quad \textit{where} \\ d_c &= a_{c0} X_c + a_{s0} \overline{X}_{cs} + a_{h0} \overline{X}_{ch} + x_c, \\ d_{cs} &= a_{s0}(X_{cs} - \overline{X}_{cs}) + x_{cs}, \quad \textit{and} \\ d_{ch} &= a_{h0}(X_{ch} - \overline{X}_{ch}) + x_{ch}. \end{aligned} \tag{9}$$

In the broadest specification of the model, each household must choose between 620 communities. In equation (9), then, d_c captures those common-value components of utility associated with the household's community choice. Each household must also choose between public and private school in its community. In an analogous way, then, d_{cs} captures those common-value components of utility associated with the household's school choice. Note that by construction d_{cs} is mean-zero within each community.

Each household must also choose a particular residence *h*. Because it is impossible to model the household decision over the full set of housing units in each community, I reduce the dimension of the available housing options to 22 distinct types in each community. Each housing type is characterized by tenure (2 categories) and annual housing cost (11 categories). So, for example, a possible choice of house would be a renter-occupied unit with annual rent of $10,000.

It is important to emphasize that the choice of such a unit in different locations will imply very different things about its unobserved quality. If this unit were located in an affluent suburban community, one would expect its unobserved quality to be quite low, while if it were located in a poor inner-city neighborhood, one would expect its unobserved quality to be quite high. More generally, then, there is a one-to-one relationship between a house's price and the unobservable quality component x_{ch}. In discussing the market clearing conditions earlier, I described how for any given price, the model predicts an implied demand for each house, given its characteristics (both observed and unobserved). The housing market clears by allowing the price of each house to adjust until its total demand is exactly one household. The estimation procedure will exploit precisely the reverse logic. That is, for any parameters of the model, given the price and observable characteristics of any house, the estimation procedure will solve for exactly the appropriate level of the unobserved characteristic so that the model perfectly predicts the demand for that type of housing unit. This one-to-one relationship between housing prices and unobservable quality x_{ch} also implies that discretizing the housing choice set by observable characteristics and price is essentially equivalent to discretizing the choice set by the observable characteristics and the unobservable characteristics.

Estimating the Utility Parameters

For expositional clarity, I distinguish between the *common-value* parameters contained in d and the *interaction* parameters subsumed in m. Given the specification of the utility function in equation (2), the identifying assumptions needed to estimate the common-value parameters are more stringent than those needed to estimate the interaction parameters (this will be made clear below). In fact, it is possible to estimate the interaction parameters along with the set of common-value utilities, (d_c, d_{cs}, d_{ch}), by simply attempting to match predictions of the model to observable moments of the data. I begin the discussion of the estimation of the utility parameters, therefore, by describing this procedure. I then discuss the additional assumptions that are necessary in order to identify the common-value parameters.

The estimation of the utility parameters begins by setting all of the interaction parameters (i.e., those in m^i_{csh}) to a set of starting values. For any combination of interaction parameters and common-value utilities, (d_c, d_{cs}, d_{ch}), the model predicts the probability that each household makes each particular choice *(c, s, h)*. Thus, summing this probability over all households and aggregating appropriately yields the predicted number of households in each community N_c, the predicted number of households in community c that choose private versus public school, N_{cs}, and that choose a house of type h, N_{ch}.

$$\hat{N}_c = \sum_i \sum_s \sum_h P^i_{csh}(d_c, d_{cs}, d_{ch})$$

(10) $$\hat{N}_{cs} = \sum_i \sum_h P^i_{csh}(d_c, d_{cs}, d_{ch})$$

$$\hat{N}_{ch} = \sum_i \sum_s P^i_{csh}(d_c, d_{cs}, d_{ch})$$

By comparing these predicted values to their actual counterparts in the data, it is possible to adjust the common-value utility components, (d_c, d_{cs}, d_{ch}), until the model's predictions for these values fit the data exactly. A simple extension of Berry (1994) ensures that there exists a unique set of common-value utility components (d_c, d_{cs}, d_{ch}) such that the model exactly predicts the number of households who choose each community-school and community-house combination.[35]

Berry (1994) not only proves that a unique solution exists to this system of equations but also provides a quick contraction mapping that solves for these common-value components.[36] An easy way to think about these common-value utilities is to think of them as fixed effects. In essence, Berry shows that we can vary the size of these fixed effects (making them larger for choices in which we are under-predicting demand and vice versa) until the predicted number of households that make each choice lines up directly with the actual number of households who make that choice. Given any set of interaction parameters, then, the first step of this part of the estimation procedure solves for the unique vector of common-value utilities that exactly clears the housing market.

Having calculated the set of common-value utilities, the model becomes completely deterministic. That is, for the given set of interaction parameters, I can now calculate the probability that each household i makes each choice (c, s, h), P^i_{csh}, as defined in equation (4). Summing this probability statement over all of the communities in a PUMA, R, yields the predicted probability that household i makes choice (R, s, h). This is exactly the decision that is observed for each household in the household-level data. An important first set of moments for the GMM procedure, then, is based on correctly predicting these choices. It is important to emphasize that this set of moment conditions is exactly analogous to ML estimation where we would attempt to maximize the model's ability to correctly predict the actual choice made by each household in the sample.

Because I do not observe each household's community choice directly in the household-level Census data, however, I also generate moment conditions that make use of the community-level Census data. For this type of moment condition I sum P^i_{csh} over various sets of households to generate predictions about the characteristics of the households that choose each district. More specifically, summing P^i_{csh} over

[35] Since the concept of utility only has relative meaning, it is necessary to tie down the utility of one element of the d_c vector.

[36] Note that the same logic that guides the construction of the proof in this case also implies that there exists a unique set of prices which clears the market.

only those households with a particular characteristic, z^r, (e.g., parental education – less than a high school degree) yields the predicted number of households of this type who make each choice *(c, s, h)*:

$$(11) \qquad \hat{N}^r_{csh} = \sum_{i \in Z^r} P^i_{csh}$$

Here *r* indexes the household characteristics that are included in the analysis. Summing the predicted number of household of type *r* who choose housing unit of type *h* over all housing units in the community yields the predicted number of households of type *r* that choose to send their children to public (and private) school in the community. This is precisely the type of information that can be observed in the district-level data available from the Census. Consequently, an additional set of moments is formed by attempting to match the populations of each household type predicted by the model to the actual number of these same households observed in choosing public (and private) school in each community. For example, if 1,000 households send their children to public school in a community and the racial composition of these households is 20 percent white and 80 percent Hispanic, the appropriate component of the distance function would be minimized if the model correctly predicted that 200 white households and 800 Hispanic households chose the community. Keeping in mind that the first step of the estimation procedure ensures that the model correctly predicts the 1,000 households who send their children to public school in this community, this part of the loss function is equivalent to correctly predicting the racial composition of the households electing public school in this community.

In this way, by choosing the parameters of the model (excluding the common-value parameters) to minimize an optimal function of the moment conditions that I have discussed so far, it is possible to estimate the entire set of interaction parameters without any additional economic assumptions. In particular, it is possible to estimate differences in how households trade-off between various choice characteristics without making any further assumptions about the processes that determine public school quality, crime rates, the demographic composition of each community and other potentially endogenous choice characteristics in equilibrium! It is important to note, however, that this is true only because I have assumed that all households gain the same utility from the unobservable characteristic of each choice. If I were instead to assume that a household's taste for these unobservable characteristics varied in some systematic way with household demographics additional moment conditions would be required to estimate these interaction parameters. As it is, additional moment conditions are needed for the estimation of the common-value parameters of the model (i.e., the parameters in equation (8)), and so I now turn to a discussion of these additional moment restrictions.

Summary of Procedure for Estimating Only Interaction Parameters

1. Fix the interaction parameters contained in m.
2. Solve for the set of common-value utilities (d_c, d_{cs}, d_{ch}) that ensure that the model correctly predicts (N_c, N_{cs}, N_{ch}) exactly.
3. Calculate a distance function based on:
   a. correctly predicting each individual household's decision (R, s, h) observed in household-level data (*a la* a traditional ML estimation procedure for estimating a multinomial logit).
   b. correctly predicting the demographic composition of public and private school enrollments in each community.
4. Repeat steps 1-3 until the GMM distance function is minimized.

Estimating the Common-Value Parameters

Using only the moment conditions discussed above, it is possible to estimate the interaction parameters of the model without further decomposing the common-value utility components. In fact, for the primary empirical exercise of this paper it is not necessary to identify the common-value utility parameters. For many applications (including all full-equilibrium counterfactual exercises), however, it will be necessary to distinguish these parameters and so I turn now to a discussion of the assumptions that are necessary for identification in this case.

As discussed earlier, given any set of interaction parameters, a unique set of common-value utilities ensures that the model correctly predicts the number of households that make each community-school and community-house combination. Having solved for the set of common-value utilities for any set of interaction parameters, the equations that define these utilities (equation 8) are simply a series of linked regression equations. In order to estimate the parameters of this equation, then, the natural thing to do would be to generate a set of covariance restrictions between a set of instruments, w, and the unobservable choice characteristics, x. Consequently, estimation of the baseline parameters follows directly from the construction of an appropriate set of instruments.

One approach (OLS equivalent) would be to assume that the observable characteristics X_j are uncorrelated with the unobservable characteristics, x_j. But, while this may be a reasonable assumption for many included choice characteristics, it is clearly not appropriate for any choice characteristics that are directly affected by how households sort across communities. Consequently, I divide the full set of choice characteristics into those that are not closely linked to sorting, X_1, and those that are indeed directly influenced by sorting, including housing prices, crime rates, public school quality, and perhaps most obviously the demographic composition of the community. I label this latter set of variables, X_2, and now need to find additional instruments to take the place of these variables in the estimation.

Before discussing the construction of this additional set of instruments, however, it is worth mentioning one difficulty that I am not able to deal with directly. That is, if households choose to live in a community based in part on the characteristics of its residents, then any unobservable characteristics of these residents belong in the error term of the utility function. Since these characteristics are systematically related to the observable district characteristics through the choice process itself, correlation would necessarily arise between all observable district characteristics and the error term. In light of this problem, the best I can do is to include as many variables that describe the demographic and socioeconomic composition of the community as possible in order to limit the size of this source of potential bias.

In forming an additional set of moment conditions for the estimation, what is required is a set of variables that are correlated with the observable sorting-dependent characteristics in X_2 of a particular choice, but which are not directly correlated with the unobservable characteristics of that choice. The set of instruments that I use is based on the observable *exogenous* choice characteristics (i.e., those in X_1) of neighboring choices. These variables will be correlated with the sorting-dependent characteristics such as demographic composition, because they directly impact how households sort across communities. For example, a community surrounded by neighboring communities with small houses is likely to attract higher income households than it would if its neighbors contained large houses. Thus, the exogenous characteristics of neighboring choices are clearly correlated with the sorting-dependent characteristics in X_2, meeting the first requirement of an appropriate instrument.

The other requirement for an appropriate instrument is that it be uncorrelated with the unobservable characteristics. But, by our original assumption that variables in X_1 are uncorrelated with the unobservable characteristics of their own choice, it follows directly that they would also be uncorrelated with the unobservable characteristics of neighboring choices. That is, if we are willing to assume, for example, that a community's air quality is uncorrelated with its own unobservable characteristics, it is almost certainly the case that this community's air quality is also uncorrelated with its neighbors unobservable characteristics.

In this way, then, a natural set of instruments rises organically out of the choice process itself. That is, as long as the researcher is able to find a set of variables (X_1) that can reasonably be assumed to be exogenous, it is possible to construct a set of additional instruments for any observed endogenous characteristics by using functions of these exogenous characteristics for neighboring choices. It is worth emphasizing that the concept of using neighboring characteristics need not be confined to the geographic sense of the word "neighbor" used here, but can also be construed in a similar way in product space.[37] The idea to use a product's isolation in product space as an instrument for price in estimating demand in a differentiated product market dates back to Bresnahan (1981, 1987) and is also exploited by Berry, Levinsohn, and Pakes (1995, 1998). It is also worth emphasizing that the solution to the "reflection problem" discussed in Manski (1991) offered here is possible only because of the discrete nature of the available choices.

---

[37] In fact, it is natural to think of geographic space as just another dimension of product space.

This also explains why identification is possible in this case when it is not always possible when researchers attempt to estimate hedonic models, which implicitly assume a continuous set of alternatives. In many ways, then, identification here is gained by exploiting the distribution of possible choices in the physical geographic landscape of California.

<u>Summary of Procedure for Estimating All Utility Parameters</u>

1. Fix the interaction parameters contained in *m*.
2. Solve for the set of common-value utilities (d_c, d_{cs}, d_{ch}) that ensure that the model correctly predicts (N_c, N_{cs}, N_{ch}) exactly.
3. Calculate a distance function based on:
    a. correctly predicting each individual household's decision *(R, s, h)* observed in household-level data (*a la* a traditional ML estimation procedure for estimating a multinomial logit).
    b. correctly predicting the demographic composition of public and private school enrollments in each community.
    c. *ensuring that the covariance of the set of exogenous choice characteristics and additional instruments* (X_1,w) *and the unobservable choice characteristics (x) is equal to zero.*
4. Repeat steps 1-3 until the GMM distance function is minimized.

6 Results (Based on the Previous Version of Paper)

I now present the results of my analysis. This section begins by describing the specification of the model taken to the data and presenting parameter estimates. These parameter estimates are then employed to calculate the willingness of households with various characteristics to pay for a variety of goods that enter the utility function, including public school quality, commuting distance, and public safety. The following section uses these results to run a set of simple simulations that predict the school quality consumption patterns that would result if households based their residential location decisions and school choices on only a subset of the factors considered in the analysis. By comparing the patterns of consumption predicted under these alternative assumptions to the actual consumption patterns observed in the data, I am able to distinguish the factors that drive the observed differences in the consumption of school quality.

The results presented for the demand side estimates are based on a sample of 5,000 households drawn at random from my household sample.[38] These Census data provide information about the household's employment location only to the granularity of the Census PUMA. To refine the geographic granularity of this employment location data, I draw an employment zipcode for each household based on the distribution of employment by zipcode in each PUMA. The household is

[38] Importantly, the Census weights are used for this selection, so that the drawn sample is a representative sample of households with school-aged children in California.

assigned the latitude and longitude for the center of the randomly selected zipcode and this information is used to calculate the household's commuting distance to each community in the sample.

Each household chooses *(c, s, h)* to maximize its utility function:

$$(12) \qquad \underset{(c,s,h)}{Max} \quad V^i_{csh} = a_c X_c + a_s X_{cs} + a_h X_{ch} + a^1_D D^i_c + a^2_D D^{i^2}_c + a^t_p \ln(y_i - p_{ch}) + x_c + e^i_{csh}$$

First note that this previous version of the model that I had taken to the data uses only five housing possibilities for each community-tenure combination and allowed for only a community-level unobservable component x_c. The community choice characteristics (X_c) included in the analysis include the FBI violent crime index, the fraction of houses in an urban area, an air quality index, and the average high temperature in July. In addition, the community's location enters the utility function through the determination of the household's commuting distance D^i_c, and squared commuting distance.[39]

The school choice (X_{cs}) characteristics include a private school dummy (if private schooling is available in the community), the average test score of the students who attend the district's public schools, and the demographic and socioeconomic characteristics of the child's peers for each school choice.[40] In this specification, the parameters that multiply the private schooling dummy variables describe the household's valuation of the private schooling options relative to public schooling, while the parameters that multiply the public school test score describe the increment to utility associated with choosing a slightly better public school district. Importantly, this interpretation of the parameters that multiply the public school test score as the households' valuation of school quality relies critically on the fact that I have explicitly controlled for the characteristics of the child's peer groups in the analysis.[41]

The housing characteristics (X_{ch}) include the number of rooms in the house or apartment and the age of the structure. The annual price of the housing unit also enters the analysis through the all other consumption term $(y_i - p_{cth})$. Importantly, I

---

39 The inclusion of the squared distance term allows a household's valuation of communities a great distance from its employment location to decrease rapidly. This feature of the model serves the function of implicitly reducing the relevant community choices available to the household to the geographic scope that one would expect the choice to be restricted.

40 In the specification of the model presented here, I constrain a household's valuation of the characteristics of its child's peers in either private or public school to be the same. That is, a classroom of children from a highly educated background enters the utility function in the same way regardless whether the child is in a private school or public school. I explore specifications of the model that allow the household to value private school peers differently than public school peers in the next chapter.

41 Note that in completely unconstrained version of the way that the characteristics of the child's peer group enters the utility function, I am unable to distinguish the extent to which a household values its child's peer group for its effect the child's achievement versus other possible peer effects.

allow the household's valuation of the housing characteristics and other consumption term to vary with its tenure choice because of difficulties comparing the actual housing costs associated with renting versus owning.

The household characteristics that enter the analysis (z^i_r) include the following dummy variables: race – black, race – Hispanic, parental education – high school or less, and household income – $40,000 or less. The household's income also enters the analysis directly through the construction of the other consumption term (y_i-p_{cth}), and its primary employment location enters the utility specification through the construction of the commuting distance term. In addition to these household and choice characteristics, I also estimate the four inclusive value terms associated with the nested logit framework.

In order to estimate the set of community baseline parameters (as described in Section 5), I need to specify an appropriate set of instruments. For the specification of the model presented in this section, I use the average housing, air quality, climate, and urban density variables for the three closest communities to the community in question.

The parameter estimates for the full model are presented in Table 5. Before examining these estimates, it is important to clarify how one should read them. The primary set of coefficients multiply the components of the utility function that describe the household's community, school, tenure, and housing choices. The first column of the table shows the estimated coefficients for the baseline group (highly educated, high income, white or Asian households), which describe how the households in this group substitute between the various choice characteristics. The remaining columns depict the estimates of the incremental parameters, which are added to the baseline estimate if the household has the appropriate characteristic. In this way, the coefficients that multiply the choice characteristics in the utility function of a black household exactly like the baseline group in every way except race would combine the baseline parameter estimates of first column with the race-black parameter estimates of the second column. The last four rows of Table 5 show the estimates of the inclusive value (s) parameters, which determine households substitution patterns between choices within versus across branches in the decision tree depicted in Figure 2. Values close to one imply that the household substitution patterns are substantially different, while values close to zero reduce the nested framework of Figure 2 to the multinomial logit framework. A quick examination of Table 5 reveals that the precision of my estimation with a sample of 5,000 households is quite good, with a majority of the would-be *t*-statistics residing in the 2 to 5 range. I now present the primary analysis of this paper, which attempts to distinguish the underlying forces that drive the observed differences in the consumption of public school quality by households with different characteristics and endowments.

Distinguishing the Basis for the Observed Differences in the Demand for School Quality

I begin by comparing the average consumption of school quality (and other local goods) for the various household groups considered in the analysis. These consumption patterns are shown in Table 6. A comparison of the consumption patterns of black and Hispanic households with those of Asian and particularly white households reveals that black and Hispanic households consume, on average, less school quality, public safety, air quality, and housing size. In addition, black and Hispanic households are less likely to send their children to private school and more likely to rent rather than own. Interesting, black renters tend to live in much smaller apartments than white renters, while paying similar amounts in rent.[42] A comparison of the consumption patterns between less educated and more educated households reveals large differences in schooling preferences (particularly between the public and private sectors of the market). The other consumption differences between education groups follow a similar pattern to the differences observed across households of different races, although the differences across the two education groups defined here are not as pronounced. Finally, comparing the consumption patterns of households with more than $40,000 in income to those with less reveals that high-income households purchase much more expensive homes and apartments, on average. High-income households also show a very high proclivity for home-ownership and private schooling. Interestingly, the differences in the consumption of public safety across both income and education groups are not nearly as striking as the differences across racial groups.

Which Household Characteristics Drive the Observed Location Patterns?

While the figures of Table 6 shows location patterns for various subsets of households (i.e., those with a particular characteristic), they do not distinguish the independent contributions made by race, education, and income to the observed patterns. For example, the fact that Hispanic households consume less school quality than white households may be explained by differences in the preferences associated with race, but might also be explained by differences in the education and income levels of Hispanic relative to white households. In order to distinguish the independent effect of a household's income, education and race on consumption, I present a series of consumption patterns predicted by the model that allow preferences to vary only according to a particular household characteristic.[43] These patterns are shown in Table 7.

To distinguish the independent impact of race, education, and income on location patterns, I consider three simple experiments that predict the school quality

---

[42] At least some of this disparity can be explained by the fact that black households are much more likely to live in urban areas (not shown in the table).

[43] As in Table 7, in order to construct predicted consumption patterns that hold particular household characteristics constant, I select 1,000 draws from the joint distribution of these characteristics in the household-level data set. For each household in the sample I then average the consumption of school quality under each draw in order to calculate the predicted consumption for each household.

consumption patterns that would result if households shared many of the preferences that affect the location decision, but differed only according to their race, education, and income, respectively.[44] By comparing the predicted consumption patterns in these various states, I can distinguish the independent impact of these characteristics in determining a household's consumption of public school quality.

The first column shows the consumption patterns predicted by the model when the full set of parameter estimates is used. A quick glance at this column reveals that poorly educated, low-income, and minority households consume relatively low amounts of public school quality. The figures in this column match the actual consumption of school quality observed in the data quite well (compare to Table 6), implying that my model does a fairly good job of explaining the observed consumption patterns. The patterns of consumption depicted in this table display a symmetry that is hardly surprising given the correlation between race, education, and income across households. For each household characteristic – race, education, and income – preferences associated with the characteristic itself explain only about 50-60 percent of the observed differences in the consumption of school quality, while preferences associated with the other household characteristics explain a sizable portion of the overall differences in consumption. It is important to remember that in these calculations, the preferences associated with a particular household characteristic vary for the full set of choice characteristics. In this way, the predicted differences across households with different racial characteristics attributable to preferences associated with race, for example, are based on differences in preferences for housing and peer characteristics as well as school quality.

Which Choice Characteristics Drive the Observed Location Patterns?

I now consider a set of simple experiments that predict the school quality consumption patterns that would result if households shared many of the preferences and endowments that affect the location decision. The simple experiments considered in Table 7 predict the consumption patterns that would result if household differed only in their employment locations, and only in their preferences for school quality, housing characteristics, and peer characteristics, respectively.[45] By comparing the predicted consumption patterns in these various

[44] I call these *simple* experiments because I do not actually compute the new sorting equilibrium that would result if households had different preferences and sorted accordingly. Instead, each household in the sample continues to choose from the actual set of choices observed in the data, including the socioeconomic characteristics of its neighbors and child's peers. Similar experiments to the ones considered here could be completed for other local goods as well, but these are not reported here for the sake of brevity.

[45] I call these *simple* experiments because I do not actually compute the new sorting equilibrium that would result if households had different preferences and sorted accordingly. Instead, each household in the sample continues to choose from the actual set of choices observed in the data, including the socioeconomic characteristics of its neighbors and child's

states, I can distinguish the factors that drive the observed differences in the consumption of public school quality.

The first column again shows the consumption patterns predicted by the model when the full set of parameter estimates is used. The second column shows the patterns of school quality consumption that would result if the full sample of households had identical preferences, but differed in their employment location.[46] These results imply that about 15 percent of the racial school quality consumption gaps and over 20 percent of the income gap can be explained by differences in the geography of employment opportunities.[47]

The final three columns of Table 8 show the consumption patterns predicted by the model if households had identical employment locations and identical preferences except for those associated with school quality, housing characteristics, and peer characteristics, respectively.[48] When households differ only in their preferences for school quality, the predicted differences in the consumption of school quality between minority households and white/Asian households are substantially smaller than those observed in the data. In fact, differences in preferences for school quality can explain only about 24 percent of the black-white/Asian gap and 30 percent of the Hispanic-white/Asian gap. Differences in preferences for school quality do, however, drive the majority of the observed differences in the consumption of school quality across households with different parental education levels. In this case, preferences for school quality itself explain over 70 percent of the overall consumption gap. In addition, only about 44 percent of the school quality consumption gap across income groups can be explained by differences in preferences for school quality itself.

The final two columns show the differences in consumption of school quality that can be attributed to housing and peer characteristics. These patterns reveal how these other aspects of the decision-making process that determines a household's residential location and school choice affect its consumption of school quality.[49] In the case of the differences in the consumption of school quality associated with race, it is immediately obvious that differences in preferences for

---

peers. Similar experiments to the ones considered here could be completed for other local goods as well, but these are not reported here for the sake of brevity.

[46] For the figures shown in Table 7, I use the weighted average of the preference parameter estimates to construct the same set of preferences for all households in the sample.

[47] I construct the figures shown in Table 7 by sampling 1,000 draws from the spatial distribution of employment. For each household in my sample, then, I average that household's consumption of school quality for each employment location draw in order to calculate the household's predicted consumption of school quality.

[48] For the purposes of these calculations, I draw a random sample of employment locations from the distribution of these characteristics in the household sample.

[49] It is important to remember that because a household must choose from a set of discrete bundles of local goods and geographic relationships when making its residential location decision and school choice, its actual consumption of any local good is determined not by its preferences for that good alone, but by the confluence of its preferences for the full set of local goods and geographic relationships and the nature of the set of available bundles.

housing and peer characteristics drive a substantial potion of the observed differences in the consumption of school quality. A similar, although not as striking, result also holds for the differences in consumption attributable to income.

It is important to note that the differences in what I measure as *preferences* for housing characteristics do not necessarily represent differences in actual preferences across racial groups for these characteristics. Instead, the differences that present themselves as differences in housing may arise from discrimination in the housing or credit markets or other aspects of these markets that restrict the set of location and housing choices actually available to minority and low-income households. Government-subsidized housing, for example, has historically limited the set of housing and location choices available to low-income households who wish to participate in the program to very small subsets of the full urban land and housing markets. Unfortunately, I am unable to distinguish between these various explanations for what appears to be differences in preferences in my framework. Regardless of the explanation for the difference in demand for housing, however, the principal conclusion of this chapter stands – differences in the demand for housing drive a substantial portion of the differences in the consumption of school quality across households from different racial backgrounds.

In sum, then, a sizable portion of the racial and income gaps in the consumption of school quality can be attributed to differences in preferences for housing and peer characteristics as well as differences in the geography of the residential location decision. In fact, differences in preferences for school quality itself explain only 20-30 percent of the observed differences across households of different races. The majority of the education gap, on the other hand, is explained by differences in preferences for school quality itself, which drive almost 70 percent of the school quality consumption gap across education categories.

7 Some Policy Implications and Concluding Remarks

The results of my analysis have important implications for policy as they suggest that the economy is systematically under-providing school quality to minority and low-income households, relative to their preferences. In the short-run, policies aimed at increasing the mobility of those households most constrained by the current economic environment should lead to direct welfare improvements for these households. Among other things, these policies might attempt to free the household's school decision from its housing decision (through, for example, magnet schooling, open enrollment with certain segments of the urban economy, or targeted voucher systems), or might attempt to improve the flexibility of the housing decision for those who participate in government-assisted housing programs. The important contribution of this work to these on-going policy debates is to highlight those segments of the population who stand to gain the most from increased mobility. In this way, the targeting of these programs to those households most affected by the current economic environment is likely to get the greatest bang for the buck.

In the long-run, however, policies that attempt to address directly the fundamental imperfections in the provision of local public goods created by the imbedding of this provision within the structure of the urban economy may prove most effective. Unfortunately, the results presented here provide only one piece of the puzzle – namely, estimates of the household preferences for the components of the bundle of local goods and housing characteristics. A complete understanding of the root causes of this market imperfection requires a deeper understanding of the incentives that school districts have to provide quality in various segments of the urban land, labor, and housing markets. These incentives almost certainly depend directly on the structure of the school finance system as well as the competitiveness of a school district's particular location within the geographic and product space of the urban economy. Thus, future research in these areas, in combination with the analysis presented here, has the potential to greatly increase my understanding of the underlying determinants and consequences the geography of school quality provision within the urban economy.

Appendix 1 – Data Sources

The data for this paper were drawn from a wide variety of data sources and transformed in a variety of ways to best serve the analysis undertaken in the paper. This appendix documents both the data sources and the methods that I have used in constructing the full dataset.

A. School District-Level Census Data

The primary source of school district level Census data is the School District Data Book (SDDB) published by the National Center for Education Statistics (NCES). This data set combines school finance data with a special tabulation of the 1990 Census by school district. This special tabulation of the Census provides aggregate economic and demographic information at the district level. Moreover, I am able to restrict the sample in two important ways. First, I am able to isolate the characteristics of only those households with school-aged children and more specifically school-aged children within specific grade ranges. Secondly, I am able to separately identify households that send their children to public school from those that send their children to private school.[50] I select the following variables for all households with children in grades 5-8, distinguishing households with children in public school from those with children in private school:

- Parental Education Levels (4 categories): the fraction of households with: less than high school degree, high school degree, some college, college degree or more.
- Race/Ethnicity Categories (4 categories): the fraction of households: Hispanic, Asian (non-Hispanic), Black (non-Hispanic), White and all other races (non-Hispanic).
- Linguistic Isolation (2 categories): the fraction of households in which English is spoken in the home.
- Household Type (4 categories): the fraction of households: two parent household, single male parent, single female parent, non-family structure

In addition, the SDDB also provides the location of the district, its population density, and the percentage of homes that are in an urban area. In my analysis, I use the latitude and longitude of the population-weighted center of gravity (as opposed to the strict geographic center of gravity) of the district.

Because some counties did not provide the NCES with boundary files for the school districts within the county, these specially tabulated data are not available for all 58 counties in California. Consequently, it was necessary to restrict the sample to the following 30 counties:

> Alameda, Contra Costa, Fresno, Imperial, Kern, Lake, Los Angeles, Marin, Mendocino, Merced, Orange, Placer, riverside, Sacramento, San Bernardino, San Diego, San Francisco, San Joaquin, San Luis Obispo, San Mateo, Santa Clara, Santa Cruz, Shasta, Solano, Sonoma, Stanislaus, Sutter, Tulare, Ventura, and Yuba.

This list includes all of the major counties of the San Francisco Bay Area and Southern California (with the possible exceptions of Napa and Santa Barbara) are included in the sample. In fact, most of the 28 missing counties are in rural mountainous regions of California and therefore contain very few school districts or people. Consequently, these missing data do not seem to pose any serious problems for my analysis.

---

[50] Due to data limitations, many previous studies have been forced to simply use the average characteristics of the full district's population as proxies for the characteristics of the students' households.

B. Housing Data

I also draw data describing distributions of housing prices and characteristics by tenure (rented vs. owned) for each district from the SDDB:

- Housing Tenure (2 categories): the fraction of housing units owner versus renter occupied.
- Housing Values and Rents: The SDDB contains information that allows me to determine the fraction of houses in each of 20 value categories and 20 rent categories.
- Age of Housing Units (8 categories): the fraction of housing units built in the years: pre-1940, 1940-9, 1950-9, 1960-9, 1970-9, 1980-4, 1985-8, and 1989-90, for each tenure type.
- Housing Size: I construct a measure of the average number of rooms per house using two measures provided in the data source: the average number of persons per room and the average number of persons per house.

Together these data provide information about the marginal distributions of housing prices and characteristics in each school district.

C. 1990 Decennial Census – Public Use Microdata Sample

The individual household level data are drawn from the One-Percent PUMS of the 1990 Census. In order to match the 5-8th grade range of the SDDB, I draw data for all households with children aged 10–13. These data include the same household level demographic and economic variables discussed above (household income, education, race, etc.) as well as whether the household sends its children in this age range to public or private school. In addition, I draw two variables concerning the household's location: the household's Place-of-Residence Public Use Microdata Area (PUMA) and the Place-of-Work PUMA for each working member of the household. PUMAs are Census areas that contain at least 100,000 individuals. Consequently, these data provide some (if somewhat imprecise) information about the household's residential and employment locations. Finally, I draw information about the household's housing and tenure decision (rent or own) including the housing price, number of rooms, and the age and type of structure.

D. Achievement Scores and Other Schooling Data

The California Department of Education (CDE) is the source for most of the data that relate to the operation of the school districts themselves. The variables drawn from this source include:

- California Achievement Test Scores: I use each district's average sixth grade test score on the California Achievement Test (which includes a math, reading, and writing components). In order to remain as consistent as possible with the demographic data from the Census and SDDB, I average these scores over the school years 1987-1991. In this way, (to as great an extent possible), both the household- and district-level demographic data as well as the achievement test scores are associated with the same cohorts of children.
- School Size Measures: average class size, average school size, and pupil-teacher ratio.
- Expenditures: the fraction of expenditures spent on instruction, support services, and fraction of spending which is current versus capital improvements.
- Employee Composition: the fraction of full-time-equivalent (FTE) employees dedicated to teaching, administration, and pupil services. These are mutually exclusive categories.

E. Climatological and Air Quality Data

The source of the climate data used in the analysis is the Istern Regional Climate Center. These data are drawn from the 350 climatological monitoring stations in California and are historical averages over the history of the monitoring station (usually 80-100 years). I collected the following variables: January and July average high and low temperatures, average annual precipitation and snowfall. In addition, I also recorded the location of the monitoring station (latitude and longitude). In order to incorporate the climatological variables into the analysis, it was necessary to convert these observations at the monitoring station-level to observations at the school district-level. Using the precise locations of the monitoring stations and school district center of gravities, I assign each district the weighted average of the data from the two closest monitoring stations. The weight given each station's value is the inverse of the distance between the district and monitoring station. In this way, if a district lies between two stations at a similar distance from the district center, the district will be assigned the average of the two stations. If, on the other hand, a single monitoring station is located near a district, the data assigned to the district will be based primarily on the data from that single close station.

The source of the air quality data is the United States Environmental Protection Agency. I collected data from 206 monitoring stations in California specifying the fraction of ground-level air that consists of ozone, sulfur dioxide, nitrogen dioxide, carbon monoxide, and particulate matter. These five *criteria pollutants* are those for which the EPA has set health-based standards. For each monitoring station, I collected the arithmetic mean of daily observations for the year 1990.

Unfortunately, each station does not provide information for each criteria pollutant. Consequently, I constructed a pollution index for each station in the following way.

1. For each observed pollutant at a station, I calculate a relative air quality measure based on the observed pollution level for that monitoring station divided by the average level of the pollutant at all monitoring stations (e.g., a station with an average amount of pollution is assigned a value of 1.00).
2. For each station, I average the available relative pollution measures to construct an overall pollution index. Consequently, a station that measures only three pollutants would be assigned the weighted average of the relative measures for those three pollutants.

As in the case of the climatological data, I must convert these monitoring station-level data to the school district-level. I use the same procedure described for the climate variables.

F. Crime Data

The crime rate data are drawn from the California Department of Justice. I collected two indices for the analysis: the California crime index and the FBI crime index. The California index measures a group of serious offenses including willful homicide, forcible rape, robbery, aggravated assault, burglary, and motor vehicle theft. The FBI index includes these same crimes but also adds other homicides and larceny-theft. The crimes in the California index are often labeled violent crimes while the crimes involving larceny or theft are often labeled property crimes. I collected crime data for the years 1988-1992 and used an average measure across these years in the analysis in order to avoid measurement error due to the discrete nature of the data – particularly in small communities. The crime data are available for 433 jurisdictions in California. Consequently, I assign each school district's crime variables following the same procedure used to assign the climate and air quality measures.

Appendix 2 – More Details Concerning School Finance in California

Until the early 1970's, California's school districts were funded almost exclusively by local property taxes. Beginning in 1971, however, a series of landmark decisions by the Supreme Court of California and legislation by the California legislature transformed this locally financed system into one in which the majority of funding is provided by the state. In the first *Serrano v. Priest* decision in 1971, the Supreme Court decreed that the use of a local property tax for the financing of education violated the "equal protection" clause of the California Constitution. The California Supreme Court did not designate a specific finance system to replace local property taxation, and so in response to the decision, the state government instituted a school aid formula that provided more money to relatively poor districts. The state government also set revenue limits for each district, which capped the amount of spending per student. These limits were then to be increased at a faster rate for poorer communities so that funding across districts would eventually be better equalized. Importantly, however, a school district could spend more than the revenue limit, if its constituents voted for such an override in a local referendum.

Unsatisfied by these reforms, the Supreme Court of California defined tougher and clearer rules for a satisfactory school finance reform in a second *Serrano v. Priest* decision in 1976. This decision essentially required that total local and state expenditures should not differ by $100 per student (in 1976 dollars) across school districts. While the court did not explicitly mandate this $100 differential as the sole determinant of satisfactory compliance with its decision, this clearly-defined target soon became the only factor for determining compliance. The California State Legislature soon prepared to institute reforms to meet this requirement. In 1978, however, when legislation designed to equalize funding across communities was about to take effect, the voters of California passed Proposition 13 (which included an amendment to the Constitution) by a 2-1 margin in a statewide referendum. Proposition 13 severely limited the ability of local districts to generate revenue from property taxation and subsequent legislation set in place the method of school financing that exists in California today. Further description and analysis regarding California's school finance system, Proposition 13, and the *Serrano vs. Priest* decisions may be found in Downes (1992), Downes and Greenstein (1996), Fischel (1989), Rubinfeld (1995), or Silva and Sonstelie (1995).

The revenues of a school district may exceed the revenue limit set by the state for three basic reasons. First, a district may receive federal aid, provided through the Elementary and Secondary Education Act of 1965 (commonly known as the Title I act). The amount of aid is based on the socioeconomic status of a district's students and by law can not substitute for local funding. Second, some school districts are associated with private foundations that raise additional money for the district. While Brunner and Sonstelie (1997) find that the average amount raised by these foundations per child statewide is only about twenty dollars, this money is concentrated in a small number of wealthy districts and may, therefore, have an important impact on the finances of these districts. Finally, the amount of revenues raised through the one percent effective property tax exceeds the revenue limit in a small percentage of the wealthiest districts. Although these districts (known as "basic aid" districts) receive only the minimum revenues from the state, total revenues exceed the revenue limit. Currently, 8 percent of California school districts are designated "basic-aid" districts.

References

Anas, Alex, (1982), *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics and Public Policy Analysis*, Academic Press, New York.

Anas, Alex, and Chausie Chu, (1984), "Discrete Choice Models and the Housing Price and Travel to Work Elasticities of Location Demand," *Journal of Urban Economics*, Vol 15, pp. 107-123.

Bartik, Timothy, (1987),

Bayer, Patrick, (2000a), "Estimating the Education Production Function in the Presence of Tiebout Sorting," mimeo, Yale University.

Bayer, Patrick, (2000b), "Estimating the Determinants of the Public-Private School Choice in the Presence of Tiebout Sorting," mimeo, Yale University.

Bayer, Patrick and Robert McMillan, (in progress), "The Determinants and Consequences of Competition in Local Schooling Markets," work in progress, Yale University, Public Policy Institute of California, and University of Toronto.

Bayer, Patrick, Robert McMillan, and Kim Rueben, (in progress), "Examining Household Residential Location and Schooling Decisions in California," work in progress, Yale University, Public Policy Institute of California, and University of Toronto.

Bayer, Patrick and Christopher Timmins, (2000), "The Equilibrium Properties of Random Utility Models with Endogenous Sorting," mimeo, Yale University.

Bergstrom, Theodore C., Daniel L. Rubinfeld, and Perry Shapiro, (1982), "Micro-based Estimates of Demand Functions for Local School Expenditures," *Econometrica*, Vol. 50, pp.1183-1205.

Berry, Steven, (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, Vol. 25, pp. 242-262.

Berry, Steven, James Levinsohn, and Ariel Pakes, (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, Vol 63, pp. 841-890.

Berry, Steven, James Levinsohn, and Ariel Pakes, (1998), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: Autos Again," mimeo, Yale University.

Blackley, Paul and Jan Ondrich, (1988), "A Limiting Joint-Choice Model for Discrete and continuous Housing Characteristics," *The Review of Economics and Statistics*, Vol. 70, pp. 266-274.

Boozer, Michael and Cecelia Rouse, (1999), "," mimeo, Yale University.

Bradford, David J., and Harry H. Kelejian, (1973), "An Econometric Model of the Flight to Suburbs," *Journal of Political Economy*, Vol. 81, pp. 556-589.

Bradford, D., R. Malt, and W. Oates, (1969), "The Rising Cost of Public Services: Some Evidence and Reflections," *National Tax Journal*, Vol. 22, pp. 185-202.

Bresnahan, Timothy, (1981), "Departures from Marginal-Cost Pricing in the American Automobile Industry," *Journal of Econometrics*, Vol. 17, pp. 201-227.

Bresnahan, Timothy, (1987), "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," *Journal of Industrial Economics*, Vol. 35, pp. 457-482.

Brunner, Eric, and Jon Sonstelie, (1998), "Coping with *Serrano*: Voluntary Contributions to California's Local Public Schools," mimeo, San Diego State University.

Card, David, and Alan Krueger, (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, Vol. 100.

Deacon, R., (1977), "Private Choice and Collective Outcomes: Evidence from the Public sector Demand Analysis," *National Tax Journal*, Vol. 30, pp. 371-386.

Downes, Thomas, (1992),

Downes, Thomas and Shane Greenstein, (1996),

Downes, Thomas A., and Thomas P. Pogue, (1994), "Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students," *National Tax Journal*, Vol. 47, pp. 89-110.

Dynarski, Mark, Robert Schwab, and Ernest Zampelli, (1989), "Local Characteristics and Public Production: The Case of Education," *Journal of Urban Economics*, Vol. 26, pp. 250-263.

Ellickson, B., (1981), "An Alternative Test of the Hedonic Theory of Housing Markets," *Journal of Urban Economics*, Vol. 9, pp. 56-79.

Epple, Dennis, (1987),

Epple, D., R. Filimon, and T. Romer, (1984), "Equilibrium Among Local Jurisdictions: Towards an Integrated Approach of Voting and Residential Choice," *Journal of Public Economics*, Vol. 24, pp. 281-304.

Epple, D., R. Filimon, and T. Romer, (1993), "Existence of Voting and Housing Equilibrium in a System of Communities with Property Taxes," *Regional Science and Urban Economics*, Vol. 23, pp. 585-610.

Epple, Dennis and Thomas Romer, (1978), ","

Epple, Dennis, and Holger Sieg, (1998), "An Approach for Estimating Spatial Models," mimeo, Carnegie Mellon University.

Evans, William, Wallace Oates, and Robert Schwab, (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy*, Vol. 100, pp. 966-991.

Fischel, William A., (1989), "Did *Serrano* Cause Proposition 13?" *National Tax Journal*, Vol. 42, pp. 465-473.

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

Hamilton, Bruce, (1975a),

Hamilton, Bruce, (1975b),

Hamilton, Bruce, (1976),

Inman, Robert P., (1979), "The Fiscal Performance in Local Governments: An Interpretive Review," in *Current Issues in Urban Economics*, ed. by P. Mieszkowski and M. Strazheim, Baltimore, Johns Hopkins Press.

King, A. Thomas and Peter Mieszkowski, (1973), "Racial Discrimination, Segregation, and the Price of Housing," Journal of Political Economy, Vol. 81, pp. 590-606.

Manski, Charles F., (1993), "Identification of Endogenous Social Effects: The Reflection Problem," The Review of Economic Studies, Vol. 60, pp. 531-542.

McFadden, Daniel, (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, eds., *Frontiers of Econometrics*, Academic Press, New York.

McFadden, Daniel, (1978), "Modeling the Choice of Residential Location," in eds. Karlquist, A., et al., Spatial Interaction Theory and Planning Models, Elssevir North-Holland, New York.

Nechyba, Thomas J., (1997), "Existence of Equilibrium and Stratification in Local and Hierarchical Tiebout Economies with Property Taxes and Voting," *Economic Theory*, Vol. 10, pp. 277-304.

Nechyba, Thomas J., (1999), "School Finance Induced Migration and Stratification Patterns: the Impact of Private School Vouchers," *Journal of Public Economic Theory*, forthcoming.

Nechyba, Thomas, Patrick McEwan, and Dina Older-Aguilar, (1999),

Nechyba, Thomas J., and Robert P. Strauss, (1998), "Community Choice and Local Public Services: A Discrete Choice Approach," *Regional Science and Urban Economics*, Vol. 28, pp. 51-73.

Nevo, Aviv, (1998), "A Research Assistant's Guide to Random Coefficients Discrete Choice Models of Demand," mimeo, University of California, Berkeley.

Oates, Wallace, (1969), "The Effects of Property Taxes and Local Public-Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis," *Journal of Political Economy*, Vol. 77, pp. 957-971.

Pollakowski, Henry O., (1973), "The Effects of Property taxes and Local Public Spending on Property Values: A Comment and Further Results," *Journal of Political Economy*, Vol. 81, pp. 994-1003.

Quigley, John M., (1985), "Consumer Choice of Dwelling, Neighborhood, and Public Services," *Regional Science and Urban Economics*, Vol. 15(1).

Rosen, H., and D. Fullerton, (1977), "A Note on Local Tax Rates, Public Benefit Levels, and Property Values," *Journal of Political Economy*, Vol. 85, pp. 433-440.

Rubinfeld, Daniel L., (1995), "California Fiscal Federalism: A School Finance Perspective," Paper prepared for California Constitutional Reform Conference.

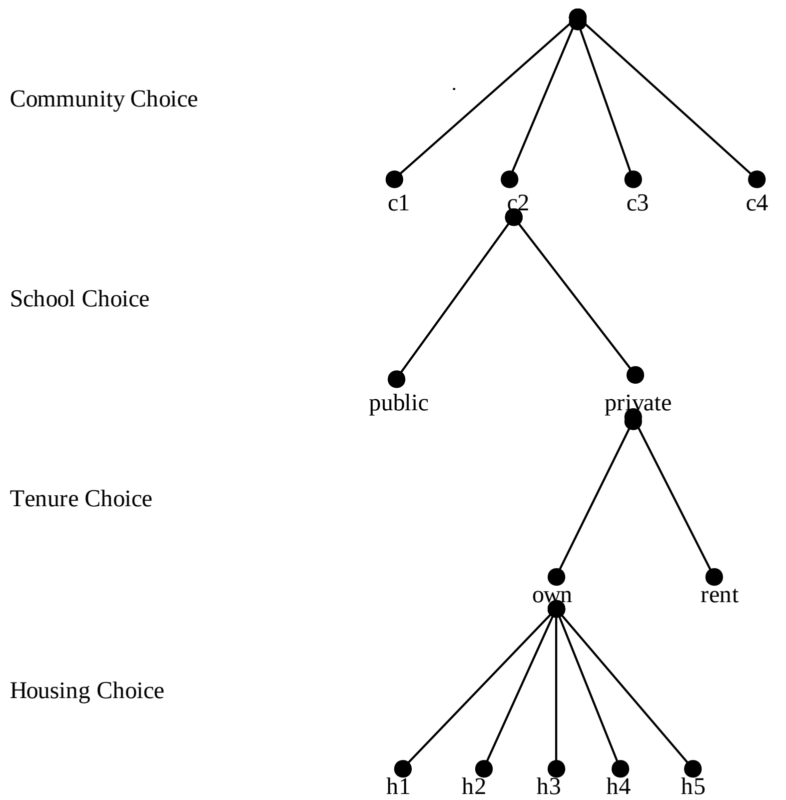
Rubinfeld, Daniel L., Perry Shapiro, and Judith Roberts, (1987), "Tiebout Bias and the Demand for Local Public Schooling," *The Review of Economics and Statistics*, pp. 426-437.

Silva, F., and J. Sonstelie, (1995), "Did Serrano Cause Decline in School Spending?" *National Tax Journal*, Vol. 48, pp. 199-215.

Tiebout, Charles M., (1956), "A Pure Theory of Local Expenditures," *Journal of Political Economy*, Vol. 64, pp. 416-424.

Witte, A., H. Sumka, and H. Erekson, (1979), "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets," *Econometrica*, Vol. 47, pp. 1151-1174.

Figure 2: The Structure of the Household's Decision Making Process



Note: As modeled, each household's choice set consists of 620 communities, 2 school choices per community, 2 tenure choices per community, and 5 housing choices per community-tenure combination. The nested structure of the household's decision-making process implies that each household first makes its community choice, then its school choice, then its tenure choice, and finally its housing choice. The nested logit specification allows for the possibility that a household may not fully count the utility that it receives from the downstream choices when making an upstream decision.

Table 1: Average Household Characteristics in the Two Organizations of the Census Data

	Household Sample[1]		District Sample[2]	
	Public	Private	Public	Private
Share	0.864	0.089	0.883	0.118
Race/Ethnicity:				
Black	0.080	0.069	0.092	0.079
Hispanic	0.361	0.224	0.351	0.242
Asian	0.089	0.094	0.107	0.114
White or Other Race:	0.470	0.613	0.450	0.564
Parental Education Levels:				
< HS degree	0.318	0.119	0.304	0.128
HS degree	0.191	0.135	0.190	0.130
Some college	0.304	0.331	0.304	0.341
College degree	0.187	0.415	0.201	0.401
Household Income:				
< $20,000	0.256	0.106	0.233	0.092
$20-40,000	0.298	0.195	0.288	0.195
$40-60,000	0.219	0.244	0.232	0.254
$60-80,000	0.115	0.172	0.183	0.279
> $80,000	0.112	0.283	0.064	0.180
Household Size:				
1-4	0.153	0.182	0.177	0.222
5-6	0.662	0.752	0.540	0.603
7+	0.185	0.066	0.283	0.175
Primary Language:				
English	0.507	0.615	0.519	0.579
Spanish	0.353	0.220	0.331	0.235
Linguistically Isolated:	0.151	0.060	0.131	0.057
Household Type:				
couple	0.724	0.822	0.723	0.815
single-male	0.068	0.021	0.058	0.033
single-female	0.208	0.157	0.211	0.148

[1]*Source: 1990 Decennial Census*
[1]*Note:* The values in the table represent the average characteristics of all households in the sample with those with children aged 10-13 in the year 1990.

[2]*Source: School District Data Book*: Special Tabulation of the 1990 Census by School District
[2]*Note:* The values for the district sample represent the average characteristics of households with children in grades 5-8 in the year 1990.

Table 2: Average Community Characteristics

	Unweighted Mean	Standard Deviation	Weighted Mean
Climate and Air Quality:			
January high temperature	58.81	5.83	62.31
July high temperature	89.50	9.00	86.64
annual precipitation (inches)	20.45	13.67	16.77
annual snowfall (inches)	9.31	59.38	4.49
pollution index	1.1510	0.303	1.2390
Urban/Density:			
% housing units in urban area	0.608	0.424	0.937
population density (persons/square mile)	665.13	1055.74	1521.12
Crime:			
California crime index	2966	6013	4358

Note: The unweighted values are averages over school districts, while the weighted values are averages over school districts weighting by the total number of households in the school district. The variables were assembled from various sources that are described in detail in the Data Appendix.

Table 3: Average Characteristics of the District's Schools

	Unweighted Mean	Standard deviation	Weighted Mean
Output Measure:			
average test score	273.44	28.25	266.14
School Characteristics			
average class size	26.14	3.99	28.10
average school size	472.65	241.65	717.77
fraction employees -- administrators	0.077	0.029	0.070
fraction employees -- pupil services	0.036	0.029	0.056
fraction expenditures -- support	0.034	0.020	0.037
fraction expenditures -- current	0.850	0.113	0.862
fraction expenditures -- instruction	0.549	0.084	0.562
district type -- K-8	0.557	0.497	0.231
district type -- K-12	0.352	0.478	0.702
Teacher Characteristics			
average experience (years)	14.22	3.31	15.04
average tenure (years)	10.46	3.02	12.35
fraction teachers -- masters degree + 60 hours	0.178	0.108	0.212
fraction teachers -- masters degree	0.124	0.787	0.160
fraction teachers -- bachelors degree + 60 hours	0.572	0.171	0.457
average salary	39,754	4,065	42,208

Note: The unweighted values are averages over school districts, while the weighted values are averages over school districts weighting by the total number of households in the school district. The sources for the school characteristics are the School District Data Book and RAND California. The source for the test score data is the California Department of Education. The construction of the variables is described in detail in the Data Appendix.

Table 5: Parameter Estimates

	Baseline	Race Black	Race Hispanic	Education HS degree or less	Income $40k or less
Community Characteristics:					
Commuting distance	0.033	0.014	0.012	0.021	0.040
	(0.010)	(0.009)	(0.007)	(0.008)	(0.012)
Commuting distance squared (/100)	-0.306	0.009	-0.114	-0.109	-0.301
	(0.140)	(0.150)	(0.060)	(0.050)	(0.050)
California crime index (/1,000)	-0.017	-0.006	0.017	-0.011	0.002
	(0.005)	(0.003)	(0.002)	(0.006)	(0.002)
Percent housing in urban area	2.453	0.503	0.086	0.496	-0.288
	(0.467)	(0.212)	(0.255)	(0.184)	(0.156)
Pollution index	-0.456	0.173	-0.012	0.585	0.499
	(0.232)	(0.172)	(0.185)	(0.160)	(0.125)
January high temperature	0.079	0.011	0.021	0.028	-0.006
	(0.042)	(0.023)	(0.021)	(0.016)	(0.007)
School Characteristics:					
Public school test score(/100)	3.469	-0.464	-0.556	-4.092	-1.802
	(0.482)	(0.350)	(0.406)	(0.718)	(0.567)
Private schooling	-4.943	-0.117	-0.871	-3.028	-1.668
	(1.481)	(0.753)	(0.668)	(1.002)	(0.611)
Peer characteristics					
Percent minority	-1.132	1.909	1.310	1.900	0.429
	(0.300)	(0.225)	(0.285)	(0.455)	(0.309)
Percent low parental education	-1.017	1.020	-0.264	1.270	-0.194
	(0.485)	(0.366)	(0.619)	(0.281)	(0.375)
Percent low income	-0.509	0.986	1.063	-0.360	0.774
	(0.333)	(0.298)	(0.313)	(0.301)	(0.169)

Note: Each cell in the table represents the parameter estimate associated with the interaction of the household characteristic specified in the column heading and the choice characteristics specified in the row heading. The baseline group is white households with at least some college degree and more than $40,000 in income. The other parameter estimates increment this baseline. Consequently, these parameter estimates must be added to the baseline estimate to calculate the appropriate coefficient if a household fits into any of these additional categories.

Note: Standard errors are shown in parentheses.

Table 5 (cont): Parameter Estimates

	Baseline	Race Black	Race Hispanic	Education HS degree or less	Income $40k or less
Tenure Characteristics:					
Rental Unit	0.020	0.052	0.054	0.047	0.106
	(0.011)	(0.014)	(0.013)	(0.015)	(0.024)
Housing Characteristics:					
Renter occupied					
ln (rooms)	0.423	-0.155	-0.192	0.212	0.151
	(0.168)	(0.132)	(0.139)	(0.111)	(0.202)
ln (age)	0.108	0.037	0.063	0.059	0.062
	(0.064)	(0.042)	(0.035)	(0.028)	(0.030)
ln (y-p)	1.000	0.326	0.299	0.073	0.053
		(0.088)	(0.120)	(0.076)	(0.046)
Owner occupied					
ln (rooms)	0.445	-0.150	-0.168	0.071	0.344
	(0.174)	(0.106)	(0.140)	(0.135)	(0.178)
ln (age)	0.041	0.016	-0.003	0.034	0.018
	(0.032)	(0.026)	(0.025)	(0.022)	(0.024)
ln (y-p)	1.223	0.205	0.193	0.065	0.040
	(0.344)	(0.091)	(0.116)	(0.070)	(0.057)
Inclusive Value Terms:					
$s^{rent}_{housing}$	0.54				
	(0.327)				
$s^{own}_{housing}$	0.32				
	(0.144)				
s_{tenure}	0.38				
	(0.134)				
s_{school}	0.65				
	(0.219)				

Note: Each cell in the table represents the parameter estimate associated with the interaction of the household characteristic specified in the column heading and the choice characteristics specified in the row heading. The baseline group is white households with at least some college degree and more than $40,000 in income. The other parameter estimates increment this baseline. Consequently, these parameter estimates must be added to the baseline estimate to calculate the appropriate coefficient if a household fits into any of these additional categories.

Note: Standard errors are shown in parentheses.

Table 6: Overall Consumption Patterms

	All	Black	Hispanic	Asian	White	HS Degree or less	Some college	Income < $40K	Income > $40K
Average test score	265.89	251.21	253.59	268.87	277.79	256.72	274.86	258.61	273.83
Private schooling	0.094	0.082	0.067	0.100	0.115	0.051	0.132	0.054	0.134
Crime index	4434	5459	5261	5226	3436	4786	4119	4641	4227
Pollution index	1.234	1.28	1.30	1.20	1.18	1.28	1.20	1.27	1.20
Home ownership	0.533	0.262	0.433	0.624	0.628	0.413	0.657	0.306	0.788
Owner occupied									
Housing value	225,015	166,136	172,934	257,045	248,831	183,800	251,435	155,541	255,235
Number of rooms	6.31	6.45	5.05	5.79	7.02	5.40	6.90	5.27	6.77
Renter occupied									
Monthly rent	560.27	596.77	514.90	568.75	598.28	514.68	639.83	515.40	724.79
Number of rooms	4.19	4.16	3.68	3.85	4.82	3.86	4.75	4.00	4.87

Note: The values in the table represent the average consumption of the choice characteristic shown in the row heading for households in the category shown in the column heading.

–

–

–

Table 7: Which Household Characteristics Explain the Observed School Quality Consumption Patterns?

	full set of choice characteristics	hhlds sort by preferences associated with[1]		
		race	parental education	household income
Race:				
Black households	252.6	258.3	261.1	261.8
Hispanic households	254.1	258.6	261.9	263.1
Asian and white households	274.3	271.0	268.8	267.7
Black-white/Asian gap	21.7	12.7	7.7	5.9
Hispanic-white/Asian gap	20.2	12.4	6.9	4.6
Parental education:				
Households with a HS degree or less	255.9	263.1	259.6	264.2
Households with at least some college	277.3	269.9	273.4	268.7
Parental education gap	21.4	6.8	13.8	4.5
Income:				
Households with income less than $40k	257.7	263.8	263.5	261.2
Households with income less than $40k	275.9	268.8	269.2	271.8
Income gap	18.2	5.0	5.7	10.6

[1] *Note:* Each column shows the predicted consumption of school quality that would result if households made their residential location decisions and school choices based only on the preferences associated with the household characteristic shown in the column heading. The first column depicts the consumption patterns predicted by the full model, while the next three columns show the patterns that would result if households sorted only on the basis of differences in preferences associated with race, parental education, and income, respectively. Importantly, in the final three columns of this table households' initial endowments (income and employment location) which enter the analysis through the construction of the all other goods (y-p) term and commuting distance are held constant.

Table 8: Which Community Characteristics Explain the Observed Consumption Patterns?

	predictions from full model	hhlds differ only in employment locations	households differ only in their preferences for		
			public school quality	housing characteristics	peer characteristics
Race:					
Black households	252.6	263.5	262.9	261.0	260.5
Hispanic households	254.1	264.1	262.1	263.0	260.3
Asian and white households	274.3	267.0	268.2	267.9	269.5
Black-white/Asian gap	21.7	3.5	5.3	6.9	9.0
Hispanic-white/Asian gap	20.2	2.9	6.1	4.9	9.2
Parental education:					
Households w/ a HS degree or less	255.9	264.5	258.8	264.2	263.1
Households w/ at least some college	277.3	267.1	273.8	267.4	268.8
Parental education gap	21.4	2.6	15.0	3.2	5.7
Income:					
Households with income < $40k	257.7	264.0	262.2	263.9	263.5
Households with income > $40k	275.9	267.8	270.2	268.1	268.6
Income gap	18.2	3.8	8.0	4.2	5.1

Note: Each column shows the predicted consumption of school quality that would result if households in the category shown in the row heading made their residential location decisions and school choices based only on the choice characteristics shown in the column heading. The first column depicts the consumption patternspredicted by the full model. The second column shows the predicted consumption patterns if households shared the same preferences but differed in their employment locations. The final three columns show the predicted consmption patterns that would result if households differed only in their initial endowments (income and employment location) and their preferences for school quality, housing characteristics, and peer socioeconomic composition, respectively.

–

–

–