CHOICE AND COMPETITION IN LOCAL EDUCATION MARKETS*

Patrick Bayer
Department of Economics
Yale University

Robert McMillan
Department of Economics
University of Toronto

July 2005

Preliminary

Abstract

Prompted by concerns about school quality, a growing empirical literature has measured the effects of greater choice on school performance. This paper contributes to the literature in three ways. First, it makes the conceptual point that the *overall* effect of greater choice, which has been the focus of the prior literature, can be decomposed into demand and supply components; knowing the relative sizes of the two has considerable relevance for policy. Second, using rich data from a large metropolitan area, it provides a direct and intuitive measure of the competition each school faces. This takes the form of a school-specific elasticity, measuring the extent to which reductions in school quality would lead to reductions in demand. Third, the paper provides evidence that these elasticity measures are strongly related to school performance: a one-standard deviation increase in the competitiveness of a school's local environment within the Bay Area leads to a 0.15 standard deviation increase in average test score. This positive correlation is robust, and is consistent with strong supply responsiveness on the part of public schools, of relevance to the broader school choice debate.

Key words: School Choice, School Competition, Education Demand, Student Achievement, School Performance.

JEL: I20, H41, R21

---

* We are grateful to Tim Bresnahan and Tom Nechyba for many valuable conversations about this research. Thanks also to Pat Bajari, Steve Berry, Dennis Epple, Fernando Ferreira, Caroline Hoxby, Brian Jacob, Jesse Rothstein, Kim Rueben, Holger Sieg, Doug Staiger, and Chris Timmins for helpful discussions. Melissa Knox (UC Berkeley) provided outstanding research assistance. This research was conducted at the California Census Research Data Center: our thanks to the CCRDC, and to Ritch Milby in particular. Financial support for this project provided by the NSF under grant SES-0137289, the Public Policy Institute of California, and SSHRC is gratefully acknowledged. Correspondence may be sent to: Patrick Bayer, 37 Hillhouse Avenue, Dept. of Economics, Yale University, New Haven, CT 06511, patrick.bayer@yale.edu, (203) 432-6292.

1 INTRODUCTION

Dissatisfaction with the quality of public education in the United States has prompted considerable interest in reforms that increase choice and stimulate competition in the school system. The standard notion, drawing on the theory of the firm, is that increasing the range of options households can choose from forces incumbent public schools to compete more vigorously for market share, better students, and financial resources; in turn (continuing the firm analogy), this should lead schools to better respond to the needs of students and their families, use available resources more efficiently, and exert greater effort, thereby improving public school performance. Because schools lack a strict profit motive, however, the extent to which public schools do actually respond to increased competition has become a central empirical question in the education literature.

In prior work, a large number of papers have examined the effects of greater choice on public school performance, typically using across-metropolitan area variation in indices of public school concentration or measures of private school availability.[1] In so doing, this literature has grappled with endogeneity problems related to the simultaneous determination of public school performance and these measures of choice, as well as omitted variable biases.[2] The evidence to emerge from this literature is mixed, ranging from little effect to a modest positive impact of increased choice on school performance.[3]

While the concepts of choice and competition are invariably linked in the broad policy and academic debates about education reform, they are not synonymous from an economic perspective.[4] Choice (in the education context) relates to the availability of schooling alternatives faced by households, while the economic notion of competition is best captured by the residual elasticity of demand that a school faces – that is, the extent to which a school's enrollment and resources are affected by a change in its performance. The greater is this elasticity, the more competitive is the environment in which it operates. In considering the effect of *competition* on

[1] See Belfield and Levin (2002) for an extensive review of forty one empirical studies in this literature.

[2] See Hoxby (1994, 2000), Figlio and Stone (1999), Jepsen (1999), Hanushek and Rivkin (2003), and Rothstein (2003).

[3] Researchers have also examined the efficiency of public school spending using alternative research designs. Barrow and Rouse (2004), for example, examine the efficiency of per pupil spending using variation in state aid while Millimet and Rangaprasad (2005) explore the spatial autocorrelation in school input decisions.

[4] The distinctness of these concepts can be made clear by noting that, in principle, choice could increase without any increase in competition - for example, private school vouchers could be given to parents wholeheartedly opposed to private schools. Likewise, the competitiveness of the education market place could change without any change in the set of available schooling options. For instance, broad changes in the demand for school quality or technological improvements in the quality of information available to households about school performance might increase residual demand elasticities even in the absence of an explicit change in choice.

school performance, one would ideally like to isolate the impact of a change in this primitive – a school's residual elasticity of demand – on school performance.

In practice, the measures of choice used in the literature tend to be related to competition for the reason that an increase in the availability of schooling options or in the ease of selecting these options is likely to increase the residual elasticity of demand faced by incumbent schools, thereby making the market more competitive. But the overall impact of increased choice on school performance depends on both demand and supply factors, combining (i) the effect of increased choice on the competitiveness of local market (i.e., the residual elasticity of demand faced by schools) and (ii) the effect of increased competition on school performance. The first of these reflects factors on the *demand side* of the education market, and is likely to vary depending on the way that choice increases.[5] The latter effect isolates the *supply-side* response to greater competition - the response of individual schools to variation in the competitiveness of the environment in which they operate.

By focusing on the overall effect of *choice* on school performance, the existing literature provides only indirect guidance as to the magnitude of the responsiveness of schools to *competition* (supply responsiveness). Yet the distinction between demand vs. supply responsiveness is important both for understanding the economics of education markets and for policy. To the extent that an overall modest positive effect of choice on performance is driven by a sizeable impact of choice on competition and a limited responsiveness of schools to increased competition, this suggests that policies targeted at improving incentives for schools to respond to competition are likely to have strong performance effects. On the other hand, if schools are highly responsive to changes in competition (as we will find) but the current system does not engender a very competitive environment, policies aimed at increasing demand elasticities would seem more promising. Such policies might, for example, increase the ability of parents to choose from a wider set of schools or provide better information about quality differences among the available options.

In terms of magnitudes, a reading of the existing education demand literature suggests that the demand side component of the overall effect of choice on school performance may in fact be quite small. This conclusion is robust across a wide variety of different research designs: (i) examining differences in housing prices along school assignment boundaries or as school assignments are changed over time;[6] (ii) estimating the heterogeneity in willingness to pay for

[5] For example, choice variation from geographic accessibility to private schools may have a quite different impact on competition than policies such as within-district school choice programs

[6] Black (1999) pioneered the use of school assignment boundary fixed effects, estimating that a school-level standard deviation in average test score is associated with approximately a 2 percent increase in house

school quality using a model of residential sorting;[7] (iii) examining the academic performance of winners versus losers following randomized school choice lotteries;[8] and (iv) examining how student performance varies with measures of school district enrollment concentration across metropolitan areas.[9] This leaves open the possibility that the responsiveness of schools to increases in competition may be sizeable despite the mixed results of the previous literature exploring the impact of choice on performance.

In this paper, we set out evidence based on a new approach for studying the effects of competition on school performance – that is, measuring supply responsiveness. At the heart of the analysis, we develop a school-specific measure that captures the extent to which reductions in school quality would lead to reductions in demand. This measure has intuitive appeal: a school is taken to operate in a competitive environment if slight reductions in school quality would lead to a substantial reduction in demand for the neighborhoods from which the school draws its students.

Our elasticity estimates are derived from a rich demand model describing household preferences for house and neighborhood attributes, including school quality. With these preference estimates in hand, we estimate the elasticity of demand faced by each of over 700 elementary schools in the San Francisco Bay Area. We then explore the relationship between these elasticity measures and public school performance, as measured by standardized tests, controlling for a wide set of student, school, household and neighborhood attributes (including all the variables that are included in the demand model). If schools simply maximize quality, measured by test scores given their available resources, then the elasticity would be irrelevant to a school's quality-setting decision. In contrast, if (at the other extreme) schools were rent-seeking or prone to make mistakes in allocating resources, then the elasticity would play a key role, as in the textbook theory of the firm.

Our results from this empirical exercise are striking and robust. We find strong evidence that higher demand elasticities are associated with increases in public school achievement scores, with little effect on resource use, indicating that productivity improves. The same findings persist regardless of which student, school, and neighborhood characteristics we condition on, and the

---

value. Kane, Staiger, and Riegg (2004) report almost an identical result using boundary fixed effects. This latter paper also reports results based on changes in school assignment, although these are less conclusive.

[7] Bayer, Ferreira, and McMillan (2003) incorporate school district boundary fixed effects in estimating a model of residential sorting, returning an estimated mean marginal willingness to pay for a standard deviation increase in average test score of approximately 2 percent of house value; they also find evidence of heterogeneity around this mean, as well as strong preferences relating to the characteristics of neighbors.

[8] Cullen, Jacob, and Levitt (2003) find little evidence that winners of randomized lotteries perform better in the schools that they subsequently select than losers who do not have the same degree of choice.

[9] Rothstein (2003) finds little evidence that sorting is related to school rather than peer characteristics.

magnitude of the estimated effect is quite large: a one standard deviation increase in the competitiveness of a school's local market within the Bay Area is associated with a 0.15 standard deviation increase in school performance. That these achievement increases are accompanied by, if anything, slight reductions in teacher quality measures also helps to allay concerns about the role of omitted variables in these regressions. In terms of heterogeneous effects, while we find that all types of schools respond to increased competition, school responsiveness is greater in more educated communities.

Two aspects of the initial research design deserve additional attention. First, the demand elasticities that we use are estimated rather than observed directly in the data. Second, the school performance regressions are estimated using OLS, thus basing the key parameter of interest on any variation in the demand elasticity not absorbed out by the controls for observable control variables. To provide additional confidence in our baseline results, we explore the variation in the competitiveness of a school's environment, relating it to measures created directly from the data that characterize the availability of close substitutes in terms of neighborhood and school characteristics in the local market. In particular, we construct dissimilarity indices that describe the extent to which the schools/neighborhoods are isolated in quality space relative to their closest neighbors. We demonstrate that our elasticity measure is strongly correlated with these dissimilarity measures and that these dissimilarity measures are correlated with school performance in a manner consistent with our baseline results.

These results have implications for the broader choice debate. That we find such sizeable effects in the San Francisco Bay Area is somewhat surprising given California's public finance system, which limits the effect of local property values on school resources. This suggests that supply responsiveness may be even larger in other states. Moreover, it suggests that mechanisms at work more generally, such as monitoring which might respond to changes in local property values, may provide enough of an incentive for schools even in the absence of the direct tying of resources to property values. In light of our results, policies that increase residual elasticities of demand would seem promising, either through reforms that improve choice or through improving information.

The rest of the paper is organized as follows: the next section motivates our empirical approach. Section 3 sets out the demand model and Section 4 describes its estimation and addresses important identification issues. Section 5 describes the rich data used in the analysis, and Section 6 discusses our demand estimates and the way these are used to construct a residual elasticity of demand for each school, based on meaningful variation in the data. Section 7 discusses the supply-side regressions that yield our main results, and Section 8 concludes.

2 SCHOOL CHOICE AND COMPETITION

In this section, we motivate more fully the use of the residual elasticity of demand as our preferred competition measure. We then describe how we use this to shed light on the direct effect of competition on public school performance.

The residual elasticity of demand measures the change in demand in response to a change in school quality. To see why this provides natural measure of competition, consider a stylized model of a local education market. The agents on the demand side comprise households with children, with teachers and school administrators on the supply side. Household preferences are defined over consumption, housing services and school quality, and households choose where to live and where to send their children to school (there may be private school alternatives), based on quality and cost. Suppose school quality is unidimensional, measured by standardized tests. It is the output of a public school education production technology that converts student characteristics, school resources, teacher quality, and discretionary teacher and administrator effort into a performance measure.[10]

Public school objectives are potentially complex, not least because public schools tend to be heavily regulated. For simplicity, we will treat the school as a single effort-making body, and consider two polar cases. On the one hand, public schools could aim to maximize school quality, given resources; on the other, public schools could aim to maximize rents, which are increasing in school revenues (and thus enrollment) and decreasing in effort cost.

If schools were quality maximizers, then quality would be set independent of considerations about the effect that it would have on enrollment. School personnel would simply set effort at the maximal level consistent with their continued participation in the schooling sector. Significantly, under quality maximization, the school production function would be seen directly in the data, and the elasticity of demand would have no effect on school performance.

In contrast, under rent seeking, a school would face the following trade-off: by raising quality (through higher effort), it would increase enrollment, and if funding were on a per-pupil basis, this would lead to an increase in revenues. At the same time, higher quality would require higher effort, which is costly. How the school would resolve this trade-off in making its optimal effort choice would depend in part on the size of the marginal benefit of higher effort. This, in turn, would depend on the response of demand (measured by school enrollment) to higher quality – the residual elasticity of demand. Clearly, in a rent-seeking world, the elasticity of demand

[10] In practice, school personnel have other choice variables than effort. They may also reallocate resources in ways that parents prefer.

would affect school production via discretionary effort choice: as the elasticity rose, so the school would have greater incentive to raise quality so as to avoid a significant loss of enrollment.[11]

This discussion motivates our main estimating equation, which relates measured school performance to a set of relevant determinants: the characteristics of students, school inputs, teacher characteristics, and neighborhood controls, including our measure of local competition, the residual elasticity of demand. Teacher and administrator effort is not observed by the researcher, but (depending on school objectives) might be influenced by market conditions. We adopt the following linear production function specification:

$$(1) \qquad T^m = X_c^{m}{}'\gamma_c + X_s^{m}{}'\gamma_s + X_n^{m}{}'\gamma_n + \gamma_E E^m + \varepsilon^m$$

where T^m is school m's average test score, X_c^m is a vector of the characteristics of the children that attend the school m, X_s^m is a vector of school and teacher characteristics, X_n^m represents a set of neighborhood controls, E^m is the elasticity of demand, γ_c, γ_s, γ_n and γ_E are coefficients to be estimated, and ε_m is a random error term. This is the equation we will take to the data, with interest focusing on the γ_E coefficient.

Beyond the stylized model, there are a variety of reasons why public schools might be responsive to the residual elasticity of demand, consistent with a positive estimate for γ_E. A high elasticity might make homeowners more sensitive to school behavior, leading to better monitoring and more active political involvement in elections for school board and other local officials. In some school finance regimes, property values determine local property taxes and thus influence school revenue; thus schools would have an incentive in such settings to care about the effect of school quality on local property prices. Conversely, there are clear reasons why schools might not be responsive: teaching and administrative positions afford a good deal of job security; and political pressure might be ineffective if elections were for positions schooling is only one of many issues.

It is an empirical question just how strong these incentives to respond to competition are. To shed light on this issue, we estimate the key equation (1) above, using a rich data set from the San Francisco Bay Area. For this purpose, we construct a set of school-specific elasticities in

[11] In practice, it is unlikely that schools are pure rent-seekers. However, to the extent that they have some discretion over quality setting, and depart from pure quality maximization, so the elasticity of demand will, to some degree, influence school quality setting.

which demand is measured in terms of local housing values,[12] and quality is measured using school average test scores, conditioning on school and student characteristics. The corresponding residual elasticity for a given school captures the change in local house prices as school quality changes.

The residual elasticities are not directly observed in the data. Rather, we estimate them in the following way: first, rich Census data are used to estimate a flexible demand system, taking careful account of endogeneity issues on this side of market and making explicit the way that individual demands aggregate up to form a housing market equilibrium using an equilibrium model of the housing market. We then use these demand estimates in combination with the equilibrium model to perform a series of simple counterfactual experiments. In particular, for each of 708 elementary schools in the Bay Area, we use the equilibrium model to conduct a simple counterfactual simulation, raising its average test score by 0.1 standard deviation (7.762 test score points) and calculating the new housing market equilibrium. This has the effect of increasing house values in the corresponding neighborhood catchement area; the resulting predicted change in house values given the change in school quality is then used in the estimating various specifications of the regression shown in equation (1).

Unlike prior work, it is important to note that our approach provides estimates of school-specific elasticities, rather than MSA-level average choice indices. Constructing our elasticities from a single financing regime, rather than looking across MSAs, has the advantage that one would expect the incentives to respond to competition to vary with financing: our approach will better allow us to identify the direct effect of competition. We note that under California's local public finance system, the marginal dollar comes from the state rather than from local property taxes. This might be expected to provide weaker incentives to respond to competition, as schools are less able to take advantage of quality improvements directly. In turn, it is likely to provide lower bound on incentives to respond to competition, worth remembering when it comes to interpreting the economic significance of our results.

3 DEMAND

This section sets out our demand model in some detail. The model consists of two key elements: the household residential location decision problem and a market-clearing condition. While it has a simple structure, the model allows households to have heterogeneous preferences defined over housing and neighborhood attributes in a very flexible way; it also allows for

---

[12] Alternatively, demand could be measured based on student enrollment (or even enrollment of specific types of student).

housing prices and neighborhood sociodemographic compositions to be determined in equilibrium.

The Residential Location Decision. We model the residential location decision of each household as a discrete choice of a single residence from a set of house types available in the market. The utility function specification is based on the random utility model developed in McFadden (1973, 1978) and the specification of Berry, Levinsohn, and Pakes (1995), which includes choice-specific unobservable characteristics. Let X_h represent the observable characteristics of housing choice h, including characteristics of the house itself (e.g., size, age, and type), its tenure status (rented vs. owned), and the characteristics of its neighborhood (e.g., school, crime, land use, and topography). We use the notation capital letter Z_h to represent the average sociodemographic characteristics of the corresponding neighborhood, writing it separately from the other housing and neighborhood attributes to make explicit the fact that these characteristics are determined in equilibrium. Let p_h denote the price of housing choice h and, finally, let d_h^i denote the distance from residence h to the primary work location of household i. Each household chooses its residence h to maximize its indirect utility function V_h^i:

$$(2) \qquad \underset{(h)}{Max} \quad V_h^i = \alpha_X^i X_h + \alpha_Z^i Z_h - \alpha_p^i p_h - \alpha_d^i d_h^i + \xi_h + \varepsilon_h^i .$$

The error structure of the indirect utility is divided into a correlated component associated with each housing choice that is valued the same by all households, ξ_h, and an individual-specific term, ε^i_h. A useful interpretation of ξ_h is that it captures the unobserved quality of each housing choice, including any unobserved quality associated with its neighborhood.

Each household's valuation of choice characteristics is allowed to vary with its own characteristics, z^i, including education, income, race, employment status, and household composition. Specifically, each parameter associated with housing and neighborhood characteristics and price, α^i_j, for $j \in \{X, Z, d, p\}$, varies with a household's own characteristics according to:

$$(3) \qquad \alpha_j^i = \alpha_{0j} + \sum_{k=1}^{K} \alpha_{rj} z_k^i ,$$

with equation (3) describing household i's preference for choice characteristic j.

This specification of the utility function gives rise to a horizontal model of sorting in which households have preferences defined distinctly over each choice characteristic including both school quality and neighborhood sociodemographic characteristics.[13] This contrasts with vertical models, which restrict households to have preferences over a single locational index, thereby constraining households to have the same preference ordering across locations.[14] The additional flexibility of the horizontal model is especially relevant for this paper as it is the magnitude of the heterogeneity in preferences for neighborhood sociodemographic characteristics that will determine the extent to which the initial stratification induced by the variation in school quality across the metropolitan area is reinforced by additional sorting due to preferences for one's neighbors.

Characterizing the Housing Market. As with all models in this literature, the existence of a sorting equilibrium is much easier to establish if the individual residential location decision problem is smoothed in some way. To this end, we assume that the housing market can be fully characterized by a set of housing types that is a subset of the full set of available houses, letting the supply of housing of type h be given by S_h. We also assume that each household observed in the sample represents a continuum of households with the same observable characteristics, with the distribution of idiosyncratic tastes ε^i_h mapping into a set of choice probabilities that characterize the distribution of housing choices that would result for the continuum of households with a given set of observed characteristics.[15]

Given the household's problem described in equations (2)-(3), household i chooses housing type h if the utility that it receives from this choice exceeds the utility that it receives from all other possible house choices - that is, when

---

[13] The horizontal specification also captures the geography of the urban housing market very naturally, allowing households to have preferences over neighborhoods depending on the distance from their employment locations. This gives rise to variation in the aggregate demand for housing in various neighborhoods throughout the metro area, thereby increasing equilibrium housing prices in neighborhoods near employment centers.

[14] It is important to point out that this flexible feature of our model is made possible because we abstract from issues related to local politics. As Epple, Filimon, and Romer (1993) note, incorporating local politics into models of residential sorting requires restrictions to be placed on preferences in order to guarantee the existence of an equilibrium. Important recent papers by Epple and Sieg (1999) and Epple, Romer and Sieg (2001) estimate equilibrium models that include voting over the level of public goods, restricting households to have shared rankings over a single public goods index. We view our model as having a comparative rather than absolute advantage over the papers in this line of the literature, better suited for an institutional setting such as that which holds in Californian, where Proposition 13 leaves almost no discretion over property tax rates or the level of public goods spending at the local level.

[15] For expositional ease and without loss of generality, let the measure of this continuum be one.

$$(4) \qquad V_h^i > V_k^i \quad \Rightarrow \quad W_h^i + \varepsilon_h^i > W_k^i + \varepsilon_k^i \quad \Rightarrow \quad \varepsilon_h^i - \varepsilon_k^i > W_k^i - W_h^i \qquad \forall \quad k \neq h$$

where W^i_h includes all of the non-idiosyncratic components of the utility function V^i_h. As the inequalities in (4) imply, the probability that a household chooses any particular choice depends in general on the characteristics of the full set of possible house types. Thus the probability P^i_h that household i chooses housing type h can be written as a function of the full vectors of housing and neighborhood characteristics (both observed and unobserved) and prices {**X**, **Z**, **p**, ξ}:[16]

$$(5) \qquad P_h^i = f_h(z^i, \mathbf{Z}, \mathbf{X}, \mathbf{p}, \xi)$$

as well as the household's own characteristics z^i.

Aggregating the probabilities in equation (5) over all observed households yields the predicted demand for each housing type h, D_h:

$$(6) \qquad D_h = \sum_i P_h^i \ .$$

In order for the housing market to clear, the demand for houses of type h must equal the supply of such houses and so:

$$(7) \qquad D_h = S_h, \quad \forall h \quad \Rightarrow \quad \sum_i P_h^i = S_h \quad \forall h \, .$$

Given the decentralized nature of the housing market, prices are assumed to adjust in order to clear the market. The implications of the market clearing condition defined in equation (6) for prices are very standard, with excess demand for a housing type causing price to be bid up and excess supply leading to a fall in price. Given the indirect utility function defined in (1) and a fixed set of housing and neighborhood attributes, Bayer, McMillan, and Rueben (2004a) show that a unique set of prices (up to a scale) clears the market.

When some neighborhood attributes are endogenously determined by the sorting process itself, we define a sorting equilibrium as a set of residential location decisions and a vector of housing prices such that the housing market clears and each household makes its optimal location

[16] For the purposes of characterizing the equilibrium properties of the model, we include an individual's employment location in z^i and the residential location in X_h.

decision given the location decisions of all other households. In equilibrium, the vector of neighborhood sociodemographic characteristics along with the corresponding vector of market clearing prices must give rise to choice probabilities in equation (4) that aggregate back up to the same vector of neighborhood sociodemographics.[17] Whether this model gives rise to multiple equilibria depends on the distributions of preferences and available housing choices as well as the utility parameters. In general, it is impossible to establish that the equilibrium is unique *a priori*. However, estimation of the model does not require the computation of an equilibrium nor uniqueness more generally, as we describe in the next section. We discuss the issue of uniqueness further in the context of the counterfactual simulations in Section 6 below.

4 DEMAND ESTIMATION

Estimation of the model follows a two-stage procedure closely related to that developed in Berry, Levinsohn, and Pakes (1995). This section outlines the estimation procedure; a rigorous presentation is contained in Bayer, Ferreira, and McMillan (2003) and is included in a technical appendix. It is helpful in describing the estimation procedure to first introduce some notation. In particular, rewrite the indirect utility function as:

$$(8) \qquad V_h^i = \delta_h + \lambda_h^i + \varepsilon_h^i$$

where

$$(9) \qquad \delta_h = \alpha_{0X} X_h + \alpha_{0Z} Z_h - \alpha_{0p} p_h + \xi_h$$

and

$$(10) \qquad \lambda_h^i = \left(\sum_{k=1}^{K} \alpha_{kX} z_k^i \right) X_h + \left(\sum_{k=1}^{K} \alpha_{kZ} z_k^i \right) Z_h - \left(\sum_{k=1}^{K} \alpha_{kp} z_k^i \right) p_h - \left(\sum_{k=1}^{K} \alpha_{kd} z_k^i \right) d_h .$$

In equation (9), δ_h captures the portion of utility provided by housing type *h* that is common to all households, and in (10), *k* indexes household characteristics. When the household characteristics included in the model are constructed to have mean zero, δ_h is the mean indirect utility provided by housing choice *h*. The unobservable component of δ_h, ξ_h, captures the portion of unobserved preferences for housing choice *h* that is correlated across households, while ε_h^i represents unobserved preferences over and above this shared component.

[17] Bayer, McMillan, and Rueben (2004a) establish the existence of a sorting equilibrium as long as (i) the indirect utility function shown in equation (2) is decreasing in housing prices for all households; (ii)

The first step of the estimation procedure is equivalent to a Maximum Likelihood estimator applied to the individual location decisions taking prices and neighborhood sociodemographic compositions as given,[18] returning estimates of the heterogeneous parameters in λ and mean indirect utilities, δ_h. This estimator is based simply on maximizing the probability that the model correctly matches each household observed in the sample with its chosen house type. In particular, for any combination of the heterogeneous parameters in λ and mean indirect utilities, δ_h, the model predicts the probability that each household i chooses house type h. We assume that ε_h^i is drawn from the extreme value distribution, in which case this probability can be written:

(11) $$P_h^i = \mathbf{exp}(\delta_h + \hat{\lambda}_h^i) \Big/ \sum_k \mathbf{exp}(\delta_k + \hat{\lambda}_k^i)$$

Maximizing the probability that each household makes its correct housing choice gives rise to the following quasi-log-likelihood function:

(12) $$\tilde{\ell} = \sum_i \sum_h I_h^i \, \mathbf{ln}(P_h^i)$$

where I^i_h is an indicator variable that equals 1 if household i chooses house type h in the data and 0 otherwise. The first stage of the estimation procedure consists of searching over the parameters in λ and the vector of mean indirect utilities to maximize $\tilde{\ell}$. Notice that the quasi-likelihood function developed here is based solely on the notion that each household's residential location is optimal given the set of observed prices and the location decisions of other households.

The Endogeneity of School Quality and Neighborhood Sociodemographic Composition. Having estimated the vector of mean indirect utilities in the first stage of the estimation, the second stage of the estimation involves decomposing δ into observable and unobservable

indirect utility is a continuous function of neighborhood sociodemographic characteristics; and (iii) ε is drawn from a continuous density function.

[18] Formally, the validity of this first stage procedure requires the assumption that the observed location decisions are individually optimal, given the collective choices made by other households and the vector of market-clearing prices *and* that households are sufficiently small such that they do not interact strategically with respect to particular draws on ε. This ensures that no household's particular idiosyncratic preferences affect the equilibrium and the vector of idiosyncratic preferences ε is uncorrelated with the prices and neighborhood sociodemographic characteristics that arise in any equilibrium. For more discussion, see the Technical Appendix.

components according to the regression equation (9).[19] In estimating equation (9), important endogeneity problems need to be confronted. To the extent that house prices partly capture house and neighborhood quality unobserved to the econometrician, so the price variable will be endogenous. Estimation via least squares will thus lead to price coefficients biased towards zero, producing misleading willingness-to-pay estimates for a whole range of choice characteristics. This issue arises in the context of any differentiated products demand estimation and we describe the construction of an instrument for price in the Technical Appendix.

A second identification issue concerns the correlation of neighborhood sociodemographic characteristics $\boldsymbol{Z}$ and school quality with unobserved housing and neighborhood quality, ξ_h. To properly estimate preferences in the face of this endogeneity problem, we adapt a technique previously developed by Black (1999). Black's strategy makes use of a sample of houses near school attendance zone boundaries, estimating a hedonic price regression that includes boundary fixed effects. Intuitively, the idea is to compare houses in the same local neighborhood but on opposite sides of the boundary, exploiting the discontinuity in the right to attend a given school.

There are, however, good reasons to think that households will sort with respect to such boundaries. Thus, while boundary fixed effects are likely to do a good job of controlling for differences in unobserved fixed factors, neighborhood sociodemographics are likely to vary discontinuously at the boundary. In this way, the use of boundary fixed effects isolates variation in both school quality and neighborhood sociodemographics in a small region in which unobserved fixed features, (e.g., access to the transportation network) are likely to vary only slightly, thereby providing an appealing way to account for the correlation of both school quality *and* neighborhood sociodemographics with unobservable neighborhood quality.

We incorporate school district boundary fixed effects when estimating equation (11). In particular, we create a series of indicator variables for each Census block that equal one if the block is within a given distance of each unique school district boundary in the metropolitan area (e.g., Palo Alto-Menlo Park).[20] We show the variation in school quality and neighborhood sociodemographics at school district boundaries in the next section after first describing the basic features of the dataset.

---

[19] Notice that the set of observed residential choices provides no information that distinguishes the components of δ. That is, however δ is broken into components, the effect on the probabilities shown in equation (10) is identical.

[20] A number of empirical issues arise in incorporating school district boundary fixed effects into our analysis. A central feature of local governance in California helps to eliminate some of the problems that naturally arise with the use of school district boundaries, as Proposition 13 ensures that the vast majority of school districts within California are subject to a uniform effective property tax rate of one percent.

5 DATA

The analysis conducted in this paper is facilitated by access to restricted Census microdata for 1990. These restricted Census data provide the detailed individual, household, and housing variables found in the public-use version of the Census, but also include information on the location of individual residences and workplaces at a very disaggregate level. In particular, while the public-use data specify the PUMA (a Census region with approximately 100,000 individuals) in which a household lives, the restricted data specify the Census block (a Census region with approximately 100 individuals), thereby identifying the local neighborhood that each individual inhabits and the characteristics of each neighborhood far more accurately than has been previously possible with such a large-scale data set.

We use data from six contiguous counties in the San Francisco Bay Area: Alameda, Contra Costa, Marin, San Mateo, San Francisco, and Santa Clara. We focus on this area for a number of reasons. First, this geographic area is reasonably self-contained, and is sizeable along a number of dimensions, including over 1,100 Census tracts, and almost 39,500 Census blocks, the smallest unit of aggregation in the data. The sample consists of 242,100 households. More importantly, the use of data from California makes it reasonable to use school district rather than school attendance zone boundaries in the analysis. In particular, Proposition 13 ensures that local jurisdictions have almost no discretion over property tax rates or the level of public good spending including school spending. In this way, unlike almost anywhere else in the country, one would not expect much variation in property values across school district boundaries to arise due to differential property tax rates in California. This same feature the public finance system may also diminish the overall strength of Tiebout-type sorting in California, as households are not free to select different tax rates and local public goods packages in each jurisdiction. For this reason we expect our analysis to generally provide a lower bound on the importance of school-related sorting relative to other states.

The Census provides a wealth of data on the individuals in the sample – race, age, educational attainment, income from various sources, household size and structure, occupation, and employment location. In addition, it provides a variety of housing characteristics: whether the unit is owned or rented, the corresponding rent or owner-reported value,[21] number of rooms,

---

Concerning the width of the boundaries, we experimented with a variety of distances and report the results for 0.25 miles, as these were more precise due to the larger sample size.

[21] As described in the Data Appendix, we construct a single price vector for all houses, whether rented or owned. Because the implied relationship between house values and current rents depends on expectations about the growth rate of future rents in the market, we estimate a series of hedonic price regressions for

number of bedrooms, type of structure, and the age of the building. We use these housing characteristics directly, and also construct neighborhood variables, such as neighborhood racial, education and income distributions, based on the households within the same Census block group (a Census region containing approximately 500 housing units). We merge additional data describing local conditions with each house record, constructing variables related to crime rates, land use, local schools, topography, and urban density. For each of these measures, a detailed description of the process by which the original data were assigned to each house is provided in a Data Appendix. The list of the principal housing and neighborhood variables used in the analysis, along with means and standard deviations, is given in the first two columns of Table 1.

School Assignment and School Quality. While we have an exact assignment of Census blocks to school districts, in the absence of comprehensive information about within-district school attendance areas, we employ the following approach for linking each Census block to a school: For a given Census block, we calculate the distance to each school in its district. We then first assign the Census block to the closest school within its district. Using this closest school assignment, we can then calculate a predicted enrollment for each school (calculated by summing over the school-aged children in each Census block assigned to a school) and compare this measure to the actual enrollment of the school. To correct discrepancies in predicted versus actual enrollment, we then use an intuitive procedure to adjust the assignment of Census blocks to schools so as to ensure that predicted enrollments equal their actual counterparts in each school in each district. We describe this procedure in detail in the Data Appendix; the results are not sensitive to this adjustment.

As our measure of school quality, we use the average test score for each school, averaged over two years. Averaging helps to reduce any year-to-year noise in the measure. When variables that characterize the sociodemographic composition of the school or surrounding neighborhood are included in the analysis, the estimated coefficient on average test score picks up what households are willing to pay for an improvement in average student performance at a school holding the sociodemographic composition constant. While the average test score is an imperfect measure of school quality, it has the advantage of being easily observed by both parents and researchers and consequently has been used in most analyses that attempt to measure demand for school quality.

---

each of over 40 sub-regions of the Bay Area housing market. These regressions return an estimate of the ratio of house values to rents for each of these sub-regions and we use the average of these ratios for the

Boundaries. Table 1 displays descriptive statistics for various samples related to the boundaries. The first two columns report means and standard deviations for the full sample while the third column reports means for the sample of houses within 0.25 miles of a school district boundary.[22] Comparing the first column to the third column of the table, it is immediately obvious that the houses near school district boundaries are not fully representative of those in the Bay Area as a whole. To address this problem, we create sample weights for the houses near the boundary.[23] Column 7 of Table 1 shows the resulting weighted means, indicating that using these weights makes the sample near the boundary much more representative of the full sample.

The fourth and fifth columns report means for houses within 0.25 miles of a boundary, comparing houses on the high versus low average test score side of the each boundary; the sixth column reports t-tests for the difference in means. Comparing these differences reveals that houses on the high side cost $53 more per month and are assigned to schools with test scores that are 43-point higher on average.[24] Moreover, houses on the high quality side of the boundary are much more likely to be inhabited by white households and households with more education and income. These types of across-boundary differences in sociodemographic composition are what one would expect if households sort on the basis of preferences for school quality. While far less significant, other housing characteristics do vary across the boundaries as well. Consequently, we expect the use of boundary fixed effects to control for much of the variation in unobserved housing and neighborhood quality, thereby giving rise to more accurate estimates of preferences for neighborhood sociodemographics and school quality.[25]

Bay Area, 264.1, to convert monthly rent to house value for the purposes of reporting results at the mean.

[22] We experimented with a variety of distances and report the results for 0.25 miles, as these were more precise due to the larger sample size.

[23] The following procedure is used: we first regress a dummy variable indicating whether a house is in a boundary region on the vector of housing and neighborhood attributes using a logistic regression. Fitted values from this regression provide an estimate of the likelihood that a house is in the boundary region given its attributes. We use the inverse of this fitted value as a sample weight in subsequent regression analysis conducted on the sample of houses near the boundary.

[24] As described in the Data Appendix, we construct a single monthly price vector for all houses, whether rented or owned.

[25] In terms of the estimates related to neighborhood sociodemographic characteristics, the key point about using school district boundary fixed effects rather than Census tract fixed effects is that in the boundary case we have a clear sense of what fundamentally leads to the sorting of households across neighborhoods within the region upon which the fixed effect is based. Because we control directly for that cause of the sorting - schooling in this case - we are less concerned that the variation in sorting is related to variation in unobservables within the region upon which the fixed effect is based.

6 DEMAND ESTIMATES

We noted in Section 3 that estimation of the full model proceeds in two stages. The first stage recovers interaction parameters and a vector of mean indirect utilities; the second stage returns the components of mean indirect utility. To give the reader a sense of the interplay between the estimates of the demand for school quality and sociodemographic characteristics of neighbors, we report results for two main specifications, which respectively exclude and include variables that characterize the racial composition, average educational attainment, and average income of the neighborhood (Census block group). To make the discussion of these estimates more transparent, we transform the estimates so that they can be described in terms of marginal willingness-to-pay measures (MWTP), reporting these estimates in Tables 2-4.

Table 2 reports the implied measures of the mean MWTP for school average test scores and other selected housing and neighborhood attributes for six specifications of the mean indirect utility regression.[26] Results are reported for the two main specifications, which exclude and include neighborhood sociodemographic variables, respectively. For each of these main specifications, results are reported for the full sample and for a sample of houses within 0.25 miles of school district boundaries, with and without including boundary fixed effects. In all cases, when the sample of houses is restricted to those within 0.25 miles of a boundary, sample weights (as described in Section 4) are used in order to make this sample as close to representative of the full sample as possible. Comparing columns 1 and 2 or columns 4 and 5, it is clear that results are very similar whether the full sample or the weighted sub-sample of houses near a school district boundary is used.

The first three columns of Table 2 report results for specifications that exclude neighborhood sociodemographics. In this case, the estimated mean MWTP for a one standard deviation increase in school average test score declines from \$123 to \$82 in monthly rent (\$21,500 in house value) when boundary fixed effects are included in the analysis. The final three columns report results for analogous specifications that include neighborhood sociodemographic characteristics. The coefficient on the average test score in these specifications returns the average of what households are willing to pay for a standard deviation increase in the average test score *conditional* on the sociodemographic characteristics of the neighborhood, which are in most cases also indicative of the sociodemographic characteristics of the local school. The estimated MWTP for a one standard deviation increase in the average test

[26] The specifications of the mean indirect utility regressions are reported in Appendix Table 1. The mean MWTP measures reported in Table 2 are calculated by dividing the coefficient associated with each choice characteristic by the coefficient on price.

score in this case declines to $26 in monthly rent ($6,900 in house value), which is approximately 2.4 percent of the average house price index for our Bay Area sample.[27]

These results make clear that much of what initially appears in the specification without neighborhood sociodemographics to be a significant willingness on the part of households to pay for school quality is instead more properly attributed to the characteristics of neighbors or peers. That the resulting MWTP for school average test scores is relatively small is exactly what one would expect if households have difficulty inferring the quality of a school from published average test score data.[28] That is, one would expect households to place a relatively small weight on this measure when choosing neighborhoods if the signal that the published average test score provided about actual school quality were small relative to the noise that it contains related to differences in the underlying composition of individuals taking the test. In fact, some of the weight that parents place directly on neighborhood sociodemographics may result from a belief that these provide a better indication than does the test score itself of the quality of the education that their children will receive in the local schools, especially if parents perceive peer effects to be important.

Before turning to the results related to heterogeneity in preferences, it is important to point out that the final two columns of Table 2 also show the impact of including boundary fixed effects on the estimates of mean preferences for neighborhood sociodemographic characteristics. Comparing these columns reveals the pattern of results that one would expect if boundary fixed effects control in part for unobserved neighborhood quality and unobserved quality is positively correlated with neighborhood income and education and negatively correlated with the fraction of non-white households.[29] Thus boundary fixed effects seem to be effective in controlling for fixed aspects of unobserved neighborhood quality that are correlated with neighborhood sociodemographics, and thus provide an attractive way of estimating preferences for both school quality *and* neighborhood sociodemographic characteristics in the presence of this important endogeneity problem.

Heterogeneity in Willingness-to-Pay

---

[27] This is very similar to the estimates of Black (1999) and Kane, Staiger, and Riegg (2004).

[28] This is especially true in 1990, which pre-dates most concerted efforts on the part of states to provide information to households about the quality of the local school. In many cases, such information simply reflects that average test score anyway.

[29] The fact that the estimated coefficient on the average test score rises from $20 to $26 when boundary fixed effects are included reflects that fact it is positively correlated with neighborhood income and education and negatively correlated with the fraction of non-white households. Thus, the estimated coefficient on the average test score tends to rise as the coefficients on these other variables change, as they do in moving from column 5 to column 6 in Table 2.

The interaction parameters estimated in the first stage for the specifications that exclude and include neighborhood sociodemographics, respectively, are reported in Appendix Tables 2 and 3. These specifications correspond to the mean MWTP estimates reported in columns 3 and 6 of Table 2, respectively. In each case, the model controls simultaneously for the effect of each of a series of household characteristics (income, education, race, work status, age, and household structure) on the marginal willingness-to-pay for a variety of housing and neighborhood attributes, including the average test score of the local school. The model also captures the spatial aspect of the housing market by allowing households to have preferences over commuting distance.[30]

Table 3 converts the estimates of the heterogeneity in MWTP for the average test score into dollar terms. The two columns of the table report results for specifications that exclude and include neighborhood sociodemographic characteristics, respectively. The first row of this table reports the estimated mean MWTP for the change listed in each column heading: for example, the first entry reports the $82 mean MWTP for a one standard deviation increase in the average test score conditional on neighborhood sociodemographics initially reported in column 3 of Table 2. The remaining rows report the difference in MWTP associated with the comparison of household characteristics shown in the row heading. Thus, the second row shows how a household's MWTP changes with an increase of $100,000 in income.

Analogously to the effect of including neighborhood sociodemographic variables on the estimates of the mean MWTP reported in Table 2, including these variables decreases the magnitude of the estimated heterogeneity in demand for school quality, decreasing the coefficient associated with household income by 60 percent, with education by 75 percent, and with race (which may proxy in part for wealth) by upwards of 80 percent. The estimates of the specification that includes neighborhood sociodemographics also returns the expected positive additional MWTP associated with the presence of school-aged children. Table 4 reports analogous measures of the heterogeneity in MWTP for selected housing and neighborhood attributes from the specification that includes neighborhood sociodemographic characteristics, revealing for example that increases in household income are typically associated with large increases in the demand for housing attributes, and that college educated households are willing to pay a substantial premium relative to less educated households to live with more educated neighbors. Specifically, they are willing to pay around $59 per month more than less educated

[30] We treat a household's primary work location as exogenous, calculating the distance from this location to the location of the neighborhood in question. MWTP estimates for other housing and neighborhood attributes based on a specification without commuting distance are qualitatively similar except for variables that are strongly correlated with employment access, such as population density.

households for a 10 percent increase in college-educated neighbors. Not surprisingly, strong racial interactions are also at work in the housing market, leading to significant sorting along this dimension as well.[31]

Constructing Measures of the Residual Elasticity of Demand Faced by Each School

Given the estimates of the demand system, we now calculate a measure of the residual elasticity faced by each school. For each of 708 elementary schools in the Bay Area, we use the equilibrium model to conduct a simple counterfactual simulation. In particular, we raise each school's average test score by a 0.1 standard deviation (7.762 points on a mean of 522) and then calculating the new housing market equilibrium. In every case, this has the effect of increasing house values in the corresponding neighborhood catchment area. The elasticity measure that we use in the subsequent analysis consists in estimated change in average local property values per standard deviation increase in test score. Across the 708 elementary schools, the corresponding increase in house values following this increase ranges from $1,400 to $19,600 ($5-$80 in monthly house price), with a mean of $11,800 ($45.5) and a standard deviation of $3,200 ($12.4).[32]

Figure 1 shows the geographic distribution of the estimated elasticities across the Bay Area. In the figure, the area of the circle indicates the magnitude of the estimated elasticity. In general, a school's residual elasticity of demand (as measured by the elasticity of house values with respect to school quality) will be a function of two features of its environment: (i) the willingness of the households that it serves to pay for improvements in school quality and (ii) the availability of close substitutes in geographically proximate neighborhoods. The former type of variation in the elasticity measure is problematic from the point of view of estimating the productive effect of competition in that a household's willingness-to-pay (WTP) for school quality is likely to be correlated with the performance of its children on standardized tests for

[31] Note that the strength of the estimated racial interactions reported in Table 4 may reflect the presence of centralized discriminatory practices in the market in addition to the direct preferences of households to live with others of the same race. See Bayer, McMillan, and Rueben (2004b) for more discussion of this issue.

[32] The basic structure of the computation of the new equilibrium consists of a loop within a loop. Having lowered the test score of a given school by a standard deviation, we first calculate a new set of prices that clears the market; Berry (1994) ensures that there is a unique set of market clearing prices up to scale. Using these new prices and the initial sociodemographic composition of each neighborhood, we calculate the probability that each household chooses each housing type, and aggregating these choices to the neighborhood level, the corresponding predicted sociodemographic composition of each neighborhood. We replace the initial neighborhood sociodemographic measures with these new measures and start the loop again – i.e., calculate a new set of market clearing prices with these updated neighborhood sociodemographic measures. We continue this process until the neighborhood sociodemographic measures converge. The household location decisions corresponding to the final sociodemographic measures along with the vector of housing prices that clears the market then represents the new equilibrium.

reasons that have nothing to do with the school itself. The latter form of variation is much less problematic in that a school and its corresponding neighborhood in any quality range can be located such that it has either many similar or dissimilar neighboring school catchment areas. It is this latter form of variation that we would like to exploit in our analysis.

Looking directly at the model of residential sorting estimated above, it is clear that the household sociodemographic characteristics included in the model, such as parental education, increase a household's estimated WTP for school quality and therefore increase the estimated elasticity of demand for schools that serve these households. Thus, in every specification of the analysis that follows, we include a complete set of controls for neighborhood averages of *all* of the household sociodemographic and other housing and neighborhood measures included in our demand estimation. Without including this full set of controls, the elasticity measure would be mechanically correlated with the average school test score - because highly educated households are willing to pay more for school quality *and* select in to schools with higher test scores, the elasticity of demand for these schools is mechanically greater. When this full set of controls is included in estimating our main regression equations, however, this mechanical correlation is eliminated; doing so has the effect of reducing the effective variation in our main elasticity measure to the variation closely related to the availability of close substitutes in the local market.

To demonstrate this, we re-estimated the regressions reported below in Table 7 using an estimate of the residual elasticity of demand drawn from a model of residential sorting that did not include distance to work (i.e., ignored geography). When these elasticities were included in average test score regressions with the full set of controls described here, the coefficients were essentially zero and highly insignificant in every case. This gives us confidence that any mechanical correlation associated with the increased willingness of some households to pay for school quality (e.g., highly-educated) is eliminated by the inclusion of a full set of controls for the variables used in the demand estimation. As the importance of including geography in the demand estimation makes clear, it is the notion of the availability of close substitutes in the local market that forms the basis for the remaining variation in our demand estimation.

Figure 2 shows the geographic distribution of the remaining variation in the estimated elasticities across the Bay Area once the full set of sociodemographic and other housing and neighborhood attributes included in the demand model have been conditioned out. For those familiar with the Bay Area, even a quick glance comparing Figure 1 and Figure 2 reveals that the remaining variation in the elasticity measure is not a simply a function of neighborhood socioeconomic conditions. For expositional ease, we work with a standardized version of this conditional elasticity measure throughout the remainder of our analysis. (Means and standard deviations for the school-related variables summarized for the 708 elementary schools used in the subsequent analysis are shown in Table 5.)

To demonstrate that the variation in these conditional elasticity measures is indeed related to the notion of the proximity of close substitutes, we construct a series of dissimilarity indices. Using the neighborhood catchment areas for the nearest 10 schools, we construct a measure of the average difference between the measure of the school in question and each of these neighbors. Thus, a high measure for a given dissimilarity index indicates that a school is quite distinct from its geographic neighbors.

These dissimilarity indices provide a simple (albeit crude) measure of the availability of close substitute neighborhoods in the local market. Table 6 shows the correlation between these dissimilarity indices and the variation in our conditional elasticity measure (once the full set of sociodemographic, housing, and other neighborhood measures used in estimating the demand model have been conditioned out). In each case, the correlation is negative, indicating that schools that are differentiated from their neighbors in terms of sociodemographic and housing characteristics tend to have lower elasticities *ceteris paribus*. This suggests that the remaining variation in our elasticity measure is indeed picking up the type of variation that we would like to exploit in estimating equation (1).

7 SUPPLY-SIDE REGRESSIONS

We now turn to our main regression analysis. In particular, we report results for the specification shown in equation (1) using a variety of school characteristics (output and input) as the dependent variable and a series of six distinct sets of control variables. (The variables included in each set of control variables are listed in Appendix Table 4.)

Table 7 reports results when the average 4^{th} grade test score is used as the dependent variable. As mentioned above, because we are using the conditional elasticity measure throughout our analysis, this amounts to including a complete set of controls for the neighborhood average of all household sociodemographic, housing, and other neighborhood variables used in estimating the demand side of our model in all specifications.

The first column of Table 7 includes only these variables as controls. The second column adds seventeen additional controls for 4^{th} grade school sociodemographics. These are included to account for the fact that the sociodemographic characteristics of the students in the school (important in the production of the test score) might differ from the neighborhood average sociodemographics. The third column adds controls for five measures of the employment access by education level as well as four direct geographic measures. These controls are included to account for the possibility that the households who live within the core of the Bay Area may be systematically different than those who live in outlying regions.[33] A full set of parameter estimates for the specification shown in the third column are shown in Appendix Table 5. The fourth column adds controls for interactions between neighborhood race, education and income measures. This ensures that the remaining variation in the elasticity measure is not an artifact of non-linearities in these important household sociodemographics. Column five adds higher-order house price and income terms, and finally, column six adds four local land-use measures.

The estimates reported in the six columns of Table 7 reveal a consistent pattern of results with the residual elasticity coefficient estimate falling in the range of 8.6-12.8 and the t-stat ranging between 3.0 and 4.0. These coefficients are reported for a standardized conditional elasticity measure and thus the interpretation is that a one standard deviation increase in the competitiveness of a school's local environment within the Bay Area is associated with a 10-12 point increase in the average test score of the school – about 0.15 standard deviations. This is a sizeable effect, indicating that a school's performance is indeed strongly linked to the competitiveness of its local environment.

We use a specification corresponding to the third column of Table 7 for the remaining analysis conducted in the paper. The first six columns of Table 8 report a series of such regressions using a various school input measures as the dependent variable. The first three columns relate to teacher experience and reveal that schools that face a greater elasticity of demand actually have significantly fewer of the most productive teachers (those with experience between 5-9 years) and tend to have a higher number of teachers who are just beginning and nearing the end of their careers. The point estimates for the effect of the elasticity on the dependent variables considered in the next three columns, the pupil-teacher ratio and teacher education variables, are highly insignificant and very small in magnitude. Thus, in general, schools facing a higher elasticity of demand appear to produce higher test scores without any significant advantages in terms of observable school inputs.

As discussed above, that school input decisions are not strongly related to the residual elasticity of demand in the Bay Area is not all that surprising given the restrictive financing regime in place in California. That schools facing greater amounts of local competitive pressures do not attract observably better teachers, however, provides some assurance that the remaining variation in our elasticity measure is not simply picking up unobserved student or neighborhood characteristics. Because teachers can sort across schools, we might be worried, for example, if a

[33] It is also worth noting that the inclusion of county fixed effects also does not affect the results.

higher residual demand elasticity was correlated with the presence of more experienced or better-educated teachers.

The final column of Table 8 provides another way to evaluate the possibility that the residual elasticity measure is correlated with unobserved student/household characteristics. In particular, column 7 reports the results for a specification that uses the average amount of income from capital sources in the corresponding neighborhood (the best proxy for wealth available in the Census) as the dependent variable. Importantly, this measure was not used in estimating the demand side of model and therefore serves as an appropriate test of whether the remaining variation in our elasticity measure proxies for the type of unobserved household characteristic that might be expected to positively affect test scores. As the results reveal, the point estimate is actually negative in this case (with a t-stat of -1.3) indicating that higher elasticity schools actually serve households with lower levels of capital income *ceteris paribus*. This evidence provides a further indication that the remaining variation in our elasticity measure is not simply proxying for unobserved household characteristics.

In the final two tables of the paper, we explore the relationship between the dissimilarity indices described above and the average 4th grade test score. Again, it is this type of variation in our elasticity measure that we would ideally like to exploit in estimating the main specification shown in Tables 7 and 8. Table 9 reports the results of a series of specifications that relate the standardized school elasticity measure to various dissimilarity indices and additional control variables. Not surprisingly, given the negative correlations in Table 6, these measures are negatively related to the elasticity measure and are significant in most cases. When various combinations of these dissimilarity measures are included directly in the average 4th grade test score regression in Table 10, they collectively enter negatively and significantly. This provides further assurance that the positive coefficient estimates reported in Table 7 are in fact reasonable.

8 CONCLUSION

Numerous studies have addressed the policy-relevant question of whether greater choice will lead to improvements in school quality. The typical focus in the prior literature has been on the *overall* effect of increased choice on school performance.

In the current paper, we began by making the conceptual point that this overall effect can be decomposed into a component measuring demand responsiveness (how increased choice affects school competition) and a second component measuring supply responsiveness (the way that increased competition affects school performance). By focusing on the overall effect of choice, the previous literature has provided little guidance as to effect of competition itself on school performance. Moreover, existing research that suggests demand responsiveness to increased choice may be weak, leaving open the possibility that supply responsiveness may be altogether more powerful.

This paper then presented a new approach for measuring the direct effect of competition on school performance – the strength of supply responsiveness. Central to this approach is the construction of a residual elasticity of demand for each school – our preferred measure of local competition – which captures the change in demand each school faces in response to a change in that school's quality. We do so using a flexible demand model, estimated using very rich Census data. This competition measure is then used in a regression framework that relates measured school performance to student, school and neighborhood controls, including our competition measure.

Our results provide the first estimates in the literature of the direct effect of increased competition on public school performance. We find evidence of a marked increase in test scores in response to an increase in the residual elasticity of demand: a one standard deviation increase in the competitiveness of a school's local market within the Bay Area is associated with a 0.15 standard deviation increase in school performance. At the same time, these achievement increases are accompanied by, if anything, slight reductions in important inputs, helping allay concerns about the role of omitted variables in these regressions.

These findings are robust to inclusion of many types of controls. Lending support to the notion that our preferred competition measure is not simply picking up unobserved household characteristics, we show that our residual elasticity has no effect in predicting neighborhood wealth. In contrast, it is correlated with similarity indices that describe the extent to which a school is isolated geographically (and in terms of product space): the residual elasticities increase the less isolated a school becomes. And as one might expect, these similarity measures also have positive effect on school performance. In terms of heterogeneous effects, we find that school responsiveness to increased competition is greater in more educated communities, suggesting that educated parents may be better able to monitor school personnel as competition increases.

Overall, our evidence is consistent with strong supply responsiveness on the part of public schools. This is relevant to the broader school choice debate, suggesting that policies that increase residual elasticities of demand may be promising.

REFERENCES

Barrow, Lisa and Cecelia Rouse (2004) "Using Market Valuation to Assess Public School Spending," *Journal of Public Economics*, 88:9-10, pp. 1747-69.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2003), "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," Economic Growth Center, Yale University Working Paper No. 872.

Bayer, Patrick, Robert McMillan, and Kim Rueben, (2004a), "An Equilibrium Model of an Urban Housing Market," unpublished manuscript, Yale University.

________, (2004b), "Residential Segregation in General Equilibrium," Economic Growth Center, Yale University Working Paper No. 885.

Belfield, Clive and Henry Levin (2002), "The Effects of Competition on Educational Outcomes: A Review of US Evidence," National Center for the Study of Privatization in Education Report, Columbia University.

Benabou, Roland, (1993), "The Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics*, 108(3), pp.619-652.

________, (1996), "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, 86(3), 584-609.

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.

Black, Sandra, (1999), "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, May 1999.

Borland, M and R. Howson (1992), "Students' Academic Achievement and the Degree of Market Concentration in Education," *Economics of Education Review*, 11: 31-39.

Cullen, Julie, Brian Jacob, and Steven Levitt, (2003), "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries," NBER Working Paper No. 10113.

Debreu, Gerard (1969), "Economies with a Finite Set of Equilibria," Irving Fisher lecture at the Brussels meeting of the Econometric Society, September, reprinted in *Mathematical Economics: Twenty Papers of Gerard Debreau*, Cambridge University Press, 1983.

Epple, Dennis, David Figlio, and Richard Romano, (2004), "Competition between Private and Public Schools: Testing Stratification and Pricing Predictions," *Journal of Public Economics* 88, 1215-45.

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

________, (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 Richard Romano, (1998), "Competition Between Private and Public Schools, Vouchers, and Peer Group Effects," *American Economic Review* 88(1): 33-62.

Epple, Dennis, Thomas Romer and Holger Sieg, (2001), "Interjurisdictional Sorting and Majority Rule: An Empirical Analysis," *Econometrica*, Vol. 69, No. 6., pp. 1437-1455.

Epple, Dennis and Holger Sieg, (1999), "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy*, Vol. 107, No. 4., pp. 645-681.

Fernandez, Raquel and Richard Rogerson, (1996), "Income Distribution, Communities, and the Quality of Public Education." *Quarterly Journal of Economics*, Vol. 111, No. 1., pp. 135-164.

________, (2003), "Equity and Resources: An Analysis of Educational Finance Systems," *Journal of Political Economy*, vol. 111, no. 4.

Ferreira, Fernando V. (2004), "*You Can Take It with You:* Transferability of Proposition 13 Tax Benefits, Residential Mobility, and Willingness to Pay for Housing Amenities." Center for Labor Economics, University of California, Berkeley, Working Paper No. 72.

Ferreyra, Maria, (2003), "Estimating the Effects of Private School Vouchers in Multi-District Economies," mimeo, Carnegie-Mellon University.

Figlio, David, and Joe Stone (1999) "School Choice and Student Performance: Are Private Schools Really Better?" *Research in Labor Economics*, 18: 115-40.

Hanushek, Eric and Steve Rivkin (2003) "Does Public School Competition Affect Teacher Quality," in *The Economics of School Choice*, edited by Caroline Hoxby, Chicago, IL: University of Chicago Press.

Hoxby, Caroline M., (1994) "Do Private Schools Provide Competition for Public Schools?" NBER Working Paper No. 4978.d

Hoxby, Caroline M., (2000) "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90 (5), 1209-38.

Jepsen, Christopher, (1999) "The Effects of Private School Competition on Student Achievement," mimeo, Northwestern University.

Kane, Thomas, Douglas Staiger, and Stephanie Riegg, (2004) "Do Good Schools or Good Neighbors Raise Property Values?" mimeo, University of California, Los Angeles.

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

Millimet, Daniel and Vasudha Rangaprasad (2005) "Strategic Competition Amongst Public Schools," mimeo, Southern Methodist University.

Nechyba, Thomas J., (1999), "School Finance-Induced Migration and Stratification Patterns: the Impact of Private School Vouchers," *Journal of Public Economic Theory*, Vol. 1.

________, (2000), "Mobility, Targeting and Private School Vouchers," *American Economic Review* 90(1), 130-46.

Rothstein, Jesse, (2003), “Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions,” mimeo, Princeton University.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzaf and Randall Walsh, (forthcoming) “Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods,” *International Economic Review*.

Tiebout, Charles M., (1956), “A Pure Theory of Local Expenditures,” *Journal of Political Economy*, 64: 416-424.

Timmins, Christopher (2003), “If You Can't Take the Heat, Get Out of the Cerrado... Recovering the Amenity Cost of Non-Marginal Climate Change in Brazil,” mimeo, Yale University.

Walsh, Randall (2004), “Endogenous Open Space Amenities in a Locational Equilibrium,” CEA Working Paper No. 04-03, February.

Figure 1: Actual Elasticities of Demand: San Francisco Bay Area Elementary Schools

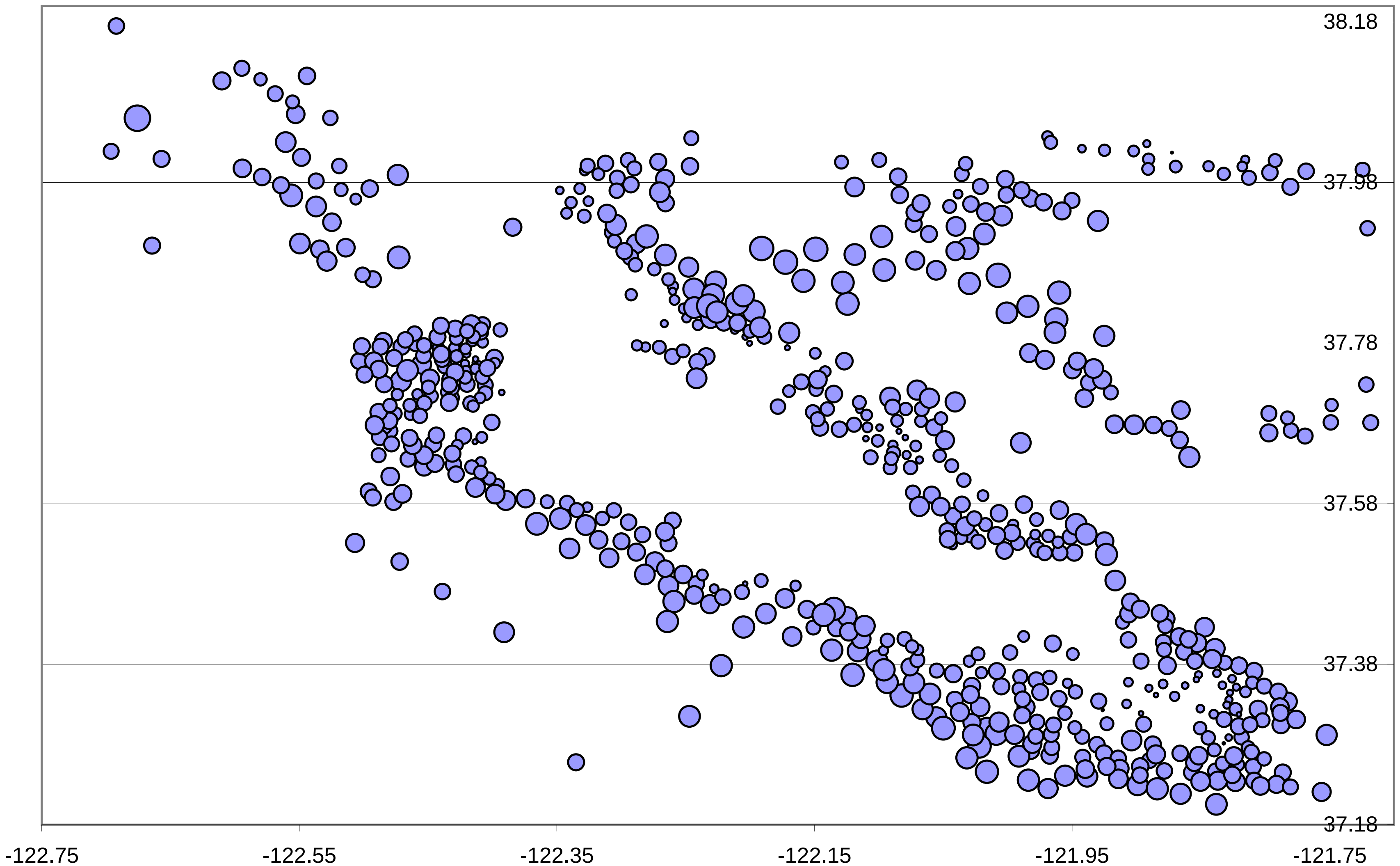

Figure 2: Conditional Elasticities of Demand: San Francisco Bay Area Elementary Schools

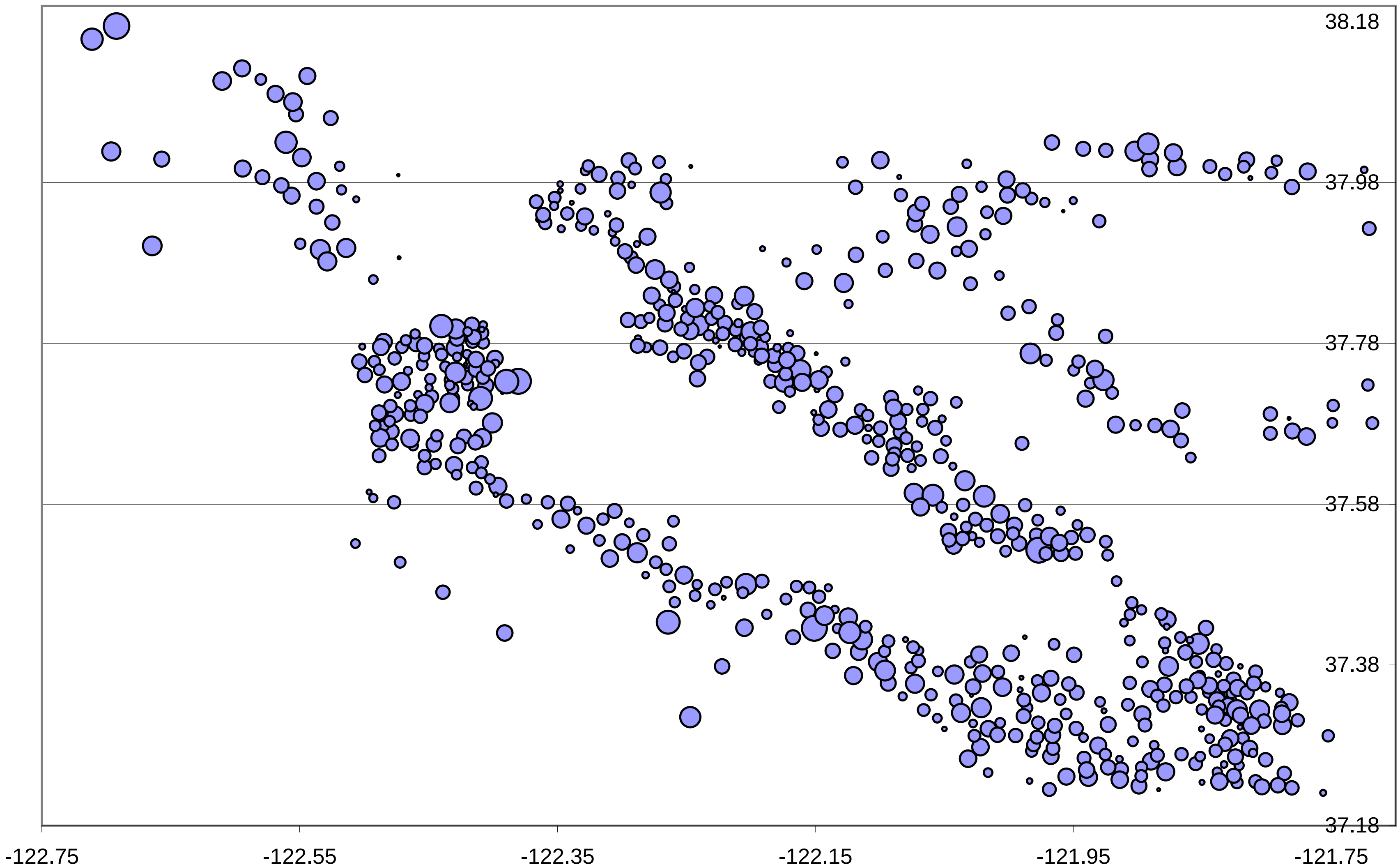

Table 1. Full Sample, and School District Boundary Sub-Sample

Sample	full sample		within 0.25 miles of boundaries				
Boundary/Weights			**actual sample**	**high test score side***	**low test score side***	**t-test for**	**weighted sample**
Observations	**242,100**		**27,958**	**13,348**	**14,610**	**difference in**	**27,958**
	(1)	(2)	(3)	(4)	(5)	**means**	(6)
	Mean	S.D.	Mean	Mean	Mean	((4) versus (5))	Mean
Housing/Neighborhood Characteristics							
monthly house price	1,087	755	1,130	1,158	1,105	5.71	1,098
average test score	527	74	536	558	515	50.96	529
1 if unit owned	0.597	0.491	0.629	0.632	0.626	1.04	0.616
number of rooms	5.114	1.992	5.170	5.207	5.134	3.13	5.180
1 if built in 1980s	0.143	0.350	0.108	0.118	0.099	5.09	0.148
1 if built in 1960s or 1970s	0.391	0.488	0.424	0.412	0.437	4.22	0.406
elevation	210	179	193	194	192	1.14	212
population density	0.434	0.497	0.352	0.349	0.355	2.08	0.374
crime index	8.184	10.777	6.100	6.000	6.192	2.36	7.000
% Census block group white	0.681	0.232	0.704	0.712	0.686	4.69	0.676
% Census block group black	0.081	0.159	0.071	0.065	0.076	3.01	0.080
% Census block group Hispanic	0.110	0.114	0.113	0.107	0.119	4.56	0.117
% Census block group Asian	0.122	0.120	0.112	0.110	0.113	1.12	0.121
% block group college degree or more	0.438	0.196	0.457	0.463	0.451	2.89	0.433
average block group income	54,744	26,075	57,039	58,771	55,457	4.65	55,262
Household Characteristics							
household income	54,103	50,719	56,663	58,041	55,405	4.20	55,498
1 if children under 18 in household	0.333	0.471	0.324	0.322	0.325	0.54	0.336
1 if black	0.076	0.264	0.066	0.062	0.070	2.69	0.076
1 if Hispanic	0.109	0.312	0.111	0.102	0.119	4.54	0.115
1 if Asian	0.124	0.329	0.112	0.114	0.110	1.06	0.121
1 if white	0.686	0.464	0.706	0.717	0.696	3.86	0.682
1 if less than high school	0.154	0.361	0.141	0.134	0.147	3.12	0.152
1 if high school	0.184	0.388	0.176	0.177	0.175	0.44	0.183
1 if some college	0.223	0.417	0.222	0.222	0.223	0.20	0.225
1 if college degree	0.291	0.454	0.294	0.295	0.294	0.18	0.286
1 if more than college	0.147	0.354	0.166	0.172	0.161	2.46	0.155
age (years)	47.607	16.619	47.890	48.104	47.699	1.99	47.660
1 if working	0.698	0.459	0.705	0.702	0.709	1.28	0.701
distance to work (miles)	8.843	8.597	8.450	8.412	8.492	0.82	8.490

Notes: Columns 1 and 2 report the mean and standard deviation for key variables for the full sample. Column 3 reports means for the sample of houses within 0.25 miles of a school district boundary. Columns 4 and 5 report means on the high versus low test score side of boundaries. Column 6 provides a t-statistic for a test of whether the means reported in columns 4 and 5 are equal. Column 7 reports weighted means for the sample of houses within 0.25 miles of a school district boundary. Weights are constructed so as to make the boundary sample more representative of the full sample and are described in the main text. In constructing columns 4 and 5, we assign each house in the full sample to the nearest school district boundary, noting whether its local school has a higher test score than the school associated with the closest Census block on the other side of the boundary.

Table 2: Implied Mean MWTP Measures

	Without Neighborhood Sociodemographics			With Neighborhood Sociodemographics		
Sample	full sample	within .25 mile of boundaries		full sample	within .25 mile of boundaries	
Boundary Fized Effects	No	No	Yes	No	No	Yes
Observations	242,100	27,958	27,958	242,100	27,958	27,958
	(1)	(2)	(3)	(4)	(5)	(6)
average test score (in standard deviations)	**126.08**	**122.89**	**81.53**	**20.17**	**20.19**	**26.22**
	(1.96)	**(5.36)**	**(7.72)**	**(1.72)**	**(4.77)**	**(6.13)**
1 if unit owned	209.76	178.37	184.54	165.38	150.77	161.05
	(3.29)	(8.99)	(11.39)	(3.19)	(8.76)	(9.24)
number of rooms	148.98	149.36	138.71	122.03	121.12	118.93
	(1.51)	(4.24)	(5.49)	(1.48)	(4.23)	(4.40)
1 if built in 1980s	129.93	74.74	106.17	99.69	85.50	95.55
	(3.94)	(10.87)	(14.41)	(3.79)	(10.69)	(11.84)
1 if built in 1960s or 1970s	28.48	9.46	15.39	13.79	7.40	4.50
	(2.78)	(8.03)	(10.48)	(2.67)	(7.71)	(8.51)
elevation (/100)	21.09	-4.82	46.46	-1.06	-18.04	12.83
	(0.81)	(2.48)	(6.35)	(0.75)	(2.46)	(5.04)
population density	-100.43	-153.53	-133.08	19.41	41.68	30.33
	(4.23)	(15.64)	(23.85)	(4.30)	(15.76)	(20.09)
crime index	-2.95	-2.30	1.78	0.00	-1.39	1.96
	(0.18)	(0.70)	(2.20)	(0.20)	(0.81)	(1.91)
% Census block group black				-324.67	-318.83	-267.08
				(10.14)	(32.15)	(39.84)
% Census block group Hispanic				-4.42	18.06	138.95
				(14.35)	(46.87)	(63.13)
% Census block group Asian				-97.39	-96.22	155.27
				(11.15)	(37.39)	(55.73)
% block group college degree or more				286.02	206.02	137.71
				(10.50)	(30.58)	(44.53)
average block group income				87.08	96.11	87.61
				(1.25)	(3.86)	(4.00)
F-statistic for boundary fixed effects			5.349			4.162

Note: Specifications shown in the table also include controls for land use (% industrial, % residential, % commercial, % open space, % other) in 1, 2, 3, 4, and 5 mile rings around location and six variables that characterize the housing stock in each of these rings. MWTP measures are reported in terms of a monthly house price. Standard errors are in parentheses.

Table 3. Heterogeneity in Marginal Willingness to Pay for School Average Test Score

	Without Neighborhood Sociodemographics in Model	**With Neighborhood Sociodemographics in Model**
	One Standard Deviation Increase in Average Test Score	
Mean MWTP	81.53	26.22
	(7.72)	(6.13)
Heterogeneity in MWTP		
Household Income (+$100,000)	40.45	15.66
	(0.28)	(0.35)
Children Under 18 vs.	-11.90	7.10
No Children	(3.07)	(3.78)
Black vs. White	-93.84	-18.05
	(5.15)	(7.50)
Hispanic vs. White	-40.75	-4.64
	(4.64)	(5.80)
Asian vs. White	-9.08	5.79
	(3.99)	(5.08)
College Degree or More vs.	57.65	14.12
Some College or Less	(3.46)	(4.24)
Householder Working vs.	1.91	6.63
Not Working	(3.16)	(4.02)
Age (+10 years)	1.02	0.86
	(0.09)	(0.11)

Note: The first row of the table reports the mean marginal willingness-to-pay for the change reported in the column heading. The remaining rows report the difference in willingness to pay associated with the change listed in the row heading, holding all other factors equal. Standard errors are in parentheses.

Table 4. Heterogeneity in Marginal Willingness to Pay for Select Housing and Neighborhood Attributes

Specification Includes Neighborhood Sociodemographic Charactersitics

	House Characteristics			Neighborhood Sociodemographics				
	Own vs. Rent	+1 Room	Built in 1980s vs. pre-1960	+10% Black vs. White	+10% Hisp vs. White	+10% Asian vs. White	+10% College Educated	Blk Group Avg Income + $10,000
Mean MWTP	161.05	118.93	95.55	-26.71	13.90	15.53	13.77	87.61
	(9.24)	(4.40)	(11.84)	(3.98)	(6.31)	(5.57)	(4.45)	(4.00)
Household Income (+$100,000)	218.37	61.19	105.07	-15.32	7.73	-0.50	26.17	15.44
	(7.13)	(1.70)	(7.62)	(3.89)	(4.13)	(2.54)	(2.18)	(1.13)
Children Under 18 vs.	-12.87	40.06	-24.52	10.38	15.03	12.17	-14.18	5.05
No Children	(6.67)	(1.80)	(7.94)	(2.70)	(3.41)	(2.51)	(2.15)	(1.06)
Black vs. White	-63.55	1.56	2.95	96.82	46.13	48.02	16.99	-0.45
	(13.25)	(3.40)	(16.38)	(3.62)	(5.75)	(4.84)	(4.40)	(2.27)
Hispanic vs. White	-6.44	-14.14	-8.07	28.89	81.36	18.01	5.43	2.07
	(9.53)	(2.63)	(12.00)	(3.68)	(4.01)	(3.81)	(3.19)	(1.41)
Asian vs. White	113.65	-32.92	43.94	27.74	21.95	92.49	-0.05	1.99
	(8.96)	(2.27)	(10.77)	(3.64)	(4.67)	(2.78)	(2.91)	(1.41)
College Degree or More vs.	33.83	4.50	42.06	8.34	-4.16	-12.70	59.29	3.66
Some College or Less	(7.67)	(2.05)	(9.57)	(3.27)	(3.94)	(2.91)	(2.37)	(1.29)
Householder Working vs.	42.72	3.69	60.60	-4.71	-2.17	-2.81	-12.62	3.88
Not Working	(7.31)	(1.94)	(8.92)	(2.88)	(3.65)	(2.82)	(2.27)	(1.04)
Age (+10 years)	6.49	0.30	-2.07	-0.15	-0.56	-0.03	-0.12	0.11
	(0.21)	(0.06)	(0.25)	(0.08)	(0.10)	(0.08)	(0.06)	(0.03)

Note: The first row of the table reports the mean marginal willingness-to-pay for the change reported in the column heading. The remaining rows report the difference in willingness to pay associated with the change listed in the row heading, holding all other factors equal. Standard errors are in parentheses.

Table 5 - Summary Statistics for School-Related Variables (N=708)

Variable Description	Mean	Std Dev
School Characteristics		
Residual Elasticity Measure	45.53	12.37
Average 4th grade test score	522.27	77.62
Proportion of teachers with 0-4 years experience	19.63	13.29
Proportion of teachers with 5-9 years experience	14.90	10.09
Proportion of teachers with more than 10 years experience	65.46	17.02
Pupil teacher ratio	23.98	3.09
Proportion of teachers with Max B.A. or less	9.53	12.81
Proportion of teachers with Min M.A. or more	26.28	13.92
4th Grade School Sociodemographics		
% AFDC	14.87	15.95
% Moved in Past Year	16.09	10.34
% Asian	14.70	17.49
% Black	13.50	22.34
% Hispanic	22.13	24.62
% White or Other Race	49.66	36.79
% Parent Educ Category 1	23.59	21.62
% Parent Educ Category 2	20.01	11.32
% Parent Educ Category 3	29.08	15.76
% Parent Educ Category 4	13.59	14.52
% Parent Educ Category 5	6.84	12.72
% Parent Educ Category 6	2.46	5.10
% Parent Educ Category 7	4.42	11.08
% Parent Language Category 1	72.44	21.36
% Parent Language Category 2	17.35	14.59
% Parent Language Category 3	10.23	13.48
% Parent Language Category 4	3.56	7.83
School Dissimilarity Indices		
avg diff btwn average income in school catchment area vs the 10 closest schools	9677.34	11018.11
avg diff btwn proportion of householders with higher ed in school catchment area vs the 10 close:	0.077	0.064
avg diff btwn average number of rooms in school catchment area vs the 10 closest schools.	0.54	0.43
avg diff btwn proportion of home owners in school catchment area vs the 10 closest schools.	0.12	0.09
avg diff btwn average monthly housing price in school catchment area vs the 10 closest schools	204.78	238.07

Note: This table reports means and standard deviations for school-related variables summairzed for the 708 schools used in the subsequent analysis.

Table 6 - Correlation Matrix Conditional Elasticity Measure and Dissimilarity Indices

Obs=708

	Cond. Elasticity	**D10 - Income**	**D10 - Education**	**D10 - Rooms**	**D10 - Ownership**	**D10 - Price**
Conditional Elasticity	1.000					
D10 - Income	-0.069	1.000				
D10 - Education	-0.086	0.510	1.000			
D10 - Rooms	-0.111	0.681	0.374	1.000		
D10 - Ownership	-0.098	0.427	0.249	0.716	1.000	
D10 - Price	-0.138	0.859	0.485	0.649	0.393	1.000

D10 - Income	dissimilarity index: avg. absolute diff. between average income in school catchment area vs. the 10 closest schools.
D10 - Education	dissimilarity index: avg. absolute diff. between % householders with higher education in school catchment area vs. the 10 closest scho
D10 - Rooms	dissimilarity index: avg. absolute diff. between average number of rooms in school catchment area vs. the 10 closest schools.
D10 - Ownership	dissimilarity index: avg. absolute diff. between proportion of home owners in school catchment area vs. the 10 closest schools.
D10 - Price	dissimilarity index: avg. absolute diff. between average monthly housing price in school catchment area vs. the 10 closest schools.

Note: The conditional elasticity measure used in this table is the residual from a regression of the actual elasticity measure on the full set of sociodemographic, school, housing, and neighborhood controls used in column 3 in Table 7. This table summarizes the correlation between this conditional elasticity measure and five dissimilarity indices that measure the average absolute difference between the measure associated with a given school and those of the ten nearest schools.

Table 7 - Regressions of Test Score on School Elasticity Measure with Control Variables

Dependent Variable	**Average 4th Grade Test Score**					
St. Dev. of Dep. Var.	77.62					
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Standardized Conditional Elasticity Measure	12.839	8.637	9.651	10.826	11.649	11.696
(St. Dev. = 1.0)	(3.312)	(2.952)	(3.256)	(3.468)	(3.929)	(3.926)
Control Variables Included in Specification						
Neighborhood Sociodemographics	Yes	Yes	Yes	Yes	Yes	Yes
Housing and Neighborhood Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School Sociodemographics		Yes	Yes	Yes	Yes	Yes
Employment Access and Geographic Variables			Yes	Yes	Yes	Yes
Neighborhood Race-Education-Income Interactions				Yes	Yes	Yes
Higher-Order House Value and Income Terms					Yes	Yes
Land-Use Variables						Yes
Obs	708	708	708	708	708	708
R2	0.703	0.788	0.793	0.805	0.807	0.807

Note: This table reports the results of six specifications of a regression of the average 4th grade test score on the standardized elasticity measure and seven sets of control variables. The variables included in each set of controls variables are listed in Appendix Table 4. The complete results for the specification reported in the third column are shown in Appendix Table 5. Standard errors are reported in parentheses.

Table 8 - Regressions of School and Neighborhood Characteristics on Elasticity Measure with Control Variables

Dependent Variable	**% Teachers <5 years exp.**	**% Teachers 5-9 years exp.**	**% Teachers >= 10 years exp.**	**Pupil-Teacher Ratio**	**% Teachers w/ Max BA**	**% Teachers w/ Min MA**	**Avg. N'hood Capital Income (/10.000)**
St. Dev. Dependent Variable	13.29	10.09	17.02	3.09	12.81	13.92	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Standardized Conditional Elasticity Measure	1.033	-2.477	1.444	0.055	-0.071	-0.055	-0.398
(St. Dev. = 1.0)	(0.857)	(0.853)	(1.217)	(0.180)	(0.602)	(0.721)	(0.316)
Control Variables Included in Specification							
Neighborhood Sociodemographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Housing and Neighborhood Chars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Sociodemographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment Access and Geographic Vars.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	708	708	708	708	708	708	708
R^2	0.160	0.186	0.199	0.188	0.217	0.293	0.834

Note: This table reports the results of seven specifications of a regression of various school inputs and neighborhood capital income on the standardized elasticity measure and four sets of control variables. The variables included in each set of controls variables are listed in Appendix Table 4. Standard errors are reported in parentheses.

Table 9 - Regressions of Elasticity Measure on Control Variables and Dissimilarity Measures

Dependent Variable	**Standardized Conditional Elasticity Measure**			
St. Dev. of Dep. Var.	1.00			
	(1)	(2)	(3)	(4)
Neighborhood Income Dissimilarity Index (/10,000)	-0.108			0.208
(St. Dev. = 1.102)	(0.079)			(0.225)
House Price Dissimilarity Index (/100)		-0.100		-0.172
(St. Dev. = 2.381)		(0.023)		(0.099)
Neighborhood Education Dissimilarity Index			-1.295	-0.416
(St. Dev. = 0.064)			(0.689)	(0.727)
Control Variables Included in Specification				
Neighborhood Sociodemographics	Yes	Yes	Yes	Yes
Housing and Neighborhood Characteristics	Yes	Yes	Yes	Yes
School Sociodemographics	Yes	Yes	Yes	Yes
Employment Access and Geographic Variables	Yes	Yes	Yes	Yes
Obs	708	708	708	708
R^2	0.953	0.954	0.953	0.955

Note: This table reports the results of four specifications of a regression of the standardized elasticity measure on various dissimilarity indices and four sets of control variables. The variables included in each set of controls variables are listed in Appendix Table 4. Standard errors are reported in parentheses.

Table 10 - Regression of Test Scores on Control Variables and Dissimilarity Measures

Dependent Variable	**Average 4th Grade Test Score**					
St. Dev. of Dep. Var.	77.62					
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Neighborhood Income Dissimilarity Index (/10,000)	6.203	6.591	-3.473	-3.466		
(St. Dev. = 1.102)	(3.539)	(3.924)	(2.092)	(2.108)		
House Price Dissimilarity Index (/100)	-3.379	-3.919			-9.620	-8.499
(St. Dev. = 2.381)	(1.603)	(1.545)			(2.828)	(2.600)
Neighborhood Education Dissimilarity Index	-93.269	-77.123	-43.583	-31.262	-37.878	-26.809
(St. Dev. = 0.064)	(35.970)	(33.277)	(26.278)	(27.135)	(24.450)	(25.405)
Control Variables Included in Specification						
Neighborhood Sociodemographics	Yes	Yes	Yes	Yes	Yes	Yes
Housing and Neighborhood Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
School Sociodemographics	Yes	Yes	Yes	Yes	Yes	Yes
Employment Access and Geographic Variables	Yes		Yes		Yes	
Obs	708	708	708	708	708	708
$\mathbf{R^2}$	0.791	0.785	0.785	0.780	0.789	0.783

Note: This table reports the results of six specifications of a regression of the average 4th grade test score on various combinations of dissimilarity measures control variables. The variables included in each set of controls variables are listed in Appendix Table 4. Standard errors are reported in parentheses.

Appendix Table 1: Choice-Specific Constant Regressions

	Without Neighborhood Sociodemographics			With Neighborhood Sociodemographics		
Sample	full sample	within .25 mile of boundaries		full sample	within .25 mile of boundaries	
Boundary Fized Effects	No	No	Yes	No	No	Yes
Observations	242,100	27,958	27,958	242,100	27,958	27,958
monthly housing price (/1000)	-10.23	-9.73	-11.34	-15.94	-15.97	-16.19
	(1.39)	(1.13)	(1.36)	(1.71)	(1.56)	(1.69)
average test score (in standard deviations)	1.29	1.20	0.92	0.32	0.32	0.42
	(0.02)	(0.05)	(0.01)	(0.03)	(0.08)	(0.01)
1 if unit owned	2.15	1.74	2.09	2.64	2.41	2.61
	(0.03)	(0.09)	(0.01)	(0.05)	(0.14)	(0.01)
number of rooms	1.52	1.45	1.57	1.95	1.93	1.93
	(0.02)	(0.04)	(0.01)	(0.02)	(0.07)	(0.01)
1 if built in 1980s	1.33	0.73	1.20	1.59	1.37	1.55
	(0.04)	(0.11)	(0.02)	(0.06)	(0.17)	(0.02)
1 if built in 1960s or 1970s	0.29	0.09	0.17	0.22	0.12	0.07
	(0.03)	(0.08)	(0.01)	(0.04)	(0.12)	(0.01)
elevation (/100)	0.22	-0.05	0.53	-0.02	-0.29	0.21
	(0.01)	(0.02)	(0.01)	(0.01)	(0.04)	(0.01)
population density	-1.03	-1.49	-1.51	0.31	0.67	0.49
	(0.04)	(0.15)	(0.03)	(0.07)	(0.25)	(0.03)
crime index	-0.03	-0.02	0.02	0.00	-0.02	0.03
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
% Census block group black				-5.18	-5.09	-4.32
				(0.16)	(0.51)	(0.06)
% Census block group Hispanic				-0.07	0.29	2.25
				(0.23)	(0.75)	(0.10)
% Census block group Asian				-1.55	-1.54	2.51
				(0.18)	(0.60)	(0.09)
% block group college degree or more				4.56	3.29	2.23
				(0.17)	(0.49)	(0.07)
average block group income				1.39	1.53	1.42
				(0.02)	(0.06)	(0.01)
F-statistic for boundary fixed effects			4.545			3.963

Note: Specifications shown in the table also include controls for land use (% industrial, % residential, % commercial, % open space, % other) in 1, 2, 3, 4, and 5 mile rings around location and six variables that characterize the housing stock in each of these

Appendix Table 2. Interaction Parameter Estimates - Model Without Neighborhood Sociodemographics

		House Characteristics					Neighborhood Attributes			
Household Characteristics	**Average Test Score (+1 s.d.)**	Monthly House Price (/1000)	Owner Occupied	Number of Rooms	Built in 1980s	Built in 1960-1979	Elevation (/100)	Population Density	Crime Index	Distance to Work
household income	**0.050**	0.121	0.305	0.074	0.142	0.038	0.016	0.028	-0.001	-0.004
(/10,000)	**(0.004)**	(0.003)	(0.010)	(0.002)	(0.011)	(0.009)	(0.001)	(0.013)	(0.001)	(0.001)
1 if children under 18 in household	**-0.190**	0.063	-0.102	0.544	-0.316	0.146	0.010	-0.740	0.015	0.036
	(0.047)	(0.065)	(0.094)	(0.025)	(0.112)	(0.083)	(0.022)	(0.101)	(0.005)	(0.005)
1 if black	**-1.395**	-0.941	-0.510	0.152	0.004	0.401	-0.062	-1.285	0.110	-0.023
	(0.080)	(0.127)	(0.167)	(0.044)	(0.211)	(0.144)	(0.041)	(0.159)	(0.007)	(0.011)
1 if Hispanic	**-0.642**	0.168	-0.036	-0.268	-0.180	-0.157	-0.104	-0.155	0.050	0.014
	(0.072)	(0.122)	(0.130)	(0.036)	(0.164)	(0.115)	(0.040)	(0.136)	(0.007)	(0.007)
1 if Asian	**-0.167**	0.315	1.765	-0.503	1.037	0.686	-0.015	0.941	0.030	0.003
	(0.062)	(0.080)	(0.122)	(0.031)	(0.145)	(0.108)	(0.028)	(0.095)	(0.006)	(0.007)
1 if college degree or more	**0.787**	0.917	-0.032	-0.012	0.489	-0.045	0.225	-0.007	0.031	-0.006
	(0.053)	(0.071)	(0.108)	(0.029)	(0.135)	(0.093)	(0.024)	(0.111)	(0.006)	(0.006)
1 if working	**0.007**	0.244	0.563	0.032	0.641	0.406	-0.048	-0.437	-0.027	-0.858
	(0.049)	(0.067)	(0.103)	(0.027)	(0.125)	(0.086)	(0.025)	(0.097)	(0.005)	(0.008)
age (years)	**0.015**	0.010	0.090	0.004	-0.034	-0.009	0.003	-0.006	0.001	-0.001
	(0.001)	(0.002)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)	(0.003)	(0.000)	(0.000)

Note: The parameters shown describe the elements of the utility function that interact household characteristics, shown in row headings, with choice characteristics, shown in column headings. Standard errors are in parentheses.

Appendix Table 3. Interaction Parameter Estimates - Model With Neighborhood Sociodemographics

		House Characteristics					Neighborhood Attributes			Neighborhood Sociodemographics					
Household Characteristics	**Average Test Score (+1 s.d.)**	Monthly House Price (/1000)	Owner Occupied	Number of Rooms	Built in 1980s	Built in 1960-1979	Elevation (/100)	Population Density	Crime Index	% Block Group Black	% Block Group Hisp	% Block Group Asian	% Blk Group College	Blk Group Avg Income	Distance to Work (miles)
household income	**0.020**	0.121	0.303	0.076	0.144	0.028	0.010	0.011	-0.001	-0.223	0.113	-0.009	0.385	0.012	-0.004
(+10,000)	**(0.005)**	(0.004)	(0.011)	(0.003)	(0.012)	(0.009)	(0.002)	(0.017)	(0.001)	(0.060)	(0.064)	(0.039)	(0.034)	(0.002)	(0.001)
1 if children under 18 in household	**0.102**	0.231	-0.238	0.582	-0.399	0.095	0.051	-0.947	0.002	1.594	2.294	1.857	-2.171	0.055	0.027
	(0.058)	(0.075)	(0.103)	(0.028)	(0.122)	(0.092)	(0.025)	(0.127)	(0.006)	(0.416)	(0.527)	(0.387)	(0.331)	(0.016)	(0.005)
1 if black	**-0.282**	0.143	-1.006	0.002	0.027	0.577	-0.068	-1.106	0.045	14.874	7.082	7.371	2.607	-0.023	-0.010
	(0.116)	(0.170)	(0.205)	(0.053)	(0.253)	(0.184)	(0.052)	(0.228)	(0.009)	(0.560)	(0.888)	(0.747)	(0.680)	(0.035)	(0.013)
1 if Hispanic	**-0.077**	0.204	-0.138	-0.246	-0.147	-0.248	-0.067	-0.128	0.005	4.435	12.471	2.757	0.830	0.011	0.012
	(0.089)	(0.139)	(0.147)	(0.041)	(0.185)	(0.131)	(0.045)	(0.169)	(0.008)	(0.568)	(0.620)	(0.587)	(0.492)	(0.022)	(0.008)
1 if Asian	**0.072**	0.558	1.633	-0.571	0.612	0.457	-0.006	-0.053	0.006	4.236	3.330	14.060	-0.016	-0.022	0.012
	(0.078)	(0.095)	(0.138)	(0.035)	(0.166)	(0.123)	(0.033)	(0.132)	(0.007)	(0.562)	(0.721)	(0.429)	(0.449)	(0.022)	(0.007)
1 if college degree or more	**0.200**	0.501	0.428	0.006	0.588	0.106	0.031	0.486	0.022	1.279	-0.638	-1.935	8.986	0.009	0.009
	(0.065)	(0.079)	(0.118)	(0.032)	(0.148)	(0.101)	(0.027)	(0.134)	(0.007)	(0.504)	(0.607)	(0.450)	(0.366)	(0.020)	(0.007)
1 if working	**0.093**	0.272	0.604	0.021	0.897	0.425	0.023	-0.515	-0.019	-0.712	-0.335	-0.434	-1.931	0.033	-0.896
	(0.062)	(0.074)	(0.113)	(0.030)	(0.138)	(0.096)	(0.028)	(0.125)	(0.007)	(0.444)	(0.563)	(0.436)	(0.350)	(0.016)	(0.009)
age (years)	**0.013**	0.011	0.097	0.003	-0.033	-0.010	0.003	-0.011	0.001	-0.022	-0.085	-0.005	-0.018	0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.001)	(0.004)	(0.003)	(0.001)	(0.003)	(0.000)	(0.013)	(0.016)	(0.013)	(0.010)	(0.001)	(0.000)

Note: The parameters shown describe the elements of the utility function that interact household characteristics, shown in row headings, with choice characteristics, shown in column headings. Standard errors are in parentheses.

Appendix Table 4 - List of Variables Included in Each Set of Controls

Neighborhood Sociodemographics
percent black in catchment area
percent asian in catchment area
percent hispanic in catchment area
percent of householders with higher education
average household income in catchment area
average age in catchment area
proportion of households with children under 18 in catchment area
proportion of householders who do not work in catchment area

Housing and Neighborhood Characteristics
average monthly house price in catchment area
proportion of homeowners in catchment area
average number of rooms in psuedo-catchment area
crime index in catchment area
proportion of houses built in the 1980's in catchment area
proportion of houses built in the 1960's & 70's in catchment area
population density in catchment area

School Sociodemographics
percent 4th grade AFDC
percent 4th grade moved in past year
percent of asian students in grade 4
percent of black students in grade 4
percent of hispanic students in grade 4
percent of white students in grade 4
Seven grade 4 parental education categories
Four grade 4 parental language categories

Employment Access and Geographic Variables
employment accessibility index for high school dropouts
employment accessibility index for high school graduates
employment accessibility index for those with some college
employment accessibility index for college graduates
employment accessibility index for those with an advanced degree
school latitude
school longitude
school latitude squared
school longitude squared

Neighborhood Race-dEucation-Income Interactions
percent asian*average income in catchment area
percent black*average income in catchment area
percent hispanic*average income in catchment area
percent asian*proportion high school graduates in catchment area
percent asian*proportion with some college in catchment area
percent asian*proportion with college degree
percent asian*proportion with advanced degree
percent black*proportion high school graduates in catchment area
percent black*proportion with some college in catchment area
percent black*proportion with college degree
percent black*proportion with advanced degree
percent hispanic*proportion high school graduates in catchment area
percent hispanic*proportion with some college in catchment area
percent hispanic*proportion with college degree
percent hispanic*proportion with advanced degree

Higher-Order House Price and Income Terms
average house price squared
average household income squared
average household income cubed
average household income to the fourth power

Land Use Variables
proportion of industrial land use in 1 mile radius
proportion of commercial land use in 1 mile radius
proportion of other urban land in 1 mile radius
proportion of open space in 1 mile radius

Note: This table lists the control variables included in each set of controls used in the analysis reported in Tables 7-10.

Appendix Table 5 - Regressions of Test Score on School Elasticity Measure with Control Variables

Dependent Variable	Average School Test Score	
St. Dev. of Dep. Var.	77.62	
Variable	Coefficient	Std Error
standardized conditional elasticity measure	9.651	3.256
percent black in catchment area	49.696	43.234
percent asian in catchment area	-50.009	32.223
percent hispanic in catchment area	34.204	42.690
percent of householders with higher education in catchment area	19.461	40.046
average income in catchment area	0.000	0.000
average age of catchment area	0.936	1.313
proportion of households with children under 18 in catchment area	77.380	32.513
proportion of householders who do not work in catchment area	8.275	63.310
average monthly house price in catchment area	0.003	0.019
proportion of homeowners in catchment area	-30.637	25.717
average number of rooms in catchment area	-14.155	7.093
crime index in catchment area	0.062	0.636
proportion of houses built in the 1980's in catchment area	-17.401	14.452
proportion of houses built in the 1960's & 70's in catchment area	-3.915	11.689
population density in catchment area	-3.928	10.449
c4pafdc	-0.544	0.248
c4pmobil	-0.423	0.176
percent of asian students in grade 4	0.372	0.124
percent of black students in grade 4	-0.331	0.214
percent of hispanic students in grade 4	-0.361	0.101
percent of white students in grade 4	0.320	0.106
percent grade 4 parental education category 1	1.237	1.826
percent grade 4 parental education category 2	1.103	1.794
percent grade 4 parental education category 3	0.452	1.824
percent grade 4 parental education category 4	0.380	1.866
percent grade 4 parental education category 5	0.025	1.804
percent grade 4 parental education category 6	-0.149	1.961
percent grade 4 parental education category 7	0.796	1.875
percent grade 4 parental language category 1	-4.281	3.397
percent grade 4 parental language category 2	-4.057	3.432
percent grade 4 parental language category 3	-4.496	3.437
percent grade 4 parental language category 4	0.198	0.195
employment accessibility index for high school dropouts	-0.049	0.014
employment accessibility index for high school graduates	0.015	0.017
employment accessibility index for those with some college	0.007	0.022
employment accessibility index for college graduates	0.000	0.013
employment accessibility index for those with an advanced degree	0.009	0.016
school latitude	-1025.008	3331.935
school longitude	5781.304	24468.960
school latitude squared	13.305	44.360
school longitude squared	11.841	50.081
Obs	708	
R2	0.793	

Note: This table reports the full specification corresponding to that report in column 3 of Table 7.