Parental Preferences and School Competition: Evidence from a Public School Choice Program

by

Justine S. Hastings
Yale University
and NBER

Thomas J. Kane
UCLA
and NBER

Douglas O. Staiger
Dartmouth College
and NBER

PRELIMINARY DRAFT
COMMENTS WELCOME
DO NOT QUOTE OR CITE WITHOUT PERMISSION

February 2005

The authors thank Lee Kindberg of the Charlotte-Mecklenburg School Board as well as Superintendent James Pughsley and his excellent staff for making this project possible. We would also like to thank Sean Hundtofte, Jeffrey Weinstein, and Orkun Sahmali for outstanding research assistance. The project is funded by a grant from the Institution for Social and Policy Studies at Yale University.

I. Introduction

School choice plans are intended to improve both equity and efficiency - to provide incentives for schools to compete on the basis of academic achievement and to provide broader access to quality schools. Yet the degree of competition generated by a choice plan will depend in large part on parents' preferences regarding proximity, school test scores and racial composition. For instance, if parents care primarily about travel convenience, a given school may be competing with only one or two nearby schools-- and not all the schools in the district. In such cases, the incentives created for individual schools to raise performance may be limited.

At the heart of the school choice debate is this important question: How will school quality be distributed under the new equilibrium? At one end of the debate, choice proponents predict that public school choice will lead to intense competition on academic quality. They predict a "tide that lifts all boats", with higher equilibrium quality at all schools and less concentration of students in the higher quality schools. At the other end of the debate, opponents predict that public school choice will result in "cream-skimming" and "vertical separation" – an equilibrium outcome where inner-city schools become worse as the few top students leave to go to better schools and the local school is left without pressure from local students to improve quality.

These quite disparate predictions are generated by two very different assumptions about the nature of preferences. The 'tide that lifts all boats' outcome will occur if all parents value school quality much more highly than anything else. Alternatively, if parents have very heterogeneous preferences for school quality, 'vertical separation' and/or 'cream skimming' equilibria may occur. Specifically, suppose that poor and/or minority families do not have strong preferences for school quality and, perhaps due to parental time constraints, strongly value proximity. Then neighborhood schools serving these students will have a 'captive audience' without competitive pressure to improve quality. They will face little competition for their students' attendance choices. In addition, suppose that students in disadvantaged minority groups who have high academic achievement relative to their peers value school quality substantially more. These students will attend non-local schools with higher quality. Local schools will serve the remainder (majority) of local students without pressure to improve quality due to the relatively low value local parents place on quality.

In this paper, we employ unique data from a school choice plan in Charlotte-Mecklenburg School District (CMS) in North Carolina to estimate parental preferences for school quality. We then investigate the implications of those preference estimates for competition on quality under school choice.

CMS introduced public school choice in the fall of 2002, after a race-based bussing plan was terminated by the courts. Under the choice plan, parents in the district were asked to submit their top three choices of schools for their children. The school district provided us with data on those choices, along with data on individual students (demographics, test scores, residential locations) and on the schools themselves (mean test scores, racial composition and location).

Using the multiple rankings submitted by parents to the school district, we estimate the distribution of preferences over school academic quality, school proximity, and school racial composition. We allow these preference distributions vary with student socio-economic status and academic ability. The estimates are generated from a mixed logit random coefficients demand model, allowing for heterogeneous preferences over school attributes. Preference heterogeneity is important for estimating realistic substitution patterns and demand elasticities in response to changes in school characteristics.

In addition to the multiple observations on parent's school choices, this policy experiment and data provide rich independent variation in key variables of interest. First, there is substantial variation in student characteristics. Charlotte-Mecklenburg school district is a large district (approximately 110,000 students) and quite diverse racially and economically. Approximately 95% of the students submitted choices for the choice plan. Within the student body, there are substantial fractions of Caucasian and African-American students both receiving and not receiving school lunch subsidies, allowing us to estimate rich demand models for all four socio-economic groups. Second, using data on location of students and schools, we were able to calculate minimum travel distances from each residence to each school. This geographic differentiation effectively varies the product attributes and choice set across students, while the multiple choices provide variation in the choice set within student. It is these two sources of variation that will aid estimation and identification of preference distribution parameters.

In addition, discontinuities in school assignment under the former race-based bussing program, the geographic placement of schools near the inner city to facilitate the bussing policies, and the sudden changes in districting for the transition to the school choice plan yield rich variation in school proximity and quality within student-level socio-economic groups. Much of this variation is generated independently from residential sorting across districts designed for the bussing for integration regime. This variation will aid in identification of preferences for quality versus preferences for proximity across each socio-economic category.

Our preliminary results indicate that i) parents value proximity highly, ii) the preference attached to a school's mean test score is substantially lower for low-income students (those qualifying for the federal Free and Reduced Price Lunch program) and for those living in low-income neighborhoods, iii) preferences for school test scores are increasing with the baseline academic ability, and iv) once controlling for preferences over racial composition of schools, African Americans and Caucasians have very similar preferences over school test scores and distance.

With demand estimates from the mixed logit model, we calculate the elasticity of demand for each school with respect to quality. In particular, we simulate the estimated increase in number of students choosing each school first if it were to increase the academic performance of its students by 10 points, all else equal. We find that demand at high-performing schools is very responsive to increases in quality, and that demand at low-performing schools is very un-responsive to increases in quality. However, lower-performing schools draw in students with higher average academic performance and who are less likely to receive lunch subsidies than their average student. We discuss the potential implications of our estimates and simulations for models of competition on quality under school choice, and the potential for integrating theses results into school choice mechanism design to maximize competitive pressure and rewards for academic improvement.

II. Previous Literature on School Choice and Competition

A number of papers in the economics of education use aggregate measures of market concentration in different parts of the country to try to infer the extent of school competition, and relate those indirect measures of competition to various academic outcomes. Borland and Howson (1992), Hanushek and Rivkin (2003), and Hoxby (2000) used Herfindahl indices to document competition among schools is associated with higher student performance on tests and higher teacher quality. Other papers have found that measures of market share in non-public schools (charter or private schools) are similarly associated with improved school performance (Hoxby, 1994, 2003; CITES). This approach is similar to cross-market studies of the relationship between market structure, conduct and performance in industrial organization. Estimating the relationship between market structure and performance generates difficulty in interpreting equilibrium relationships.

The more recent literature in industrial organization has focused on estimating underlying preference parameters of consumer's indirect utility to understand the demand, substitution patterns, and nature of competition between firms in differentiated products markets. In our context, estimates of preference parameters will yield estimates of elasticities faced by individual schools, allowing insights into the nature of competition on quality under school choice.

There is a substantial literature using surveys to elicit parental preferences. Typically parents are offered a list of school attributes—such as academic rigor, school safety, religious affiliation, school size, class size, extracurricular options, physical condition of facilities, racial composition, and travel convenience—and simply asked to rank their importance.[1] On such surveys, researchers have typically found that academic standards and teacher quality loom large in parents' minds.[2] Even among those attending private religious schools, parents often report academic quality to be paramount. (Convey (1986), Nelson (1988), Goldring and Bauch (1995)).

---

[1] For a review of the literature in this field prior to 1990, see Maddaus (1990) and Carnegie Foundation (1992).

[2] See Armor and Peiser (1998), Vanourek, Manno and Finn (1998), Greene, Howell and Peterson (1998), Kleitz, Weiher, Tedin and Matland (2000), Schneider, Marschall, Teske and Roch (1998), Nault and Uchitelle (1982), Williams, Hancher and Hutner (1983), Darling-Hammond and Kirby (1985).

Nevertheless, because stated preferences may not reflect behavior, inferring parents' actual preferences from such questionnaires can be misleading. For example, parents' may implicitly be limiting their choice sets in ways not apparent to the researcher. Among the schools on a parent's short list and, therefore, highly salient, academic excellence may have been a determining factor in their eventual choice-- but they may have only been considering schools within a certain travel range or within a certain range demographically. Moreover, parents' may tailor their responses to fit social norms - for example, over-reporting the importance they place on academic quality and under-reporting their potentially discriminatory views on the racial composition of schools.

A few studies have taken different approaches: For example, Schneider and Buckley (2002) monitored the search behavior of parents on an internet web site providing information on public schools in Washington, DC. Fossey (1994) studied the characteristics of the school districts that gained and lost students in a Massachusetts inter-district choice program. Van Dunk and Dickman (2002) not only asked parents to report what they valued in schools, but also tested their knowledge of those characteristics at the school their children were attending.

A smaller set of studies have exploited the actual choices parents make to infer parental preferences.[3] Bayer, Ferreira and McMillan (2003) use household location decisions to estimate household preferences over a broad range of housing, neighborhood and school characteristics. They find evidence of considerable preference heterogeneity, but do not derive the implications that this has for school demand elasticities. In a more directly related study of a school choice program in Minneapolis, Glazerman (1997) found that while test scores mattered in driving parental choices, parents tended to avoid schools in which their children's racial group represented less than 20 percent of all students. Glazerman used a conditional logit model to estimate parental preferences. However, the conditional logit estimates are subject to the well known independence of irrelevant alternatives property, with substitution patterns driven by the additive random component and the logit functional form assumption (Hausman and Wise 1978, Berry Levinsohn and Pakes 1995 and 2004). Using our multiple response choice data, we will

[3] See Manski and Wise (1983) for an early application to college choice.

instead be able to estimate a much more flexible mixed logit model, allowing for heterogeneity in preferences and generating much more realistic substitution patterns in responses to changes in the school characteristics and the school choice set.

III. Details of Public School Choice Plan in CMS

Before the introduction of a school choice plan in the fall of 2002, the Charlotte-Mecklenburg public school district (CMS) operated under a racial desegregation order for three decades. In September 2001, the U.S. Fourth Circuit Court of Appeals declared the school district "unitary" (as opposed to operating a "dual" system of education for white and black students) and ordered the district to dismantle the race-based student assignment plan by the beginning of the next school year. Several months later, in December of 2001, the school board voted to approve a new district-wide public school choice plan.

In the spring of 2002, parents were asked to submit their top three choices of school programs for each child. Each student was assigned a "home school" in their neighborhood, typically the closest school to them, and was guaranteed admission to this school if they were not admitted to any of their top three choices. Students were similarly guaranteed admission to continue in magnet programs in which they were enrolled in spring 2002. Admission to home schools and continuation magnet schools was guaranteed, while admission to other schools was limited by grade-specific capacity limits set by the district. These capacity limits were allowed to be substantially higher than past enrollment in many schools. The district allowed significant increases in school enrollment size in the first year of the school choice program in an expressed effort to give each child one of their top three choices. In the spring of 2002, the district received choice applications for approximately 105,000 of 110,000 students. Approximately 95% of parents received admission to one of their top three choices. Admission to over-subscribed schools was determined by a lottery system as described in Hastings, Kane and Staiger (2004).

Change in School Assignment Zones

In Charlotte, the creation of the public school choice system coincided with the dismantling of the racial desegregation plan. School assignment zones—which were often drawn to capture non-contiguous black and white neighborhoods to achieve racial balance—were thoroughly redrawn as a result of the Fourth Circuit court's ruling. In the fall of 2002 (the first year under the choice plan), for 43 percent of parcels, the default elementary school (or "home school") differed from the school assigned under the desegregation plan the year before. At the middle school and high school levels, 52 and 35 percent of parcels were assigned a different default school than they had been assigned the previous year under the desegregation plan. Moreover, even when the home school remained the same as under the desegregation plan, the composition of students with that school assigned as their home school changed dramatically due to changes in boundaries elsewhere.

Therefore, in our analysis, the home school for many students is often <u>not</u> the school their parents had previously chosen through their location. Similarly, the school that each student attended in 2002 is often not the home school in 2003, and may lie quite far from home. This dramatic change in school assignment zones implies that residential location was less likely to reflect endogenous sorting based on family preferences for a nearby school, in the sense that location near a school would not have guaranteed student enrollment in the prior year.

Evidence of Heterogeneity of Preferences

Interestingly, there was little unanimity in parents' choice of schools. For example, among those who would be in grades two through five during the 2002-03 school year, parents listed 93 different schools as their first choice. No single school represented the top choice for more than 2.7 percent of these parents. Some of the variance in parents' top choices of elementary schools is driven by differences in travel times to a given set of schools. But, even among those assigned to the same home school for 2002-03 (home schools were assigned by neighborhood), there was considerable heterogeneity in parental choices. Among those with the same elementary home school for 2002-03, parents on average listed 14.6 different elementary schools as their first choices. Such a diversity of choices implies that there is a considerable amount of heterogeneity in preferences, making the mixed logit modeling approach important.

Potential for Strategic Choice

The lottery mechanism used by the Charlotte-Mecklenburg schools was not strategy-proof (Abdulkadiroglu and Sonmez (2003)). For example, a student with a particularly undesirable default school might not list a very desirable school as first choice if there was a low probability of admission, and instead list a less desirable school with a better chance of admission, in an effort to hedge against not receiving admission to one of their choices and being assigned to their home school. Such strategic behavior would imply that student choices would not reflect true preference orderings for schools—to the extent that students are *not* listing their preferred match (in the absence of capacity constraints) at the top of their list.

However, the choice application was quite vague in describing how slots in oversubscribed schools were to be allocated and how the lottery system would be operated. In addition, the school district made an effort to communicate to parents that they would make every attempt to give each student admission to one of their chosen schools. Because of the uncertainty around the rationing rules and the efforts by the district to expand capacity at popular schools, we believe that the extent of strategic manipulation in the first year was limited and that parents were generally reporting their true preferences. We test for strategic hedging by using the exogenous redistricting of home schools under the School Choice Plan. Using the geographic boundaries for the 2001-2002 school year, and the new boundaries for 2002-2003, we test if students who were redistricted to lower-performing schools chose on average schools with lower test scores, relative to students in their same 2001-2002 district area who did not experience a negative shock to their default school quality. We did not find evidence that students who faced worse home schools after redistricting chose schools with significantly lower average quality than students in the same former district who were given better home school assignments.[4]

[4] In subsequent years of school choice, when capacities at schools were no longer changed to accommodate demand, strategy may have become more important. In the second year of choice, CMS no longer made an

IV. Data

Working with the Charlotte Mecklenburg School (CMS) Board and district officials, we have obtained secure access to a wide range of administrative data for all students in grades kindergarten through twelfth grade in the years surrounding the implementation of the choice program. The data falls in to five broad categories: (1) information from student choice forms, (2) geographic information, (3) student demographic information, (4) student test score information, and (5) information on school characteristics. These data are described in more detail below.

Choice Forms

We began with the choice forms submitted by 105,706 students in the first year. Reflecting the district's intensive outreach efforts, choice forms were received for 96 percent of all the students enrolling that fall. We dropped those who applied to special programs—including those designed for autistic children, the behaviorally/emotionally disabled, hearing impaired, learning disabled, orthopedically disabled, mentally disabled, hearing impaired, and English as a second language. This left a sample of 100,486.

For each student, we have the choice forms submitted to CMS, allowing a student to specify up to 3 choices for their school. We use the term "school" inclusively, to include all distinct academic programs (including distinct magnet programs that share a building) to which a student could apply. Overall, 35,754 students filled out only a first choice, 18,486 students listed only a first and second choice, and 46,246 students listed completely all three choices.

Geographic Information

We were given information on the exact location of each student's residence. We used that information as well as the exact location of all the schools to calculate the driving distance in miles (on the shortest route) to each of the schools in the district. In

---

effort to accommodate choices by changing school capacities. Many parents received none of their three choices, and expressed frustration because they had made choices without knowing the probability of admittance. In response, in the third choice year, CMS provided published probabilities of admittance to each school, so that parents could incorporate this information into their decision making process. We are currently using the choice responses across the three years to examine the effects of strategy on school choice.

addition, residential location was used to assign each student the median family income in his or her census block group for his or her race (from the 2000 census). Direct measures of family income are not available in the CMS administrative data, so this variable serves as a reasonable proxy.

Student Demographics

The CMS administrative data provided us with information on each student's grade, race (five categories, which we collapse into white and nonwhite), and eligibility for free or reduced lunch. The administrative data also provided information on which school each child was attending in the spring of 2002, at the time that the choice form was submitted.

Student Test Scores

We use data on reading and math scores on North Carolina end-of-grade exams for students in grades 3 through 8.[5] Test scores are reported as percentiles, based on grade level and year. We use reading and math scores from the spring of 2002 (before the choice program) as well as for the spring of 2003 and 2004—the first and second year after students were assigned to schools. Analyses that rely on these scores are limited to grades where the scores are available. For example, models that condition on baseline test scores are restricted to students who were in grades 3-8 in 2002.

As a measure of academic ability, we average each student's math and reading percentile score from the spring of the 2001-2002 school year, and then standardize by the mean and standard deviation of test scores for other students who attended the same school (e.g. a z-score). This measure captures how many standard deviations a given student was above (or below) their peers at the school they attended in the year prior to choice.

School Characteristics

[5] Students in kindergarten through 2nd grade do not take the state exams, and high school students only take end-of-course exams in the subjects they choose.

We constructed three school level measures that, while far from being exhaustive, capture the main dimensions that are often believed to influence school choice and that are central to the literature.[6] First, as already discussed, we calculated distance to each school. Second, to proxy for the academic quality of a school, we constructed the average test score (averaging both math and reading percentiles) of students attending each school in the first year of choice (2002-2003). Note that this calculation was done separately for each academic program whether or not they were housed in the same physical structure. By using scores from the first year of choice, we are assuming that student's made decisions based on rational expectations, but using prior year scores yields similar results. A third school characteristic that was likely to influence school choice, particularly given the history of court ordered desegregation, was racial composition of the school. We calculated the percent of the students in each school that were black, and used this to create categorical variables for schools that were 10-30%, 30-60%, and over 60% black.

V. Empirical Model

We begin this section with a brief overview of our estimation strategy and a discussion of why the CMS data is particularly useful for identifying heterogeneity in school preferences. We then describe more formally the model that we will use to estimate the preference parameters behind individual choices, how this model will be estimated, and how the results will be used to simulate school demand parameters of interest.

Overview

Each student listed as many as three school programs on his or her choice form in order of preference. Our empirical model uses these choices, along with data on each student and the school programs available, to estimate the parameters of the distribution preferences over school characteristics, and how they vary with student level SES category and academic achievement.

---

[6] In ongoing work, we plan to consider alternative measures of academic quality (including value added measures) and a more extensive list of other school features such as disciplinary problems, faculty characteristics, and characteristics of magnet programs.

Our estimates are derived from a mixed logit model (McFadden and Train 2000, Train 2003). Mixed logit discrete choice models of demand are multinomial choice models with *random coefficients* on product attributes in the indirect utility function. As discussed earlier, random coefficient discrete choice models have been used extensively in the industrial organization and marketing literatures to estimate preferences for product attributes, and thus demand elasticities and cross-elasticities for the products of interest. The mixed logit model that we estimate differs from the traditional conditional logit model in that it allows for a more flexible functional form on random preferences than the conditional logit does. In particular, the conditional logit restricts the random component of utility to enter only in an additively separable fashion. This restriction, while convenient, leads to the well-known independence of irrelevant alternatives problem. This restriction implies that, when choice sets are altered (for example by the introduction of a new school), substitution to that new school does not depend on its similarity to existing schools. That seems like an unrealistic assumption, particularly in a school choice program.

The mixed logit model allows the random component of the indirect utility function to interact with the product attributes – leading to random preference parameters for the attributes in the indirect utility function. The mixed logit can approximate any random utility model, given appropriate mixing distributions and explanatory variables (McFadden and Train 2000, Dagsvik 1994). This flexibility allows for realistic substitution patterns – allowing for credible estimates of demand elasticities and simulations – key in understanding implications of school choice for competition on quality.

This flexibility, however, comes at some cost. Because of the more complicated functional form, the likelihood function for the mixed logit does not have a closed form, and must be estimated by numerically integrating over the distribution for the random parameters using the method of maximum simulated likelihood to determine the final likelihood for each observation. In addition, as we will discuss further in the subsection on identification, in order to identify random parameters without relying heavily on the logit functional form, one needs repeated observations for individuals as well as variation in the choice set. The CMS data provide this important form of variation to help us pin down the means and variances of preference parameter distributions of interest.

Model and Estimation

Our model is based on a standard random utility framework. Let U_{ij} be the expected utility of individual *i* from attending school *j*. Individual *i* chooses the school *j* that maximizes his or her utility over all possible schools in the choice set. For the first choice, the individual chooses over the set of all available schools (denoted J_i^1), so that:

$$y_{ij}^1 = 1 \;\text{ iff } U_{ij} > U_{ik} \;\forall k \in J_i^1$$

$$y_{ij}^1 = 0 \;\text{ otherwise.}$$

The second and third choice (identified by y_{ij}^2 and y_{ij}^3) is made in a similar manner, except that the choice sets (denoted J_i^2 and J_i^3) exclude schools already chosen by individual *i*.

We assume that utility is a linear function of the observed student and school characteristics, X_{ij}, such as distance from home, average test scores, and racial composition of the school, plus an unobserved component, e_{ij}, that reflects unobserved idiosyncratic preference of student *i* for school *j*.

(1) $\quad U_{ij} = X_{ij} b_i + e_{ij}$

We assume that the unobservable component (e_{ij}) is distributed i.i.d. extreme value, which yields the usual logit form for the choice probabilities conditional on b_i.

Heterogeneity in individual preferences implies that the coefficients, b_i, in equation (1) will vary across individuals. We allow for this heterogeneity in two ways. First, we allow the parameters of equation (1) to vary randomly across individuals. We assume that $b \sim f(b \mid m_b, q)$, where $f(\cdot)$ is a mixing distribution, where m denotes the mean, and ? represents the other parameters describing the density function. Second, we separately estimate parameter distributions for students in each of the four main SES categories: White and Black by lunch subsidy status. This allows us to compare means and variances of preferences for school characteristics, such as average test scores, across the different socio-economic groups. In addition, we allow the coefficient on a school's average test score to vary with a student's baseline test score and family income by including interactions between these student characteristics and the school score in X.

In the specifications that are reported below, we assume that all the parameters are distributed *i.i.d.* normal or log normal. Hence in this specification we do not allow the random parameters to be drawn from multivariate distributions. We are currently incorporating covariance terms in our next round of results.

Given the specification above, the probability that individual *i* chooses schools (j^1,j^2, j^3) is given by:

$$P_i\left(j^1, j^2, j^3\right) = \Pr\left\{\left(U_{ij^1} > U_{ik} \forall k \in J_i^1\right) \cap \left(U_{ij^2} > U_{ik} \forall k \in J_i^2\right) \cap \left(U_{ij^3} > U_{ik} \forall k \in J_i^3\right)\right\}$$

$$(2) \qquad \int \prod_{c=1}^{3} \frac{e^{X_{ij^c} b}}{\sum_{k \in J_i^c} e^{X_{ik} b}} f\left(b \mid m, q\right) db$$

The term inside the integrand represents the probability of observing the three ranked choices conditional on the preference coefficients (ß): this is the product of three logit probabilities evaluated at b_i, corresponding to the probability of making each choice from among the remaining options.[7] This conditional probability is integrated over the distribution of b to yield the unconditional probability of observing the ranked choices.

These probabilities form the log-likelihood function:

$$(3) \qquad LL\left(X, m, q\right) = \sum_{i=1}^{N} \sum_{j=1}^{J_1} \sum_{k=1}^{J_2} \sum_{l=1}^{J_3} y_{ij}^1 y_{ik}^2 y_{il}^3 \ln\left(P_i\left(j, k, l\right)\right)$$

While equations (2) and (3) do not in general have a closed form solution, simulation methods can be used to generate draws of b from $f\left(\cdot\right)$ to numerically integrate over the distribution of preferences. Estimation was by the method of maximum simulated likelihood, using 100 draws of b from $f\left(\cdot\right)$ for each individual in the data set. The results were not sensitive to increases in the number of draws.

[7] For students submitting fewer than three choices, the likelihood is modified in an obvious way to reflect only the probability of the submitted choices.

V. Discussion of Identification and Descriptive Statistics

Before presenting the results from the mixed logit specification and the simulations based on those parameter estimates, we first discuss sources of identification in our data and research design, and provide some descriptive statistics summarizing the characteristics of parents and their choices.

The Proportion of Parents Listing More than One Choice

As discussed earlier, the availability of more than one choice for students in CMS will help identify the preference parameters. The choice form allowed parents to list up to three choices. Multiple choices are important for identifying variance of preferences in the population. Intuitively, when only a single (1^{st}) choice is observed for every individual, it is difficult to be sure whether an unexpected choice was the result of an unusual error term (e_{ij}) or unusual preferences by the individual (b_i) for some aspect of the choice. However, when an individual makes multiple choices that share a common attribute (e.g. high test scores) we can infer that the individual has strong preferences for that attribute, because independence of the error terms across choices would make observing such an event very unlikely in the absence of a strong preference.

The source of identification that comes from observing multiple choices on each individual is closely related to tests of the functional form assumption imposed by the choice model. For example, with only a single choice observed, the standard test of the IIA assumption in the logit model relies on the implication that this model, if correct, must yield the same coefficients when estimated on a limited choice set using only the sub-sample with choices from this choice set. If in fact there are random taste parameters for attributes, this will no longer be true. The sub-sample of individuals with choices from the restricted set will have different preferences than the rest of the sample. We can think of multiple choices for each individual behaving like the test of the IIA assumption: comparing the model estimated using only first choices versus only second choices. For second choices, the *same* individuals face different choice sets, so the distribution of preferences must be the same for first and second choices. Thus, the distribution of preferences that is estimated on the first choice must also fit the data on second choices – a type of out-of-sample fit where individuals face different choice sets.

As described earlier, we have choice response forms for 95% of students. However, since they were guaranteed a slot in their default school, many parents filled out only one choice. Presumably this occurred when their default school truly was their first choice. Overall, 35,754 students filled out only a first choice, 18,486 students listed only a first and second choice, and 46,246 students listed completely all three choices.

Table I and Table II describe the student body population in CMS, including the proportion of students in each socio-economic status category. Table III reports the number of choices submitted by each parent by race and free lunch eligibility. Among white students who were ineligible for the free lunch program, about half (51%) listed only one choice on their forms. Non-white and free lunch eligible students were much more likely to fill out all three choices. Among whites, the free lunch eligible students were considerably more likely to list three choices than the ineligible students (46 percent versus 29 percent). Moreover, among those who were ineligible for the free lunch program, non-white students were nearly twice as likely to list all three choices relative to white students (54 percent versus 29 percent).

There are at least two reasons why white students who are not free-lunch eligible were more likely to list only a single choice. First, the average quality of the default school for white free lunch-ineligible students is significantly higher. Table IV shows the mean test scores for the home schools by race and free lunch eligibility. The mean test scores of the default schools for white and free-lunch ineligible students are higher than those of the default schools of other groups. As a result, the more affluent students are less likely to find another school in their choice set that would dominate their default school. Hence they would be more likely to fill out only one choice. Moreover, the longer distances between schools in the suburban areas means that affluent parents are less likely to find any other school which would dominate their default school (that is, the next closest school may be quite far away and would have to be much better on some dimension to dominate).

If one's default school truly was one's first choice, there was no incentive to fill out the remaining slots on the choice form. Nevertheless, many parents specified multiple choices even if they listed their home school first. Table V reports the fraction of students in each SES category that listed their home school as their first choice.

According to Tables V and III, 64% of white, lunch-ineligible students chose their home school first, while only 51% listed only a first choice. This implies about a fifth of the white, lunch-eligible parents whose top choice was their home school actually provided additional rankings. About half of the black, lunch-eligible children whose top choice was their home school, provided some additional listings. Whatever their reasons for doing so, the availability of multiple choices even for those who listed their home school first will aid in the identification of the distribution of preferences for proximity versus school quality in the population.

Location of high test-score schools relative to population

Figure 1a presents a map of school locations and their test scores against the demographic characteristics of census block groups in Mecklenburg County. The neighborhoods are shaded a deeper blue when there was a higher proportion of the population that is African American in the block group in the 2000 U.S. Census of the Population. The shading of the school location markers are a function of the average test score in the school, darker shading identifying the schools with higher average test scores. This map is helpful in visualizing variation in data that contribute to identification of the random parameters in the mixed logit model. The high-scoring schools (some of them magnet programs created as part of the desegregation plan) are dispersed around the county, located in both urban and suburban areas, and in both minority and non-minority communities. Figure 1b shows an up-close example of school locations and demographics for a particular set of neighborhoods. This up-close picture measures roughly 4.5 miles across. These neighborhoods vary greatly in their racial composition, yet are roughly the same distance to the same set of schools, which vary substantially in average test scores. By examining the relative rankings of these three schools by students of various skill levels and socio-economic backgrounds, we can identify how the valuations of and trade-offs between the school characteristics of interest vary in the population.

Figure 2 shows a histogram of average test scores in CMS schools. There is substantial variation in the quality of CMS schools as measured by the average test scores of students in each school. Table VI presents the average travel distance (in miles) to the nearest 75th percentile school program by their eligibility for the federal Free and Reduced Price Lunch program ("lunch" identifies those were eligible for that program and "no-lunch" describes those who were not eligible). On average, students in all four categories have the same approximate travel distance to get to the nearest such high-scoring school. Such variation will help us identify preferences for school proximity and school quality across the socio-economic groups of interest.

Figure 3 shows the distribution of baseline test scores of students in CMS in Spring 2002 by race and lunch-recipient status. The histogram standardized test score by the mean and standard deviation in the district for students in each grade. Hence, a value of zero implies students who scored the average relative to all other students in their grade in the district. Based on the histograms, not only is CMS a racially and economically divers district, but within each category, there is substantial variation in student ability as measured by performance on standardized tests. While the mean of the distribution for white, non-lunch-recipient students is the highest of all four categories, there is a substantial fraction of underperforming students in this category, and there is a substantial density of high-performing students each of the other 3 categories. It is also true that within-school variation in performance is greater than across school variation in performance. This variation in student level-baseline achievement with and across schools, within and across socio-economic groups, will help identify the degree to which preferences for school quality vary with own academic ability.

V. Results

The specifications were estimated separately by race (white vs. non-white) and by free lunch eligibility (lunch vs. no-lunch). To keep the estimation tractable, we focused on a subset of five school characteristics and their interactions with student characteristics. The five school characteristics and their interactions with student characteristics were focused on because of the importance of the implications of these parameters for competition on quality in the context of the school choice debate. The specification includes distance from the student's residence to the school (measured in miles), school mean test score (from the Spring 2003 testing, measured in percentiles on the statewide exam), a standardized measure of a student's test score in their school the prior year (with a mean of zero and a standard deviation of one, the score reflects a student's ranking in their school in the previous year when ordered by standardized tests), the actual percent black in the school in Spring 2003 (with categories of 10-30 percent black, 30-60 percent black, >60 percent black and a left-out category of less than 10 percent black), and the median household income in the student's neighborhood for the student's race(measured in $1000's, using their census block group in 2000, and de-meaned with the countywide median of $51,000). The means and standard deviations of these variables are reported in Table VII.

The mixed logit parameter estimates are reported in Table VIII. All of the point estimates were statistically different from zero at the 1 percent level. To preserve space, the estimated standard errors are reported in Appendix Table I.

Travel Distance

Given strong priors that the coefficient on distance would be non-positive, we imposed a lognormal distribution on the preference coefficient for distance: $b_{distance} = -\exp(a)$, where α was assumed to be normally distributed. The coefficient on distance was the only coefficient for which we felt comfortable imposing an assumption regarding sign. For white, free lunch ineligible students, we estimated α to have a mean of -.751 and a standard deviation of .527 for white, free lunch-ineligible students. This implied that $b_{distance}$ had a mean of -.542 ($(-\exp(m_a + s_a^2/2))$ and a standard deviation of .307.

The results in Table Y suggest that all groups heavily weight travel distance in choosing schools. The ratio of the coefficient on distance to that on school test score was 9.8 and 6.2 for white and non-white free lunch-ineligible students with average baseline own-test scores and average income levels respectively. In other words, among the free lunch-ineligible students, parents of students with average performance on standardized test scores and average neighborhood income levels were willing trade-off 6 to 10 percentile points of school mean test scores for each mile of distance. The range of the school test score variable ran from 30 to 97 percentile points. In other words, a 6.8 mile difference in distance (67/9.8) would be enough to force the mean white, lunch ineligible parent to switch from the highest scoring to the lowest scoring school in the district. Non-white lunch-ineligible parents would require a slightly larger difference in distance, 10.8 miles, to make that switch (67/6.2).

Among lower-income free lunch-eligible students, the willingness to trade-off school performance for distance was even greater. Among the free lunch-eligible, white parents of children with average standardized test scores were willing to trade-off 27 points of mean school test score for every mile of distance (implying that they would be willing to switch from the highest-scoring to the lowest-scoring school for a 2.5 mile difference in distance); while non-white parents of children with average standardized test scores were willing to trade off 15 points per mile of distance.

School Test Scores

In estimating the effect of school test score on parental preferences, we included interactions between school score and the standardized value of one's test score within one's school during the prior school year. We also include an interaction between neighborhood income (income in the census block group measured in $1,000's and de-meaned by the countywide average of $51,000) and school score.[8]

The coefficient on the interaction between the standardized value of one's own test score and the school mean test score is positive - implying that those with higher test scores relative to their baseline peer group valued a school's mean test score more.[9] Among white, free lunch-ineligible students, being two standard deviations above the mean in one's school in the prior year meant that one valued school mean test scores about 20 percent more than those scoring at the mean of their school last year (2*.0056/.0552). Among black, free lunch-ineligible students, being two standard deviations above the mean implied that one valued school mean test scores about 27 percent more than those scoring at the mean of their school last year (2*.0082/.0611). Among white and black free lunch-eligible students, the interaction mattered even more:

[8] Since both income and the standardized value of own score are "de-meaned", the coefficient on the main effect of school score measures the value of school test score when both neighborhood income and standardized own score are at their respective means (equal to zero).
[9] To some unknown extent, it may also be capturing the fact that those who ranked well within their school in the prior year valued school test scores more. In future versions, we will be trying to distinguish between these two effects.

being two standard deviations above the mean increased the value of attending high-scoring school by 55 percent (2*.0053/.019) and 44 percent (2* .006/.027) respectively.[10]

Higher neighborhood income was strongly associated with higher valuations of school mean test scores. Among white, free lunch-ineligible parents, those in neighborhoods with a median income of $151,000 ($100,000 above the mean in the county) valued a school mean test score 268 percent more (a 168 percentage point increase) than those in neighborhoods with a median income of $51,000 (1+100*.000929/.0552). Among black, free lunch-ineligible parents, a $100,000 difference in median neighborhood income was associated with a 106 percentage point increase in the valuation of school test scores (100*.000648/.061). In addition, recall that the coefficient on school mean test score was considerably lower for the students who were free lunch eligible than for those who were free lunch ineligible. The difference in the mean coefficient in the two specifications for whites (.019 versus .055) is equivalent to the difference associated with a $38,000 increase in family income among free lunch-ineligible.

To summarize the importance of own baseline academic achievement and income on the value of a school's test scores relative to the value of proximity consider the comparison between a student with the average baseline test score and average neighborhood income level. For white and black free-lunch ineligible students with average own-test scores and average neighborhood income, the ratio of disutility of distance to preference for school test scores is 9.8 and 6.2 respectively. The ratios fall to 4.8 and 3.4 respectively for white and black students with baseline scores that are 2 standard deviations above the mean and who live in neighborhoods where the median income is $50,000 above the mean. We can think of this as increasing the geographic radius of schools that compete with each other on quality for students with high-achievement levels in their own baseline scores and higher-than-average neighborhood income levels.

[10] Although the estimate of the mean coefficient on the interaction term was similar across lunch status, the main coefficient on test scores was considerably smaller for low-income, free lunch-eligible students.

Racial Composition

Among those who were free lunch eligible as well as among those who were ineligible for the free lunch program, there were relatively modest racial differences in the coefficients on distance and school mean test score. However, there were large differences between the races in the valuation of racial composition at the school of choice. The excluded category in these specifications is schools that were less than 10 percent African American. Therefore, a positive coefficient on the indicator for schools that were 10 to 30 percent black for white free lunch ineligible students implies that they actually preferred racially integrated schools to schools with similar test scores that were 0 to 9 percent black. Indeed, for white free lunch ineligible students, the optimum racial composition was 10 to 30 percent African American. (Among white free lunch-ineligible parents, the only category less preferred than 0 to 9 percent white schools were schools that were more than 60 percent black.) African American parents also preferred to attend racially integrated schools. However, they preferred to attend schools that were more than 30 percent black. (There was a small difference in the mean parameter estimate between schools that were 30 to 60 percent black and more than 60 percent black.)

These results are quite consistent with an earlier literature highlighting racial differences in stated preferences regarding the racial composition of neighborhoods. That literature (well surveyed in Armor (1995)) reported that both whites and blacks preferred to live in integrated neighborhoods. However, blacks and whites disagreed on the optimal amount of integration—whites preferring neighborhoods that were 10 to 30 percent black and blacks preferring neighborhoods that were roughly 50 percent black. A number of authors (e.g. Farley et. al. (1978) and Schelling (1971)) have speculated about the implications of these preferences for equilibrium levels of integration. Even though both blacks and whites prefer integration, the equilibrium outcome may yield more segregated schools than either would prefer, given the differences in preferences.

While the focus of this current paper is on preferences for school quality rather than preferences for racial composition, failing to account for racial differences in preferences regarding school racial composition can lead to misleading inferences regarding preferences for school quality. For instance, if one were to leave out racial composition of the school, blacks and whites appear to have very different preferences regarding the mean test score of the school. Figure 4 plots average school test score versus percent black in the school. Because school test score is positively correlated with the percent black in the school (with a correlation coefficient of approximately .65), failing to deal with explicit racial preferences leads us to overstate whites' valuation of school scores (since they prefer schools with below-average black enrollments) and understate black student valuation of school scores (since they prefer schools with above-average black enrollment). As we noted above, holding school racial composition constant, whites and blacks attach similar weights to school scores on the margin. However, failing to account for school racial composition would have led to the false conclusion that whites care much more about school scores than black students.

VI. Simulations

In the discussion of the results above, we focused primarily on the mean weight attached to various school attributes. However, the aggregate response to any policy change will depend not only on the mean parameter estimate, but also on the variance or distribution of that parameter in the population. As noted in the introduction, a key issue in the policy debate over school choice is the elasticity of demand with respect to school test scores. In order to shed some light on this question, we took each school individually, added 10 percentile points to its mean school score holding all else equal, and simulated the change in the number of students listing that school as a first choice.

Figure 5 plots the change in number of students listing a school as a first choice by the school's original average score (each point in the figure is the result of a simulation for a different school). The demand response is quite different for schools that were originally high and low-scoring. The upward sloping relationship implies that the demand response is greatest among schools that were already high scoring. This result reflects the parameter estimates in the mixed logit model. More affluent parents of higher-scoring kids value average test scores highly, and are willing to drive their child to higher average score schools. These parents are both i) likely to only consider high scoring schools for their children, and ii) would be willing to drive further in response to an increase in score at another high scoring school. These results imply that the incentives to focus on student performance are larger for higher performing schools, since, schools above a critical performance level compete intensely on quality for the quality-elastic segment of the population.

Figures 6 and 7 plot differences in mean characteristics between the marginal students (those who are drawn in by the 10 percentile point test score rise) and students who previously enrolled in each school. The incentive for any school to improve its performance would be dampened if, in doing so, they were swamped by lower-performing and or lower-income students, who would bring down mean performance and potentially be more costly to educate. Figure 6 graph reports differences between marginal and average students in the percentage eligible for the free lunch program; Figure 7 reports differences in mean test scores. The points above the 45° degree line in the Figure 6 indicate schools where a higher proportion of the marginal students qualified for the free lunch program than the average student. It is evident in graph that the marginal students had lower free lunch eligibility rates than the students already enrolled. In other words, the marginal students were more affluent, not less affluent that the students already enrolled in most schools. Figure 7 reports differences in mean test score between the marginal student and the average student previously enrolled in the school. The fact that most points were above the 45° line implies that the marginal students, on average, were higher performing than the students already enrolled.

The implications of these simulations are very interesting for school choice policy design. On one hand, they suggest that the absolute enrollment responses to improvements in performance are small at schools that start out low-performing. The enrollment responses are much larger at schools that start out higher performing—suggesting that market forces may lean toward greater vertical separation on test scores. In the extreme, this pattern would lead to a two tiered system, with the best schools competing heavily on the academic dimension for geographically dispersed students who highly value academic quality, while low scoring schools faced little incentives to improve scores – acting as local monopolists over students in their neighborhood whose parents' preferences over distance and school scores make them highly inelastic on school quality. On the other hand, the marginal students who are brought in when a low-scoring school improves tend to have much higher baseline performance than the students already enrolled. The net effect on the incentives for teachers and principals at these schools to improve their performance will depend on the financing scheme - how much additional funding per student they receive perhaps as a function of student quality- as well as the intangible rewards that come from drawing higher scoring students to one's school.

VII. Conclusion

This paper uses student-level data from a school choice program in Mecklenburg County, North Carolina to estimate a mixed logit model of demand for schools. The mixed logit demand model allows us to estimate the heterogeneity of preferences in the population, which is important for estimating substitution patterns and demand elasticities in response to changes in school characteristics. The results in this paper are still tentative, and the school characteristics we have considered to this point are far from an exhaustive list. The substantive findings should be treated as preliminary. Nevertheless, the resulting estimates illustrate some interesting features of the demand facing public schools in a choice environment.

Our preliminary results indicate that i) parents value proximity highly, ii) the preference attached to a school's mean test score is substantially lower for low-income students (those eligible for Free or Reduced Price Lunch or living in a low income neighborhood), and iii) preferences for school test scores increase significantly with the student's own academic ability. We also find considerable heterogeneity in preferences across individuals even after controlling for income and ability. In addition, we find that, African Americans and whites have very similar preferences over school test scores and distance conditional on family income and the student's baseline test score, once we control for preferences over racial composition of schools (which differ significantly by race).

Given our demand estimates, we simulate the elasticity of demand for each school with respect to mean test scores in the school. We find that demand at high-performing schools is more responsive to increases in mean test scores than demand at low-performing schools. This result is generated from the fact that students who value academic achievement choose high-test-score schools, and are much more willing to switch schools in response to an increase in test scores at another school. Hence, these high-performing schools would have a stronger incentive to compete for these elastic students by raising their academic performance. The less elastic students will remain to be served by the lower-performing schools. The disparate competitive pressure across high and low performing schools may result in a two tiered system. In differentiated product markets, this type of segmentation is a common phenomenon. When a low-price firm (e.g. WalMart) enters the market, it draws price elastic customers from a broad geographic region. Local stores then are left to serve the residual market of very local, price-inelastic customers. Hence their best strategy might be to increase their price to these local inelastic customers, letting WalMart serve the elastic customers from a broader geographic market.

However, our simulation results also show that, while lower-performing schools draw fewer students in response to an increase in average test scores, the students they do draw have higher average academic performance and are less likely to be poor than the school's average student. These simulation results may have interesting implications for school choice policy design. The net effect on the incentives for teachers and principals at these schools to improve their performance will depend on the financing scheme - how much additional funding per student they receive, perhaps as a function of student income or ability - as well as the intangible rewards that come from drawing higher scoring students to one's school.

References:

Abdulkadiroglu, Atila and Tayfun Sonmez "School Choice: A Mechanism Design Approach", *American Economic Review*, 93(3):729-747, June 2003.

Armor, David J. *Forced Justice: School Desegregation and the Law* (Oxford: Oxford University Press, 1995).

Armor, David J. and Peiser, B.M. "Interdistrict Choice in Massachusetts" in Paul Peterson and Bryan Hassel (eds.) *Learning from School Choice.* (Washington, DC: Brookings Institution Press, 1998).

Berry, Steven, J. Levinsohn and A. Pakes (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market," *Journal of Political Economy*, 112(1), 68–104.

Berry, Steven, J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 60(4), 889–917.

Bridge, R. Gary and Julie Blackman *A Study of Alternatives in American Education Vol. IV: Family Choice in Schooling* (Santa Monica, CA: Rand, 1978).

Carnegie Foundation *School Choice* (Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching, 1992).

Clotfelter, Charles T. *After Brown: The Rise and Retreat of School Desegregation* (Princeton, NJ: Princeton University Press, 2004).

Convey, J.J. "Parental Choice of Catholic schools as a function of religion, race and family income" Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, (April, 1986) (ERIC Document Reproducton Service No. ED 269 542) (as cited in Maddaus (1990)).

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

Dagsvik, John, "Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes," *Econometrica*, 62(5):1179-1205, September 1994.

Dahl, Gordon, "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, 70(6):2367-2420, November 2002.

Darling-Hammond, Linda and S.N. Kirby *Tuition Tax Deductions and Parent School Choice: A Case Study of Minnesota* (Santa Monica: Rand, December 1985).

Dubin, J. and D. McFadden, “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption,” *Econometrica*, 52:345-362, 1984.

Farley, Reynolds, H. Schuman, S. Bianchi, D. Colasanto, and S. Hatchett “Chocolate City, Vanilla Suburbs: Will the Trend Toward Racially Separate Communities Continue?” *Social Science Research* (1978) Vol. 7, pp. 319-344.

Fossey, Richard “Open Enrollment in Massachusetts: Why Families Choose” *Educational Evaluation and Policy Analysis* (1994) Vol. 16, No. 3, pp. 320-334.

Glazerman, Steven, “Determinants and Consequences of Parental School Choice” Unpublished working paper, University of Chicago, Harris School of Public Policy, (December 21, 1997).

Goldring, Ellen and P. Bauch “Parental Involvement and school responsiveness: Facilitating the Home-School Connection in Schools of Choice” Educational Evaluation and Policy Analysis, (1995) Vol. 17, pp. 1-22.

Greene, Jay, William Howell and Paul Peterson “Lessons frm the Cleveland Scholarship Program” in Paul Peterson and Brian Hassel (eds.) Learning from School Choice, (Washington, DC: Brookings Institution, 1998).

Henig, J. “Choice in Public Schools: An Analysis of Transfer Requests Among Magnet Schools” *Social Science Quarterly* (1990) Vol. 71, No. 1, p. 76.

Hoxby, Caroline “Does Competition among Public Schools Benefit Students and Taxpayers?” *American Economic Review* (2000) Vol. 90, No. 5, pp. 1209-1238.

Kleitz, B., G.R. Weiher, K. Tedin, and R. Matland “Choices, Charter Schools and Household Preferences” *Social Science Quarterly* (1998) Vol. 81, No. 3, pp. 846-854.

Krueger, Alan B. and Pei Zhu “Another Look at the New York City School Voucher Experiment” *National Bureau of Economic Research Working Paper* No. 9418, December 2002.

Maddaus, John “Parental Choice of School: What Parents Think and Do” *Review of Research in Education*, Vol. 16. (1990), pp. 267-295.

Manski, Charles F. and David A. Wise *College Choice in America* (Cambridge, MA: Harvard University Press, 1983).

Mayer, Daniel P., Paul E. Peterson, David E. Myers, Christina Clark Tuttle, and William G. Howell, “School Choice in New York City After Three Years: An Evaluation of the School Choice Scholarships Program," Washington, DC: Mathematica Policy Research, Inc., 2002.

McFadden, D. and K. Train (2000). "Mixed Multinomial Logit Models for Discrete Response." *Journal of Applied Econometrics*, 15(5): 447-470.

Nault, R.L. and S. Uchitelle "School choice in the public sector: A case study of parental decision-making" in M.E. Manley-Casimir (ed.) *Family Choice in Schooling: Issues and Dilemmas* (Lexington, MA: Heath, 1982).

Nelson, K.O. "Reasons given by Anglo/Hispanic parents/guardians for choosing a Catholic high school in the Southwestern United States" (November, 1988) Dissertation Abstracts International, 49, 1049-A (University Microfilms No. 8814262) (as cited in Maddaus (1990)).

Schelling, Thomas. "Dynamic Models of Segregation" Journal of Mathematical Sociology (1971) Vol. 1, pp. 143-186.

Schneider, Mark and Jack Buckley "What Do Parents Want From Schools?: Evidence from the Internet" *Educational Evaluation and Policy Analysis* (2002) Vol. 24, No. 2., pp. 133-144.

Schneider, M., M. Marschall, P. Teske and C. Roch "School Choice and Culture Wars in the Classroom: What Different Parents Seek from Education" *Social Science Quarterly* (1998) Vol 79, No. 3, pp. 489-501.

Train, Kenneth E. Discrete Choice Methods with Simulation. Cambridge University Press (2003).

Van Dunk, Emily and Anneliese Dickman, "School Choice and Accountability: An Examination of Informed Consumers in Different Choice Programs" *Urban Affairs Review* (2002) Vol. 37, No. 6, pp. 844-856.

Vanourek, G., B.V. Manno, and C.E. Finn, "Charter Schools as Seen by Students, Teachers and Parents" in Paul Peterson and Bryan Hassel (eds.) *Learning from School Choice.* (Washington, DC: Brookings Institution Press, 1998).

Weiher, Gregory .R. and Kent L. Tedin, "Does Choice Lead to Racially Distinctive Schools? Charter Schools and Household Preferences" *Journal of Policy Analysis and Management* (2002) Vol. 21, No. 1, pp. 79-92.

Williams, M.F., K.S. Hancher and A. Hutner *Parents and School Choice: A Household Survey* (Washington DC: U.S. Department of Education, 1983).

Table I: Percent of Student Population by Ethnic Group and Grade

	Ethnic Group					
Grade	Percent White	Percent Black	Percent Hispanic	Percent Asian	Percent Other	Total Number Students
Pre-School	14.45	59.64	15.05	4.69	6.17	3,793
Kindergarten	42.42	39.79	10.34	4.31	3.14	8,489
First	42.59	40.63	10.01	3.99	2.77	8,648
Second	43.21	41.01	8.82	4.25	2.70	8,658
Third	42.19	43.65	7.99	3.93	2.22	8,769
Fourth	42.57	44.20	7.25	4.07	1.92	8,819
Fifth	41.39	45.44	7.90	3.94	1.33	8,872
Sixth	42.64	45.67	6.40	4.18	1.11	8,542
Seventh	42.87	45.18	6.38	4.57	1.00	8,447
Eighth	43.44	44.85	6.17	4.40	1.15	8,188
Ninth	43.61	45.81	5.48	4.30	0.79	8,952
Tenth	49.48	39.81	4.85	4.88	0.98	7,136
Eleventh	55.63	34.24	4.39	5.16	0.58	5,330
Twelfth	53.35	37.62	3.37	5.15	0.51	5,516
Total	43.33	43.17	7.37	4.35	1.79	108,159

Table II: Free and Reduced Lunch Status by Ethnic Group

	Ethnicity					
Lunch Status	Black	Asian	Hispanic	White	Other	Total Number Students
None	37.47	63.6	29.08	90.5	54.61	66,270
Free or Reduced	62.53	36.4	70.92	9.5	45.39	41,889
Total						108,159

Table III: Percent of Students Submitting Full Set of Choices by SES Group

		1^{st} Choice Only	1^{st} and 2^{nd} Choice Only	All Three Choices
White	**No-Lunch**	0.5123	0.1985	0.2892
	Lunch	0.3311	0.2057	0.4631
Non-White	**No-Lunch**	0.2768	0.1778	0.5454
	Lunch	0.2065	0.1664	0.6271

Table IV: Average Quality of Home School by Socio-economic Status (Measured by Average 2002 End of Grade and End of Course Percentile Scores for Students in Each School)

		Reading		Math	
		Mean	**St.Dev.**	**Mean**	**St.Dev.**
White	**No – Lunch**	58.7431	10.7440	59.3801	12.6574
	Lunch	49.3351	10.5111	49.6225	10.4972
Non-white	**No – Lunch**	50.8271	10.4711	50.0099	11.0127
	Lunch	43.3369	10.6322	44.5727	8.6046

Table V: Fraction of Students Whose 1^{st} Choice was their Home School, by Socio-economic Status

		Chose Home School First	
		Yes	**No**
White	**No-Lunch**	0.6443	0.3557
	Lunch	0.5140	0.4860
Non-White	**No-Lunch**	0.4251	0.5749
	Lunch	0.3827	0.6173

Figure 1a: Thematic Map of Charlotte Mecklenburg County with Census Block Groups by Race and School Location by Average Test Score

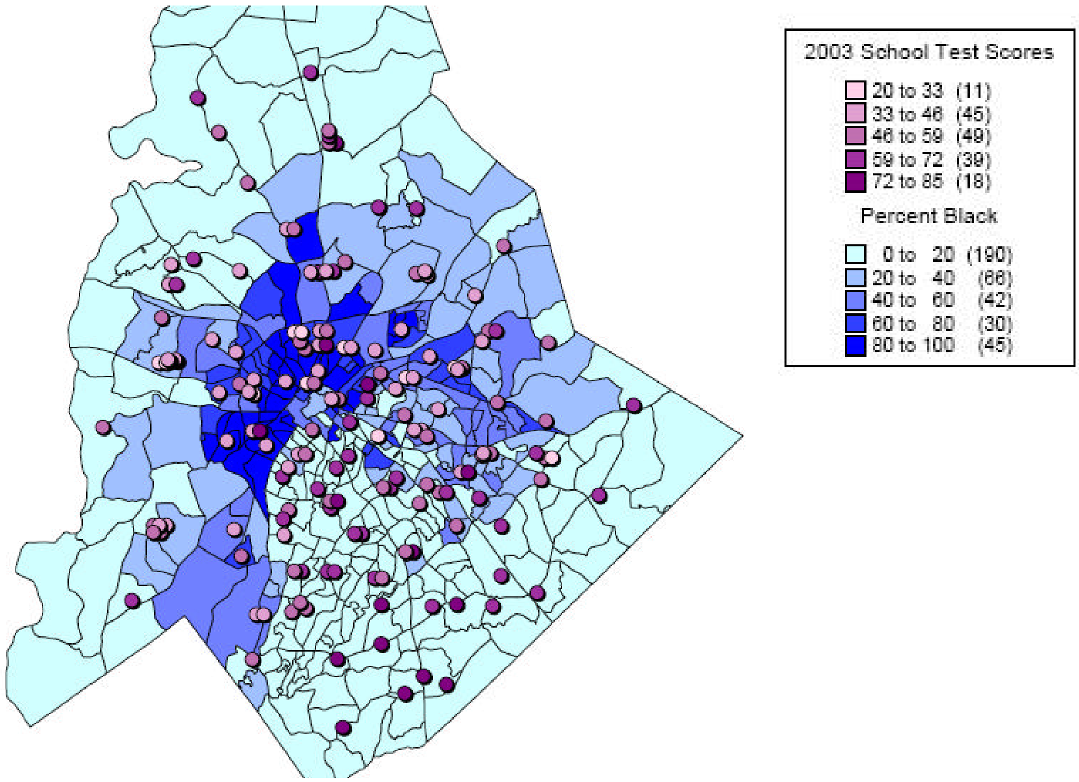


Figure 1b: Close View of Block Groups and School Choices by Average Test Score

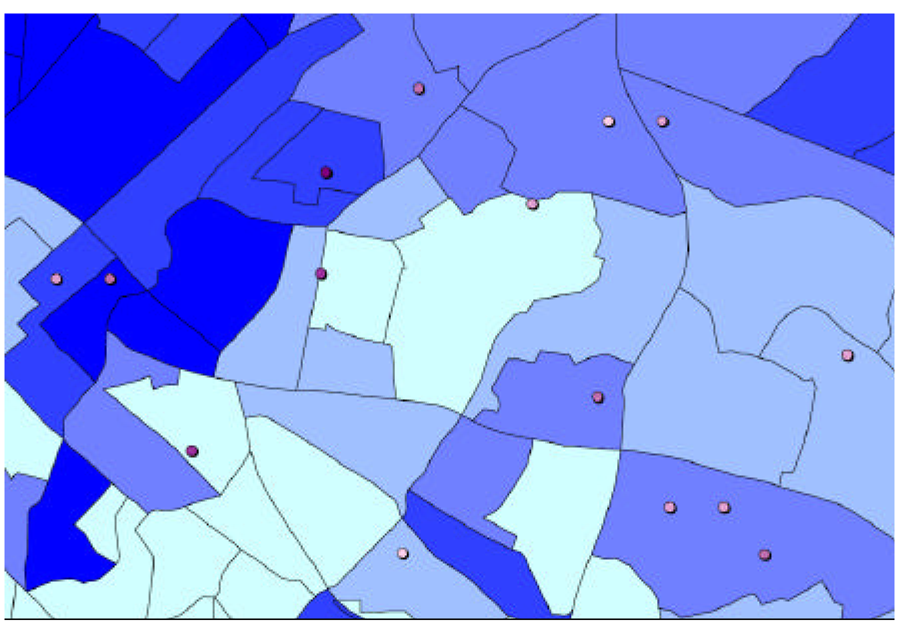


Figure 2: Distribution of Average Percentile Score for End of Grade 2002 Reading and Math Exam for School Programs in CMS.

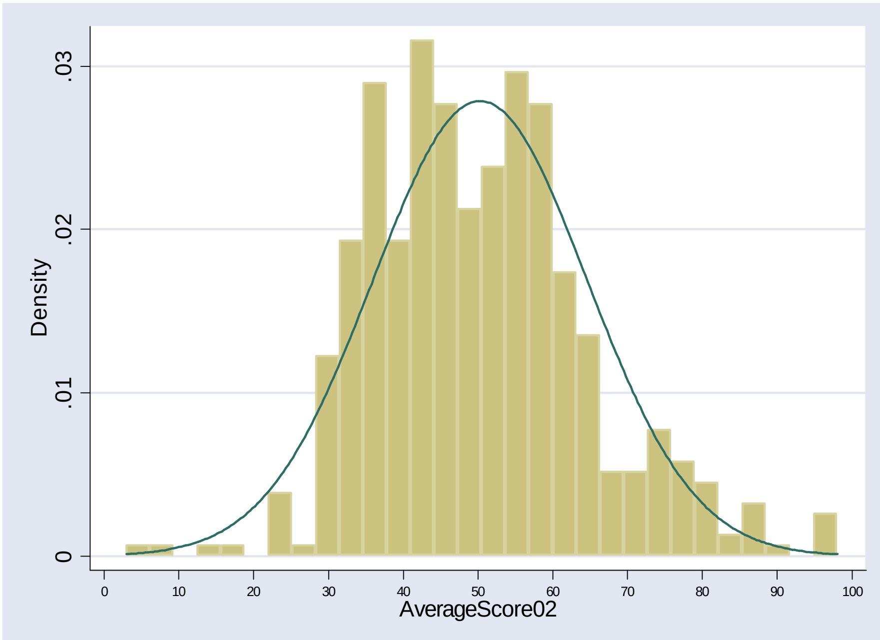
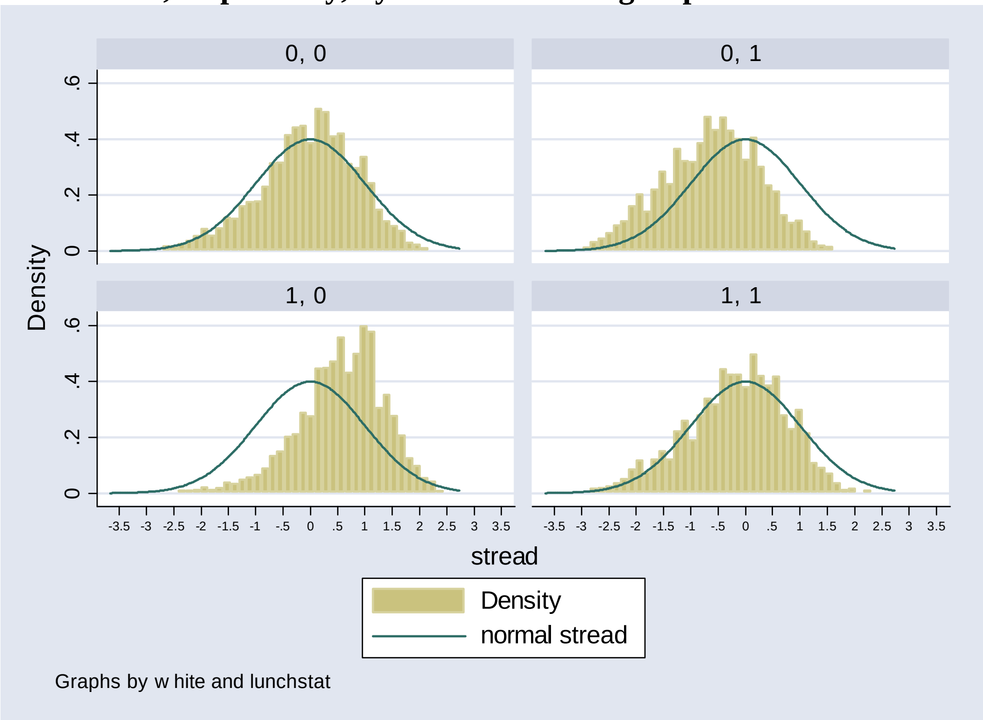


Table VI: Average Travel Distance (in miles) to a 75th Percentile Ranked School by Socio-economics Status

		N	Mean	St.Dev	Min	Max
White	**No – Lunch**	42392	2.5664	1.6134	0.0019	9.3975
	Lunch	3755	2.4523	1.4359	0.0094	7.9419
Non-white	**No-Lunch**	22014	2.6616	1.4828	0.0038	7.9300
	Lunch	32209	2.1272	1.2000	0.0010	7.6356

Figure 3: Distribution of Standardized Scale Score for End of Grade Reading and Math Exam, respectively, by Socio-economic group.



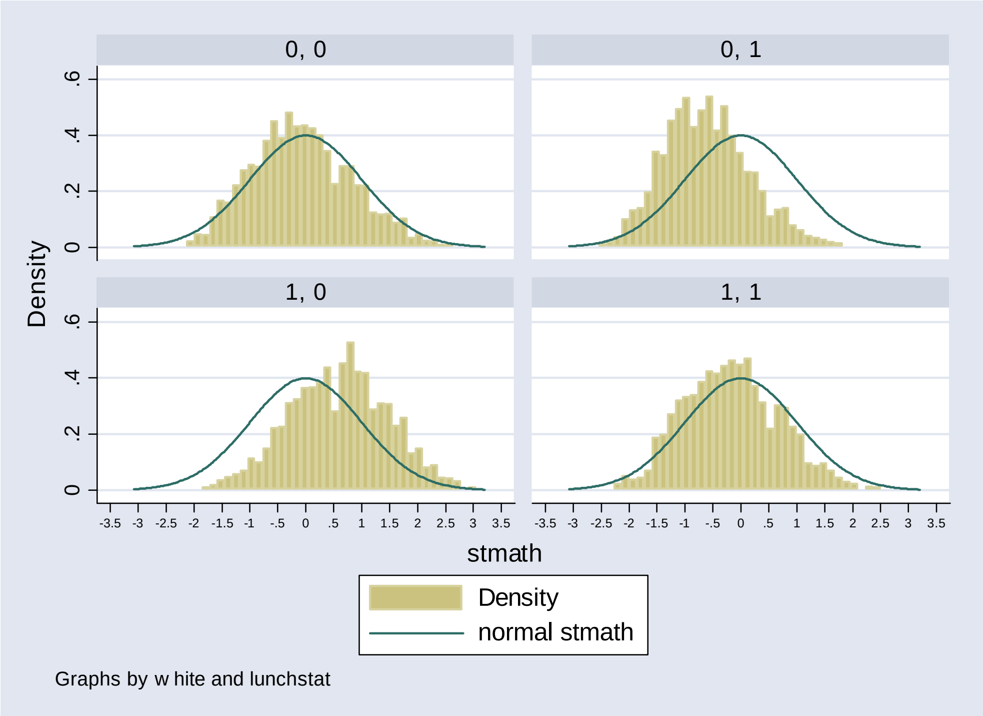


Table VII: Explanatory Variable Definitions and Summary Statistics

Variable	Description
Distance	Driving distance from student *i* to school *j* calculated using MapInfo with Census Tiger Line files.
School Score	Average percentile score of students in school *j* on math and reading End of Grade exams for 2002-2003 school year.
Z-Score	Student *i*'s percentile score on End of Grade exams in baseline year 2001-2002 standardized by test scores of students in her school.
Income	The median household income reported in the 2000 Census for households of student i's race in student i's block group. Income is demeaned by the county-wide average of approximately $51,000 and is reported in thousands of dollars.
Percent Black	The percent of students in school *j* who are black according to 2002-2003 school year administrative data.

Summary Statistics Using First Choice Data					
Variable	**Obs.**	**Mean**	**Std. Dev.**	**Min**	**Max**
Distance	3593362	12.9454	6.73231	0.001	42.4069
Last-year School	3593362	0.014859	0.120987	0	1
School Score	3593362	50.85243	12.37414	30.21118	97
Z-Score	3593362	0.018329	0.982778	-7.61337	6.982959
Z-Score*School-Score	3593362	0.946546	51.44977	-583.966	568.3639
Income	3593362	4.573575	27.59766	-48.501	149.001
Income*School-Score	3593362	232.437	1445.469	-3947.64	13896.3
Percent Black 10-30%	3593362	0.160413	0.366988	0	1
Percent Black 30-60%	3593362	0.384632	0.486508	0	1
Percent Black >60%	3593362	0.413619	0.492482	0	1

Table VIII: Estimates from Mixed Logit Model

		Parameter Estimates*			
		No Lunch		Lunch	
Variable	**Preference Parameter**	**White**	**Black**	**White**	**Black**
Distance**	**Mean (normal)**	-0.751720	-1.088200	-0.792370	-1.002800
	Std. Dev. (normal)	0.527400	0.477940	0.515680	0.613300
	Mean (lognormal)	-0.541916	-0.377575	-0.517158	-0.442758
	Std. Dev. (lognormal)	0.306882	0.191271	0.285440	0.299197
Last-year School	**Mean**	3.071700	2.666400	2.360600	2.314300
	Std. Dev.	1.495700	1.172800	1.100700	1.462300
School Score	**Mean**	0.055211	0.061113	0.019162	0.029678
	Std. Dev.	0.015422	0.000630	0.007907	0.000292
Z-Score * School Score	**Mean**	0.005631	0.008201	0.005261	0.006151
	Std. Dev.	0.005705	0.006444	0.001503	0.008539
Income*School Score	**Mean**	0.000929	0.000648	--	--
	Std. Dev.	0.000535	0.000419	--	--
Percent Black 10-30%	**Mean**	0.747710	0.783100	-0.118100	-0.063676
	Std. Dev.	0.293140	0.256130	0.934270	0.374200
Percent Black 30-60%	**Mean**	0.502160	1.141300	-0.292000	0.067487
	Std. Dev.	0.124240	0.066275	0.001501	0.068906
Percent Black >60%	**Mean**	-0.318100	1.210400	-0.745940	0.143740
	Std. Dev.	0.769640	0.021822	0.044500	0.038462

* All estimates are significant at the 1% level or higher

** Distribution of preference on distance follows a log normal distribution.

Figures 4: Scatter Plot of Average Percentile Score of Students in a School Program versus Percent Minority in the School Program

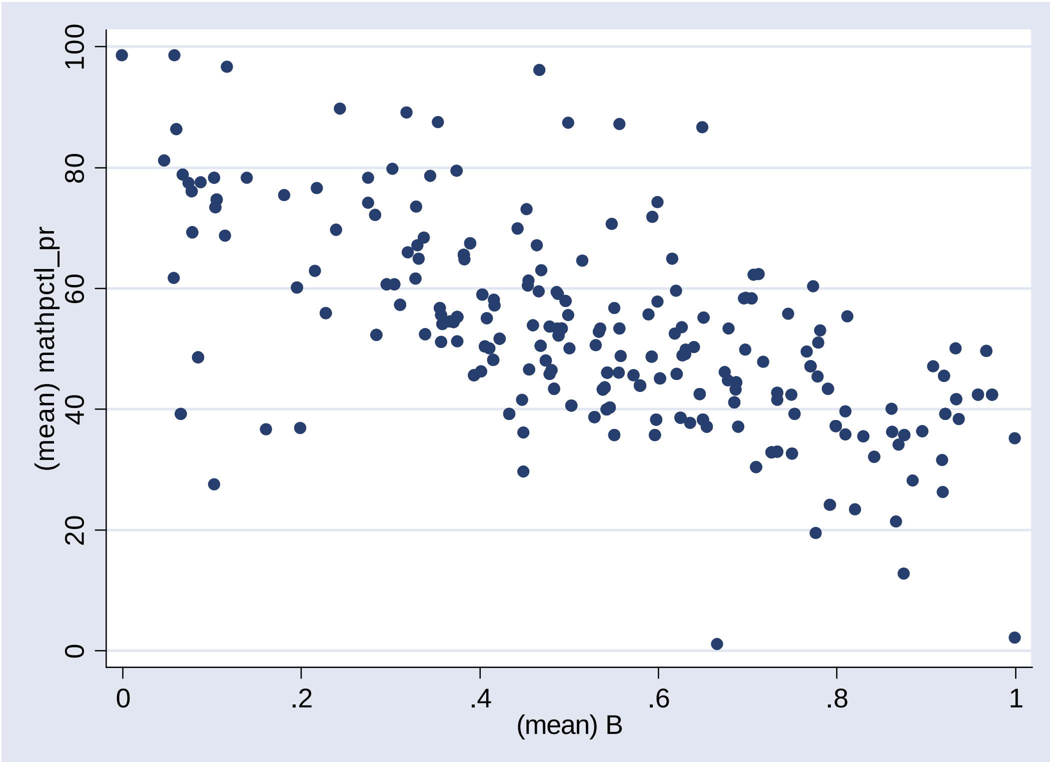


Figure 5: Elementary Schools: Simulated Change in Number of Students Choosing School *j* when the Average Percentile Score at School *j* increase by 10 points, all else equal.

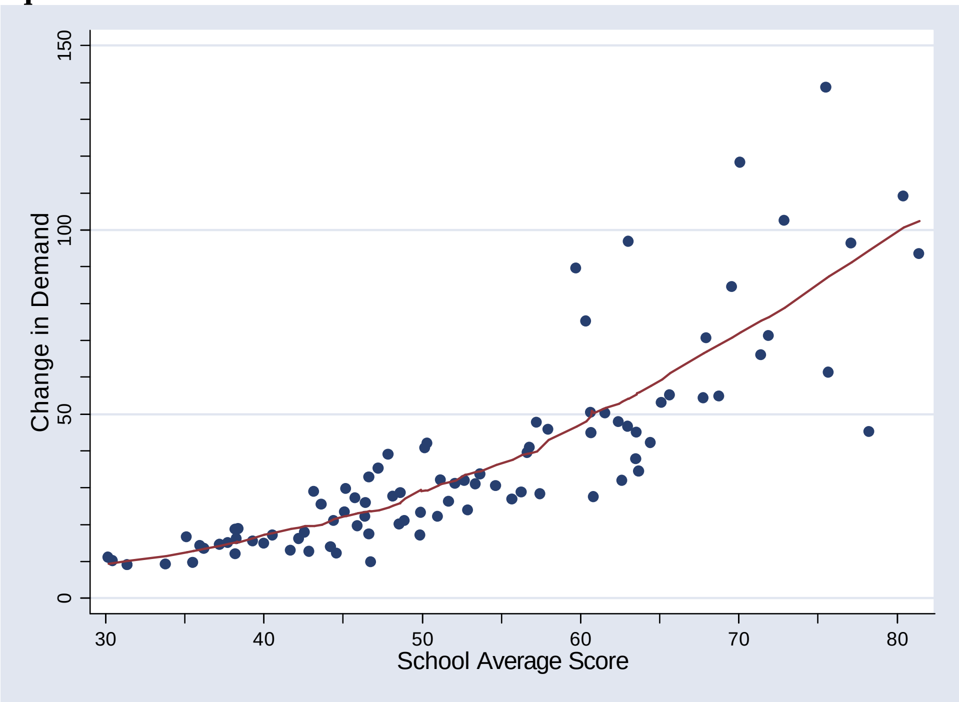


Figure 6: Middle Schools: Simulated Change in Number of Students Choosing School *j* when the Average Percentile Score at School *j* increase by 10 points, all else equal.

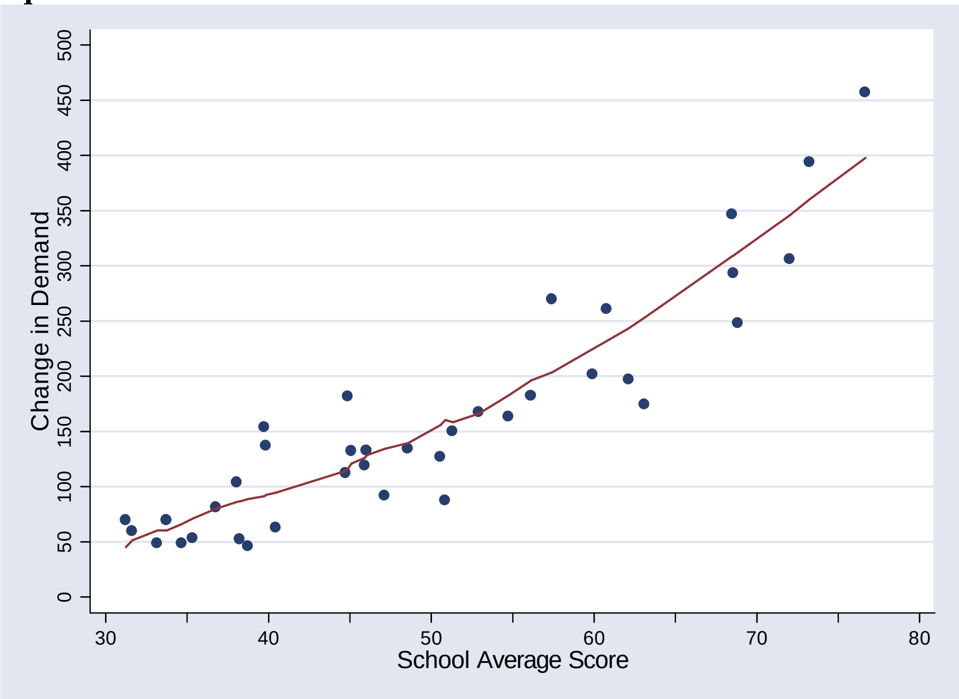


Figure 7: Percent of the Additional Students who Choose School *j* in Response to a 10 point Increase in Average Percentile Score at School *j* who qualify for Free Lunch.

Figure 8: Average 2002 Percentile Score for the Additional Students who Choose School *j* in Response to a 10 point Increase in Ave. Score at School *j*.