



**Yale University**  
Department of Economics

**Yale Working Papers on Economic Applications and Policy**

Yale University  
P.O. Box 208268

New Haven, CT 06520-8268

**DISCUSSION PAPER NO. 10**

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

Justine S. Hastings  
Yale University & NBER  
[justine.hastings@yale.edu](mailto:justine.hastings@yale.edu)

— Thomas J. Kane —  
UCLA & NBER

Douglas O. Staiger  
Dartmouth College & NBER

November 2005

\*We thank the Charlotte-Mecklenburg School Board, Office of the Superintendent, as well as the excellent school district staff for making this project possible. We would also like to thank Joseph Altonji, Patrick Bayer, Steven Berry, Philip Haile, Fabian Lange, Sharon Oster, Miguel Villas-Boas, Sofia Villas-Boas, and seminar participants at Yale University, University of California at Berkeley, and University of Chicago for helpful comments. We gratefully acknowledge Sean Hundtofte, Jeffrey Weinstein, and Orkun Sahmali for providing outstanding research assistance. The project is funded by grants from the Institution for Social and Policy Studies at Yale University and the U.S. Department of Education.

This paper can be downloaded without charge from the Social Science Research Network electronic library at: <http://ssrn.com/abstract=885537>

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

## Abstract

This paper uses data from the implementation of a district-wide public school choice plan in Mecklenburg County, North Carolina to estimate preferences for school characteristics and examine the implications for the local educational market. We use parental rankings of their top three choices of schools matched with student demographic and test score data to estimate a mixed-logit discrete choice demand model for schools. We find that parents value proximity highly and the preference attached to a school's mean test score increases with student's income and own academic ability. We also find considerable heterogeneity in preferences even after controlling for income, academic achievement and race, with strong negative correlations between preferences for academics and school proximity. Simulations of parental responses to test score improvements at a school suggest that the demand response at high-performing schools would be larger than the response at low-performing schools, leading to disparate demand-side pressure to improve performance under school choice. Moreover, given the greater sensitivity to school test scores among high-income and high-scoring youth, the marginal students attracted when a school improves performance would typically raise the average income and baseline test score at the school.

JEL CLASSIFICATIONS: I0, I20, I28

## **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 market outcome generated by any choice plan will depend in large part on parents' preferences over school characteristics, and how those preferences vary in the population. 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 demand-side 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 choice impact school quality? 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 the top students in the higher quality schools. At the other end of the debate, critics of choice plans predict that public school choice will result in "vertical separation" – an equilibrium outcome in which the top students abandon under-performing schools. Once the parents with high elasticities depart from their neighborhood schools, such schools would face little pressure from the remaining students to improve academic achievement.

These disparate predictions are generated by two very different beliefs about the nature of preferences. The 'tide that lifts all boats' outcome will occur if all parents value school quality more than other school characteristics, so that all schools must provide high quality instruction or parents will go elsewhere. Alternatively, if parents have very heterogeneous preferences for school quality, 'vertical separation' may occur. For instance, high-quality schools may compete intensely in a city-wide market for students with strong preferences for school quality, while neighborhood schools may serve the remaining students with weak preferences for school quality. These neighborhood schools will have a 'local monopoly' with little competitive pressure to improve quality. In addition, if these local monopolist schools serve less advantaged communities, vertical separation may increase disparities in public education quality across economic classes.

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 a variety of school characteristics. CMS introduced public school choice in the fall of 2002, after a race-based bussing plan was terminated by court order. Under the choice plan, parents in the district were asked to submit their top three choices of schools for their children. Approximately 95% of the students submitted choices for the choice plan. We analyze 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 a mixed logit discrete choice model, we estimate the distribution of preferences over school academic quality, school proximity, and school racial composition in a random coefficients discrete choice model of demand. We allow preference distributions to vary with student demographics and academic ability. The data from CMS and the school choice policy intervention are uniquely suited to estimating parental preferences for several reasons. First, parents were asked to submit their top three choices. The multiple responses create variation in the choice set by effectively removing the prior chosen school from the subsequent choice set. This choice-set variation allows us to estimate the distribution of preferences for school characteristics from observed substitution patterns for each individual. 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 the mean and variance of preferences in the population.

In addition to the multiple observations on parent's school choices, the data and policy experiment provide rich independent variation in key variables of interest. First, there is substantial variation in student characteristics. Charlotte-Mecklenburg school district is a large school district (approximately 110,000 students) with a diverse student population in terms of race and income. Second, the court order ending the race-based bussing plan forced the district to redraw school boundaries. Many neighborhoods which had been bussed to a distant school were re-assigned to a home school in their

neighborhood. Approximately fifty percent of parents in the county were assigned to a home school different from the school they would have been assigned the year before. In other words, the school they were assigned as their default school under the choice plan was often not the same school they would have anticipated when they chose their residence. When parents choose their residence based upon known school assignments, it is difficult to distinguish between a preference for proximity and a preference for other unobserved attributes of the neighborhood school. Because the implementation of the choice plan coincided with changes in school assignments, we are better able to identify the preference for proximity.

Our results indicate that 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). Moreover, the preference for a school with high test scores is increasing in the student's baseline academic ability and neighborhood income level. In addition, after accounting for racial differences in preferences for racial composition of schools, African Americans and whites have similar preferences over school test scores. Our results also indicate that parents value proximity highly and that the value of proximity is strongly negatively correlated with the preference for test scores. The value of proximity will be overstated if parents choose their residential location based on unobserved traits of local schools. Given the large scale redistricting that occurred with the implementation of the choice plan, we test whether the demand estimates are biased by endogenous residential location. To do so, we estimate the demand model using only the redistricted sub-sample of students and find similar results. This suggests that the strong preference for nearby schools is not an artifact of parents locating in the school zone of their preferred school.

With demand estimates from the mixed logit model, we calculate the elasticity of demand for each school with respect to school mean test score. In particular, we simulate the estimated increase in number of students choosing each school if it were to increase the academic performance of its students by a given amount, all else equal. We find that demand at high-performing schools is much more responsive to increases in academic performance, than demand at low-performing schools. Moreover, because high-income and high-scoring youth are more responsive to changes in a school's academic

performance, the marginal students who are attracted to schools that increase their academic performance tend to be students with high average academic performance and who are less likely to receive lunch subsidies than the average student. Thus, heterogeneous preferences for school quality lead to disparate competitive pressure across high and low performing schools, which could lead to a two-tiered educational system in the long run.

In the absence of public school choice, residential location alone determines school assignment. Earlier papers, such as Hoxby (2000), have focused on the competitive pressures created when school districts compete on the basis of the residential location of constituents. Public school choice advocates argue that competition on the basis of residential location alone may not be enough to spur schools to focus on student achievement, since parental responses are muted by budget constraints and desires for other neighborhood attributes. Our results measure the additional competitive pressure created by a public school choice plan when school assignments are decoupled from residential location. Rather than residential location constraints, our analysis points to heterogeneous preferences in a differentiated schools (products) market and the resulting market segmentation as the key factors limiting competition among schools.

## **II. Previous Literature on School Choice and Competition**

A number of papers in the economics of education use aggregate measures of market concentration to infer the extent of school competition in different metropolitan areas and relate those indirect measures of competition to academic outcomes. For example, Borland and Howson (1992), Hoxby (2000) and Hanushek and Rivkin (2001) find that less market concentration as measured by a Herfindahl Hirschman Index is associated with higher student performance on tests and higher teacher quality.<sup>1</sup> A related literature has studied the relationship between academic performance and the share of the local education market controlled by the public sector. Several such papers have found that the higher market share in charter or private schools is associated with improved school performance (Hoxby (1994) and Couch, Shugart and Williams (1993)).

---

<sup>1</sup> For a recent review of this literature, see Belfield and Levin (2002).

The theoretical basis for the use of the Herfindahl index comes from the equilibrium relationship between price-cost margin and the number of competitors in a symmetric, homogeneous goods product market with Cournot competition. However, the education market is not homogeneous. Schools are differentiated by test scores, racial composition and location. In differentiated markets, the number of firms is no longer a sufficient statistic for the degree of competition. The degree of competition will depend on the characteristics of the firms, consumer preferences over those characteristics, and the degree to which those preferences vary in the population.

Accordingly, the recent literature in industrial organization has focused on estimating underlying preference parameters of consumer's indirect utility to understand demand, substitution patterns, and nature of competition between firms in differentiated products markets. In our context, estimates of preference parameters will yield estimates of the demand response faced by individual schools, allowing insights into the nature of competitive pressure on quality and student sorting 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.<sup>2</sup> On such surveys, researchers have typically found that academic standards and teacher quality loom large in parents' minds.<sup>3</sup> Even among those attending private religious schools, parents often report academic quality to be paramount. (Convey (1986), Nelson (1988), Goldring and Bauch (1995)).

Nevertheless, because stated preferences may not reflect behavior, inferring parents' actual preferences from such questionnaires can be misleading. Parents may implicitly be limiting their choice sets in a manner not apparent to the researcher (such as considering only nearby schools or schools with a given racial composition), or they may tailor their responses to fit social norms. For example, they may over-report the

---

<sup>2</sup> For a review of the literature in this field prior to 1990, see Maddaus (1990) and Carnegie Foundation (1992).

<sup>3</sup> 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).

importance they place on academic quality and under-report 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.<sup>4</sup> 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 differences in preferences across observable demographics, 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), using a conditional logit framework, 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. However, the Minneapolis choice plan involved only a small percentage of parents, with very limited options, and a history of incentives and participation that may have affected parents to stated choices. The CMS school choice plan provides a unique opportunity to examine preferences across the population of students at the introduction of sweeping school choice program, using a more flexible mixed-logit discrete choice framework (Hausman and Wise (1978) and Berry Levinsohn and Pakes (1995) and (2004)).

### **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" and ordered the district to dismantle the race-based student

---

<sup>4</sup> See Manski and Wise (1983) for an early application to college choice.

assignment plan by the beginning of the next school year. 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 a seat at this school. Magnet students were similarly guaranteed admission to continue in their current magnet programs. Admission for all other students 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 (2005).

#### *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 often paired non-contiguous black and white neighborhoods to achieve racial balance, were redrawn as a result of the Fourth Circuit Court’s ruling. Boundaries could no longer be drawn on the basis of race. Under the choice plan, 43 percent of parcels were assigned to a different elementary grade ‘home school’ than they were assigned to the year before under the bussing system. At the middle school and high school levels this number was 52 and 35 percent respectively. 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 due to changes in boundaries elsewhere.

Therefore, in our analysis, the home school for many students is often not the school they would have been assigned at the time they chose their residence. 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 a given 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. After controlling for 2001-2002 school assignments, there are still on average 10.4 different first-choice elementary schools within each 2002-2003 elementary home school assignment boundary. 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 home school might not have listed their most preferred school as a first choice if there was a low probability of admission. Instead, they may have hedged their bets by listing a less preferred option with a higher probability of admission in order to avoid 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 due to strategic hedging on quality.

However, there were a number of reasons why strategic hedging was unlikely to have been a major concern in the 2002-2003 CMS choice plan. First, the choice application was vague in describing how slots in oversubscribed schools would be

allocated and how the lottery system would be operated. The school district also communicated to parents that they would make every attempt to give each student admission to one of their chosen schools. In order to accommodate demand, the district substantially expanded capacity at popular schools. In addition, the district gave a ‘priority boost’ to low-income students choosing to attend schools with low concentrations of low income students. Hence, choices for top schools by students with under-performing home schools would be given top priority. This would counteract the incentive for these students to hedge their choices as outlined above.

Nevertheless, we tested for evidence of 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 tested if students who were redistricted to lower-performing schools chose on average schools with lower test scores, relative to students in the same school zone in 2001-2002 who did not experience a negative shock to their default school quality. We did not find evidence that students who were assigned 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. For these reasons, we believe that the extent of strategic manipulation in the first year was limited and that parents were generally reporting their true preferences.<sup>5</sup>

#### **IV. Data**

Working with the Charlotte Mecklenburg School (CMS) Board and district officials, we obtained access to a wide range of administrative data for students in kindergarten through twelfth grade in the years surrounding the implementation of the choice program. The data fall into five broad categories: (1) information from student choice forms, (2) geographic information, (3) student demographic information, (4)

---

<sup>5</sup> 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 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.

student test score information, and (5) information on school characteristics. Throughout the analysis, we will focus on students entering grades 4 through 8. We focus on this segment of the student population for two reasons. First, younger students do not have baseline standardized test scores. Since we wanted to study the relationship between baseline test scores and parental preferences and the implications for student sorting in school choice, we focus on students entering fourth grade or higher. Second, we do not consider students entering high school in this analysis since high schools choices are far fewer in number with consequently less independent variation in characteristics of schools in the choice set, and high school choice is likely influenced by factors such as graduation rates and athletic programs that are not central to elementary and middle school choices, and thus would be better handled in a separate analysis. We therefore focus on choices for elementary and middle schools in this paper.

### *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, limited English proficient, orthopedically disabled, mentally disabled, or hearing impaired. This left a sample of 96,147.

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, 34,313 students filled out only a first choice, 17,396 students listed only a first and second choice, and 44,438 students listed the maximum of three choices. Approximately 8 percent of students listed only one or two choices and did not bother listing their guaranteed home school as their final choice. We assumed that these parents knew that the choice system would have taken their home school as their implied next choice and we assigned the home school as the final choice to these students. Therefore, our final sample of fourth through eighth grade students contained 12,755, 6,701 and 17,360 students with one, two or three choices identified respectively.

### *Geographic Information*

Using information on the exact location of each student’s residence along with the exact location of all the schools, we calculated the driving distance in miles from each student’s residence to each of the schools in the district. In 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 federal lunch subsidies. We also received 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 used data on reading and math scores on North Carolina end-of-grade exams for students in grades 3 through 8.<sup>6</sup> Test scores are reported based on grade level and year. We used reading and math scores from the spring of 2002 (before the choice program) as well as for the spring of 2003 - the first year after students were assigned to schools. Analyses that rely on these scores are limited to grades where the scores are available. In particular, all of the mixed-logit models focus on baseline test scores and preferences for school test scores, and are therefore restricted to students who were enrolled in grades 3 through 8 in CMS in 2002. As a measure of student-level academic ability, we added each student's math and reading score from the spring of the 2001-2002 school year, and then standardized by the mean and standard deviation of test scores for all students of the same grade level in the district.

### *School Characteristics*

We constructed three school level measures that, while not exhaustive, capture the main dimensions that are often believed to influence school choice. First, as already discussed, we calculated distance to each school. Second, to proxy for the academic quality of a school, we constructed a measure of school-level academic scores. The measure is the average standardized student test score for students attending each school in the first year of choice (2002-2003). This calculation was done separately for each academic program whether or not they were housed in the same physical structure. By using scores from 2002-2003, we implicitly assumed that families foresaw the sorting resulting from choice, and made choice decisions accordingly. However, in specification checks, we found that using the baseline year measures for school scores yield similar results.<sup>7</sup> 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 in 2002-2003 that were black.

---

<sup>6</sup> Students in kindergarten through 2<sup>nd</sup> grade do not take the state exams, and high school students only take end-of-course exams in the subjects they choose.

<sup>7</sup> In addition when we included both 2001-2002 and 2002-2003 measures are included in the analysis, we found that the 2001-2002 measures become insignificant, implying that the 2002-2003 provided a better fit for the choice data.

Finally, we constructed three variables that captured potentially important non-academic dimensions of a school. The first was an indicator of whether the student attended the school in the prior year (2001-2002). We expected that students in a continuing grade would have had a preference for remaining in their prior school because of the continuity of the peers and the instruction. The second variable was an indicator for the student's guaranteed home school. Anecdotal evidence from parent interviews we conducted in Charlotte suggested that many parents have a strong preference for their children to attend the neighborhood school. The final variable was an indicator for schools that are in a student's choice zone. Each student was assigned to one of four choice zones based on their residence, and transportation was only provided to schools within the zone. We expect that students, particularly those from poor families, would have preferred schools for which transportation was guaranteed.

## **V. Empirical Model**

We begin this section with a brief overview of our estimation strategy and a discussion of why the CMS data and policy intervention are particularly useful for identifying heterogeneity in school preferences. We then describe more formally the model that we use to estimate the preference parameters behind individual choices, how this model is estimated, and how the results are 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 mean and variance of preferences over school characteristics, and how they vary with student level demographics and baseline academic achievement.

We estimate a mixed logit discrete choice demand model (McFadden and Train 2000, Train 2003). Mixed logit models of demand are multinomial logit 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 differs from the traditional conditional or multinomial 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 (IIA) assumption. This restriction implies that, when choice sets are altered (for example by the introduction of a new school), substitution to any new school does not depend on its similarity to existing schools. This is likely to be an unrealistic assumption, particularly in a school choice program.

In contrast, the mixed logit model allows the random component of the indirect utility function to interact with the product attributes – leading to random preferences for product attributes in the indirect utility function. The mixed logit can approximate any random utility model, given appropriate mixing distributions and explanatory variables (Dagsvik (1994), McFadden and Train (2000)). 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. In addition, as we will discuss further in the subsection on identification, changes in the choice set generated by multiple choice data are often needed to identify the mean and variance of the preference parameters. The CMS data provide this important source of variation.

### *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:

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

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

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,  $\mathbf{e}_{ij}$ , that reflects unobserved idiosyncratic preference of student  $i$  for school  $j$ .

$$(1) \quad U_{ij} = X_{ij} \mathbf{b}_i + \mathbf{e}_{ij}$$

We assume that the unobservable component ( $\mathbf{e}_{ij}$ ) is distributed i.i.d. extreme value, which yields the usual logit form for the choice probabilities conditional on  $\mathbf{b}_i$ .

Heterogeneity in individual preferences implies that the coefficients,  $\mathbf{b}_i$ , 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  $\mathbf{b} \sim f(\mathbf{b} | \mathbf{m}, \mathbf{q})$ , where  $f(\cdot)$  is a mixing distribution, where  $\mathbf{m}$  denotes the mean, and  $\mathbf{q}$  represents the other parameters describing the density function. Second, we separately estimate parameter distributions for students in each of the four main demographic categories: white and African American by lunch subsidy status. This allows us to compare means and variances of preferences for school characteristics across the different socio-economic groups. In addition, we allow the coefficient on a school's standardized score to vary with a student's baseline test score and family income by including interactions between these student characteristics and the school mean test score in the vector of characteristics,  $X$ .

In the specifications that are reported below, we assume that all random parameters are drawn from a joint normal or log normal distribution. In particular, we

allow for a covariance between preferences for proximity and school mean test score, since these are key dimensions of product differentiation. Other preference parameters are assumed to be independently distributed.<sup>8</sup>

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

$$(2) \quad P_i(j^1, j^2, j^3) = \Pr\left\{U_{ij^1} > U_{ik} \forall k \in J_i^1\right\} \cap \left\{U_{ij^2} > U_{ik} \forall k \in J_i^2\right\} \cap \left\{U_{ij^3} > U_{ik} \forall k \in J_i^3\right\} \\ \int \prod_{c=1}^3 \sum_{k \in J_i^c} \frac{e^{X_{ij^c} \mathbf{b}}}{e^{X_{ik} \mathbf{b}}} f(\mathbf{b} | \mathbf{m}, \mathbf{q}) d\mathbf{b}$$

The term inside the integrand represents the probability of observing the three ranked choices conditional on the preference coefficients ( $\beta$ ): this is the product of three logit probabilities evaluated at  $\mathbf{b}_i$ , corresponding to the probability of making each choice from among the remaining options.<sup>9</sup> This conditional probability is integrated over the distribution of  $\mathbf{b}$  to yield the unconditional probability of observing the ranked choices.

These probabilities form the log-likelihood function:

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

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

---

<sup>8</sup> Allowing for general covariance structure across all parameters led to instability in the estimated covariance terms in some specifications, but did not significantly affect the remaining parameters or the substantive results that we report.

<sup>9</sup> 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<sup>st</sup>) choice is observed for every individual, it is difficult to be sure whether an unexpected choice was the result of an unusual error term ( $\mathbf{e}_{ij}$ ) or unusual preferences by the individual ( $\mathbf{b}_i$ ) 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 a strong preference for that attribute, because independence of the additive error terms across choices would make observing such an event very unlikely in the absence of a strong preference.

This 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 provides summary statistics on all of the students who submitted choice forms for the entire sample, and broken down by race and lunch subsidy status. Roughly 8500 students submitted choice forms in each grade, with somewhat less students per grade in the later high school years when school enrollment is no longer mandatory. Whites and Blacks each comprise just over 43% of the school population. Approximately 10 percent of white students receive federal lunch subsidies, while just over 60 percent of African Americans do. Among white students who were ineligible for the lunch subsidy program, about half (51%) listed only one choice on their forms, while just over a quarter (29%) of these students listed all three choices. Students who were African American or eligible for lunch subsidies were much more likely to fill out all three choices, with nearly two thirds of the students who were both African American and receiving lunch subsidies filling out all three choices

There are at least two reasons why white students who were not eligible for lunch subsidies were more likely to list only a single choice. First, the average quality of their home schools is significantly higher. Table I shows the average percentile scores for reading and math in student's home schools by race and free lunch eligibility. The average scores for home schools of students who were white and ineligible for lunch subsidies were one quarter to one half of a standard deviation higher than those 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 guaranteed school. Hence they would be more likely to fill out only one choice. Moreover, the longer distances between schools in the suburban areas may effectively limit the choice of these students.

If one's home school truly were one's first choice, there would be no incentive to fill out the remaining slots on the choice form. As seen in Table I, the proportion of each student group that choose their home school was directly related to the proportion that listed a first choice only. Nevertheless, many parents specified multiple choices even if

they listed their home school first. For instance, 64% of white students who were not receiving federal lunch subsidies chose their home school first, while only 51% listed only a first choice. This implies about a fifth of the white parents eligible for lunch subsidies whose top choice was their home school actually provided additional rankings. About half of the African American children eligible for lunch subsidies whose top choice was their home school provided some additional listings. Whatever their reasons for doing so, the availability of multiple choices from those who listed their home school first will further aid in the identification of the preference parameters.

### *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. It measures approximately 22 miles across and 30 miles north to south at the widest and longest points. 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) 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 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 I reports the average travel distance (in miles) to the

nearest top quartile school. 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.

Table I also provides the mean and standard deviation of baseline test scores of students in CMS in Spring 2002 by race and lunch-subsidy status. Student-level test scores are reported as the standardized score – standardized 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. While there are large differences in average test scores across the groups of students, there is substantial variation in student ability within each category as measured by performance on standardized tests. While the mean of the distribution for white students not receiving lunch subsidies 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 in each of the other 3 categories. Similarly, the within-school variation in performance is greater than across-school variation in performance. Such variation in student level-baseline achievement within 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 mixed logit model was estimated separately by race (white vs. non-white) and by receipt of federal lunch subsidies. Therefore, all estimates for both the means and variance-covariance matrices of the preference parameters were allowed vary across race and lunch-subsidy status. Within race and lunch-subsidy status, we included as explanatory variables measures of school and student characteristics that are central to understanding competition on quality in the context of the school choice debate. To capture the importance of proximity and travel costs, the specification included driving distance (in miles) from the student's residence to the school (measured in miles), an indicator if bussing was provided for the student to the school (the school was in the student's zone), and an indicator if the school is the student's neighborhood school. An

indicator if the student attended the school in the prior year was included to capture the importance of continuity for students who were continuing in elementary or middle school. To capture the academic quality of the school, we included a measure of average test scores in the school (the school level average of all students' standardized math and reading scores in spring of 2003). We interacted the school's average test score with the student's standardized baseline test score (standardized by grade level across the district) 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-means with the countywide median of \$51,000). These interactions allowed the effect of school test scores on school choice to vary with a student's income and academic ability. Finally, to capture the racial composition of a school, we included the percent black in the school in Spring 2003 and its square. When the quadratic term has a negative coefficient, this specification yields an implied bliss point (where the quadratic peaks) for preferences over racial mix of a school.

The final estimation sample includes 36,816 students entering grades 4-8. Estimation is limited to these grades because of the lack of test scores (either baseline or school test scores) in other grades. The means and standard deviations of these variables across the 2.4 million school choice and student interactions available to our sample of students and schools are reported in Table III. The mixed logit parameter estimates are reported in Table IV. All of the point estimates were precisely estimated and statistically different from zero at less than the 1 percent level. To preserve space, the standard errors of the estimates of the preference distribution means and standard deviations are reported in Appendix Table I.

*Proximity: Travel Distance, Neighborhood Schools, and Bussing Zones*

Given strong priors that the coefficient on distance would be non-positive, we imposed a lognormal distribution on the preference coefficient for distance:

$\mathbf{b}_{dist} = -\exp(\mathbf{a})$ , where  $\alpha$  was assumed to be normally distributed. The coefficient on distance was the only coefficient for which we felt comfortable imposing an assumption regarding sign. In the table we report the mean and standard deviation for the actual coefficient ( $\mathbf{b}_{dist}$ ) as well as for the underlying normal ( $\alpha$ ). The negative weight placed

on distance is fairly uniform across the four demographic groups. However the effect of distance is slightly larger in absolute value for white students compared to black students, and slightly more variable for students receiving lunch subsidies. For an average student, each additional mile of distance reduces the odds of choosing a school by roughly 35% among whites and 25% among nonwhites. The variation in this effect across students is large relative to the mean, implying that for some students distance is a major barrier to choice while other students place very little weight on distance.

The coefficient on the home school indicator was intended to capture a preference for the neighborhood school. Preferences for home school have a strong average effect across all demographic groups; however there is also a large variance in the idiosyncratic preference for this characteristic. The pattern for home school preference is similar to that found for distance: The mean preference for a home school is somewhat larger for whites, and somewhat more variable for students eligible for lunch subsidies. The mean effect is roughly equivalent to 6-7 miles in travel distance in each sub-sample of students.

Because the home school is often the closest school to the student, this variable may pick up a non-linearity in preferences for proximity (parents have an added preference for the closest school to them). However, it may also represent a preference for the characteristics of being the “neighborhood school”. We investigated the degree to which the estimated preference for a home school reflects a non-linear preference for proximity versus a value of neighborhood school by re-estimating the model on a sample of students living along newly-created home-school boundary borders that bisected old school assignment zones. To the extent that these students had attended the same schools in the prior year, and faced a similar distance to all schools, we hoped to better isolate the importance of the Home School designation. The preference for Home School was lower for students living on the boundary of a home school zone, but still positive and significant. Hence there appears to be a preference for the neighborhood school in addition to a non-linear component of preference for proximity.<sup>10</sup>

Another possibility is that the home school indicator is picking up a default effect, rather than a preference for neighborhood school. If parents do not want to invest the time

---

<sup>10</sup> In general, it appeared that many new home school boundaries were set natural neighborhood boundaries, often with visible discontinuities in student densities on either side of the new boundary segments, and that controlling for the home school effect a linear specification for distance best fit the data.

to fully consider all the school in the choice set, they may simply list their guaranteed school as their first choice.<sup>11</sup> This would imply a stronger preference for a home school on the first of the three choices. We estimate the mixed logit model with a separate coefficient on the home school indicator for just the first choice. Default behavior as described above should imply a positive coefficient on the home school indicator interacted with first choice. We found that the coefficient was *smaller*, and within one half of one standard deviation of the mean overall preference distribution for a home school. Hence, we interpret the coefficient on the home school indicator as picking up a preference for the ‘neighborhood school’ instead of picking up a default behavior.<sup>12</sup>

Finally, the coefficient on a school being in a student’s Choice Zone was intended to capture lower travel costs to these schools, since transportation by the district was only provided to schools within a student’s Choice Zone. All four demographic groups have a strong mean preference for schools in their choice zone, with the effect being largest among students who are eligible for lunch subsidies (as would be expected if these students had limited access to alternative transportation). The standard deviation of the coefficient on Choice Zone is of roughly the same size as the mean in each demographic group, suggesting considerable variation across students in these preferences.

Overall, the estimates for travel distance, neighborhood schools, and choice zones support the same general conclusion. While there are some differences across demographic groups, it is clear that proximity is an important determinant of school choice for the average student. At the same time, there appears to be great heterogeneity across students in the weight that they place on proximity in choosing a school, ranging from students who place virtually no weight on the proximity variables to students who weight these variables more than twice as highly as the average student.

### *School Test Scores*

Given our prior that preferences for school scores would vary with student baseline academic ability as well as student income level even within race and lunch-

---

<sup>11</sup> Note this is different than defaulting by not turning in a choice form. Recall that 95% of parents submitted the choice forms. Of these, defaulting behavior might be an over-propensity to list Home School first, and nothing else.

<sup>12</sup> Recall that there are substantial fractions of students listing their Home School first, but also completing the subsequent choices.

subsidy status, we included school test scores and their interaction with the student's baseline test score and neighborhood income level. For students who are eligible for lunch subsidies, we did not include the interaction with neighborhood income because all of these students are presumably very low income.<sup>13</sup> Both neighborhood income and the student's baseline score are "de-measured", so that the coefficient on the main effect of school score measures the value of school test score for a student with average income and baseline test score (both equal to zero). The coefficient on the main effect (the school test score) was treated as a random parameter, allowing for additional variation in preferences for school scores (beyond that explained by income and baseline test score).

These estimates reveal a number of interesting results. For an average student, the mean preference for school scores is larger for non-white students within lunch-subsidy status, but students not receiving lunch subsidies value school scores much more than their peers who were receiving lunch subsidies. The difference between those who were receiving lunch subsidies and those who were not is consistent with the coefficient on the interaction with neighborhood income. Higher neighborhood income was strongly associated with higher mean preference for school scores, with a similar effect for both whites and non-whites. Thus, the fact that the students receiving lunch subsidies are lower income should imply that they also place a lower value on school scores. Using the coefficient on the income interaction, the difference in the mean preference between students receiving lunch subsidies and students not receiving lunch subsidies (about 0.8 within each race group) is roughly what would be predicted by a \$50-\$60 thousand dollar income difference, which is roughly the right order of magnitude for the income difference between these groups. Thus, there appears to be a clear relationship between income and preference for schools with high test scores in these data.

The mean preference for school scores is also increasing in the student's baseline test score. 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 value a school's test scores more. The coefficient varies somewhat across groups (with white free-lunch students having the

---

<sup>13</sup> In initial specifications using a conditional logit, income interactions with the preference for school scores were generally insignificant for the lunch-recipient segments.

smallest coefficient), but there is no obvious pattern by race or free-lunch status. The effect of a student's baseline score on the preference for school test scores is similar in magnitude to the effect of income: A one standard deviation increase in the baseline test score is associated with a 0.3-0.6 increase in the mean preference for school test scores, while a one standard deviation increase in neighborhood income (about \$27,000) is associated with a 0.3-0.4 increase in the mean preference for school test scores.

While the mean preferences for school test scores appear to be somewhat stronger for non-whites (consistent with a lower mean income level and test scores for non-whites than for whites) the heterogeneity in preferences within racial groups is even larger. Differences in baseline test scores and income each generate a standard deviation in preferences of roughly 0.3-0.6 based on the calculations from the previous paragraph. In addition, the variance on the coefficient for the main effect of school test scores is also estimated to vary with a standard deviation of around 0.3 for non-whites and 0.65 for whites, which generates additional heterogeneity in preferences. Taken together, the total variation in preferences for school test scores across individuals is substantial.

#### *Summarizing Preference Trade-offs Between Proximity and School Scores*

For the student with the average preferences, the preference for a home school is very large relative to the preference for School Scores. For a student with average baseline test scores, and average neighborhood income, the preference for a home school is equivalent to 1 to 2 student-level standard deviations in average test scores, depending on the demographic group. However, there is a large variance in preferences for home school, and the idiosyncratic preference for home school is strongly negatively correlated with the preference for School Scores across all of the four demographic groups. This implies that, while many students are very inelastic with respect to School Scores, there is a significant density of students who highly value school scores and have low preferences for their neighborhood school. These students are willing to seek better schools over a relatively broad geography.

Table V presents the trade-offs between the home school preference and preference for school score for various values of idiosyncratic preferences, baseline test

scores, and neighborhood income levels. The first row lists the gains to test scores that would be needed to compensate an average student (average preferences, test score=0, and standardized income=0) enough to induce him to choose a non-home school. The magnitudes are very large – from 1 to 5 student-level standard deviations in test scores. Even for a student with a standardized test score of 2 (row 2), a 0.59 to 2.08 (African American students not receiving lunch subsidies and white students receiving lunch subsidies respectively) gain in school scores would be needed to induce a student to choose a school other than the home school. The gain in average test scores again falls for high-income students (row 3) who have higher mean preferences for school scores, however it remains relatively high in magnitude.

The strong negative correlation between idiosyncratic preferences for a home school and school mean score has a large impact on the trade-off between home school and average scores. Table V presents trade-offs for students with preferences for home school that are one to two standard deviations below the mean. A student with a preference that is two standard deviations below the mean, with a test score of two, would be willing to choose a school other than the home school for a gain in test scores between 0.08 and 0.35 (row 8). This is equivalent to a 2 to 10 percentile point increase in test scores, implying that there is a substantial segment of the population who are willing to choose away from their neighborhood school for a slight increase in school level test scores.

While preferences for a home school and school mean score are strongly negatively correlated, the preference for a home school and distance are only weakly negatively correlated. This implies that students who place a low value on the home school indicator, and hence a high value on school mean test score are relatively elastic on school scores with respect to proximity. For the mean preference, white and black students not receiving lunch subsidies are willing to drive an extra mile for 0.30 and 0.14 gains in school scores respectively. For students with average preferences but with a baseline test score of 2, this trade-off falls to 0.15 and 0.09 gains to school scores respectively.

Taken together, these estimates for the distribution of preferences for distance, home school, and school mean test score imply that there tend to be two types of

students: i) those who highly value proximity through their neighborhood school and are not likely to choose another school without substantial improvements to average school scores, and ii) those who place little value on proximity through their neighborhood school and place a large value on school scores. The first type of student is highly inelastic to school quality as approximated by average test scores. This type of student will be served by their local school, and will stay with that local school even in the face of potentially large losses to relative school quality. The second type of student is very elastic with respect to school quality - willing to travel over a relatively broad geography for much more modest gain in school scores. These underlying characteristics of the preference distribution have important implications for demand-side pressure for competition on school quality which we will discuss further in Section VI.

### *Racial Composition*

While student preferences for school racial composition are not the focus of this paper, they are important to account for because racial composition of a school is correlated with average test scores of the school. To capture the racial composition of a school, we included the percent black in the school and its square. When the quadratic term has a negative coefficient (which was always the case), this specification can be interpreted in terms of an implied bliss point (where the quadratic peaks) for preferences over racial mix of a school, and a quadratic cost loss function for differences from this bliss point. In Table IV, we report the estimates for the linear and quadratic terms, along with the implied bliss point. We allowed for a random coefficient on the linear term but not the squared term, which is equivalent to allowing for unobserved heterogeneity in the bliss point but not the quadratic loss around that bliss point. Other specifications, such as dummies for ranges of percent black or a spline in percent black, yielded similar results.

Not surprisingly, there were large differences between the races in their valuation of a school's racial composition (and little difference by lunch-subsidy status). The mean bliss point for whites was around 30% black, while the mean bliss point for non-whites was around 70% black. Thus, the average preferred school for each racial group was one in which 70% of the school was their own race. But there was also substantial variation in this preference within racial groups, with a standard deviation in the bliss point of

20%-30%. The quadratic term was negative for all demographic groups, but was larger for whites than non-whites, and larger for free-lunch ineligible than for eligible.

These results are quite consistent with an earlier literature highlighting racial differences in stated preferences regarding the racial composition of neighborhoods. That literature (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.

The focus of this current paper is on preferences for school quality rather than preferences for racial composition. Given the history of bussing for integration in CMS, and the higher-than average private school attendance of whites in the district, it is not clear that these racial preference estimates generalize outside of settings similar to this one.<sup>14</sup> However, 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 understate black student valuation of school scores since they prefer schools with above-average black enrollment. 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.

---

<sup>14</sup> Approximately 12-15% of students attend private schools in CMS compared with 8-9% in the state of North Carolina, and 9-10% in the U.S as a whole.

### *Redistricted Students*

For the results in Table IV, we have taken location as exogenous. If parents with strong preferences for certain schools have moved to be near those schools, we may overstate the importance of proximity. What we interpret as a strong preference for proximity as influencing school choice may actually be the opposite – strong preference for a school influencing proximity. The history of court-ordered bussing in this district may have muted this effect. School assignment boundaries prior to the school choice plan were not necessarily determined by proximity; many students were bussed long distances in order to achieve racial balance. Moreover, school assignment boundaries were redrawn on a regular basis during the three decades of race-based bussing in order to maintain racial balance in the schools, making it difficult for parents to predictably choose a school through residential location.

In addition, as shown in Table I, a substantial fraction of parents who listed their home school as their first choice also listed subsequent choices. These additional choices would assist in identification of preferences even in the presence of residential sorting by revealing parent's trade-offs in the event that their child could not attend their preferred neighborhood school. In other words, second and third choices act as a hypothetical redistricting among those who chose their home school first and may have located near it because it was their preferred school by restricting the choice set to exclude the neighborhood school.

Nevertheless, there is the potential that our preference estimates may be biased toward overstating the importance of proximity. To evaluate the effect that endogenous residential location may have on our estimates, we re-estimated the mixed-logit models using only the subset of students who were assigned to a new home school as part of the choice plan. As part of the implementation of the choice plan, the district redrew school assignment boundaries. The old boundaries had been drawn for the purpose of racial balance, while the new boundaries were drawn based on proximity to a neighborhood school. As a result, approximately 50% of students were redistricted to a new home school. It is unlikely that these students could have foreseen the redistricting to their new home school, and their residential location should not have been driven by a strong preference for the new home school.

Table VI provides summary statistics comparing the redistricted sample to the sample of students who were not redistricted. Because of the nature of the prior system of bussing, students who were redistricted were much more likely to be non-white and eligible for lunch subsidies. But within the four demographic groups, the redistricted students looked similar to those who were not redistricted in terms of baseline test scores and median income. More interestingly, the redistricted students were much less likely to choose their home school or their last year's school, and much more likely to list three choices. These facts are not necessarily evidence that redistricted students have less preference for their home school: Students who were not redistricted were more likely to have their home school be their last year's school, making it very likely that they would choose that school. In contrast, redistricted children faced a less clear choice since their last year's school was no longer their home school.

Table VII reports results from the mixed logit model estimated on the sample of students who were redistricted. The most striking feature of these estimates is their similarity to estimates from the full sample. Estimates of the mean and standard deviation of all the preference parameters are qualitatively and quantitatively similar. The mean of the parameter for home school is actually higher for all demographic groups in the redistricted sample, while the means for the distance and choice zone parameters are about equally likely to increase as decrease in the redistricted sample. Overall, these estimates suggest that endogenous residential location is not a major source of bias in this data.

#### *Other Robustness Checks*

A range of alternative specifications yielded similar quantitative and qualitative results. We have pooled elementary and middle school students for simplicity, but estimating the model separately for elementary and middle schools yielded similar parameter estimates. As already mentioned, we experimented with alternative specifications for the racial composition of the school, including dummy variables and splines in percent black. The spline estimates were very consistent with the more parsimonious quadratic specification.

We also specified distance to each school in terms of driving time (based on expected speed on each class of road) rather than driving distance, yielding nearly identical results. We experimented with a range of alternative proxies for academic quality of a school. Using closely related measures such as the average scale score or the average percentile score resulted in nearly identical estimates. Allowing for non-linearities in the effect of school scores, through a quadratic or spline term, did not change the qualitative implications of the parameter estimates. However these models fit the data poorly in the tails of the distribution, and for this mechanical reason they generated implausible results when used in simulations. Including separate terms for the school average test scores of whites and non-whites separately resulted in all students, both white and non-white, placing similar weights on the two scores, with both racial groups placing a larger weight on white test score performance. Again, the implications of the results were unchanged across these specifications. Finally, including a separate dummy variable for schools that were academic magnets (e.g. International Baccalaureate, Math and Science magnets) reduced the mean coefficient on school test scores about in half. This result highlights that average test scores are a proxy for the academic focus of a school, and not necessarily the sole causal factor driving demand.

Finally, when we estimated a general mixed-logit model with full covariance terms for the parameters, we found that some covariance terms became unstable in some specifications. For example, when we included a covariance between racial preferences and preferences for other characteristics could often be unstable, yielding corner solutions in some circumstances. However, the means and standard deviations of the preference parameters were largely unchanged, and the implications of the estimates in the demand simulations were very similar. This suggests that some of the covariance terms are poorly identified, but that these terms are not of first order importance to simulations of demand.

## **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 .33 average student-level standard deviations 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.<sup>15</sup>

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. Parents with high preferences for school scores, and thus low preferences for their neighborhood school, are sensitive to changes in school scores and willing to consider schools over a relatively broad geography. These parents are both likely to only consider high scoring schools for their children and willing to change schools in response to an increase in score at another high scoring school, even one that is located further away. 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 .33 average student-level standard deviation score increase) 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 reports differences between marginal and average students in the percentage receiving lunch subsidies; Figure 7 reports differences in mean test scores. The points below the 45° degree line in the Figure 6 indicate schools where a lower proportion of the marginal students received

---

<sup>15</sup> This is approximately equivalent to a 10 point increase in the average percentile score for students attending that school.

lunch subsidies than the average student. It is evident in graph that the marginal students had lower rates of lunch subsidy receipt than the students already enrolled. In other words, the marginal students were more affluent than 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 key features of the simulations reported in Figures 5-7 appear to be driven primarily by the estimated heterogeneity in preferences, rather than other details of the specification. In all the alternative specifications we have estimated that allowed for heterogeneity in preferences, we found that an increase in school test scores had a much larger effect on demand in high scoring schools, and attracted higher-performing and higher-income students to the school (particularly at low scoring schools). Eliminating unobserved heterogeneity in preferences (estimating a conditional logit) reduces the simulated difference in demand response to higher performing schools by roughly 10%-15%. Eliminating preference heterogeneity through observable characteristics (income, race, lunch-subsidy status and baseline test scores) further decreases the difference in demand response across high and low performing schools, leading to a low demand response across all schools. Thus heterogeneity preferences appear to be a key element in understanding the properties of parental demand for schools and their implications for student sorting and demand-side pressures for school quality in a public school choice program.

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 demand-side forces may lean toward greater vertical separation on test scores. In the long run, the new equilibrium will depend on both the incentives provided to school managers as well as demand for school quality. If a greater percentage of resources are directed towards high demand and high performance schools, top tier schools will have strong incentives to improve student performance, while lower-tier

schools may not. However, district policies that commit to close schools or replace principals in schools with shrinking enrollments coupled with financial incentives for performance may minimize the degree of vertical separation. For example, North Carolina has provided bonuses to teachers in schools with test score improvements and, with the No Child Left Behind Act of 2001, the federal government has required states to penalize schools with poor test performance. Moreover, the district has replaced principals at schools with low enrollments and moved close and reorganize schools at the lowest end of performance (and enrollment). Finally, the marginal students who are brought in when a low-scoring school improves tend to have higher baseline performance than the students already enrolled. Depending upon the net payoff to school managers to improving performance, the combination of demand and supply side factors could 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.

## **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. These results illustrate some interesting and substantive features of the demand facing public schools in a choice environment. In particular, parents value proximity highly and the preference attached to a school’s mean test score increase with income and student’s own academic ability. We also find considerable heterogeneity in preferences across individuals even after controlling for income and ability, where students with higher than average unobservable preferences for test scores have lower than average unobservable preferences for proximity.

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.

School choice programs are intended to introduce market forces to motivate school improvement through the threat of parental choice. However, in differentiated products markets, the extent of competitive forces depends on the distribution of preferences in the population of consumers (Anderson, de Palma and Thisse (1992)). A textbook example occurs when low-price firm (e.g. WalMart) enters the market. Suppose it draws price elastic customers from a broad geographic region, leaving local stores to serve the residual market of local, highly-inelastic customers. The best strategy of the local store in response to WalMart's entry might be to increase prices to local inelastic customers, and let WalMart serve the elastic customers from a broader geographic market. This paper presented evidence from actual choice decisions supporting a demand system in school choice that leads to differential competitive pressure across high and low performing schools. Based on our estimates, school choice will lead to increased pressure for improvement at only higher performing schools, leaving low-performing schools to serve locally inelastic customers.

Our simulation results also show that, while lower-performing schools draw few 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.

## References:

- Abdulkadiroglu, Atila and Tayfun Sonmez “School Choice: A Mechanism Design Approach”, *American Economic Review*, 93(3):729-747, June 2003.
- Anderson, Simon P., André de Palma, and Jacques-François Thisse, Discrete Choice Theory of Product Differentiation, The MIT Press: Cambridge, Massachusetts, 1992.
- 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 Reproduction 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 from the Cleveland Scholarship Program" in Paul Peterson and Brian Hassel (eds.) *Learning from School Choice*, (Washington, DC: Brookings Institution, 1998).
- Hastings, Justine S. Thomas Kane and Douglas Staiger (2005), "Evaluating a School Choice Lottery: the Importance of Heterogeneous Treatment Effects," manuscript.
- 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 M. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review* (2000) Vol. 90, No. 5, pp. 1209-1238.
- Hoxby, Caroline M. "Do Private Schools Competition for Public Schools?" NBER Working Paper No. 4978, 1994.
- 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).

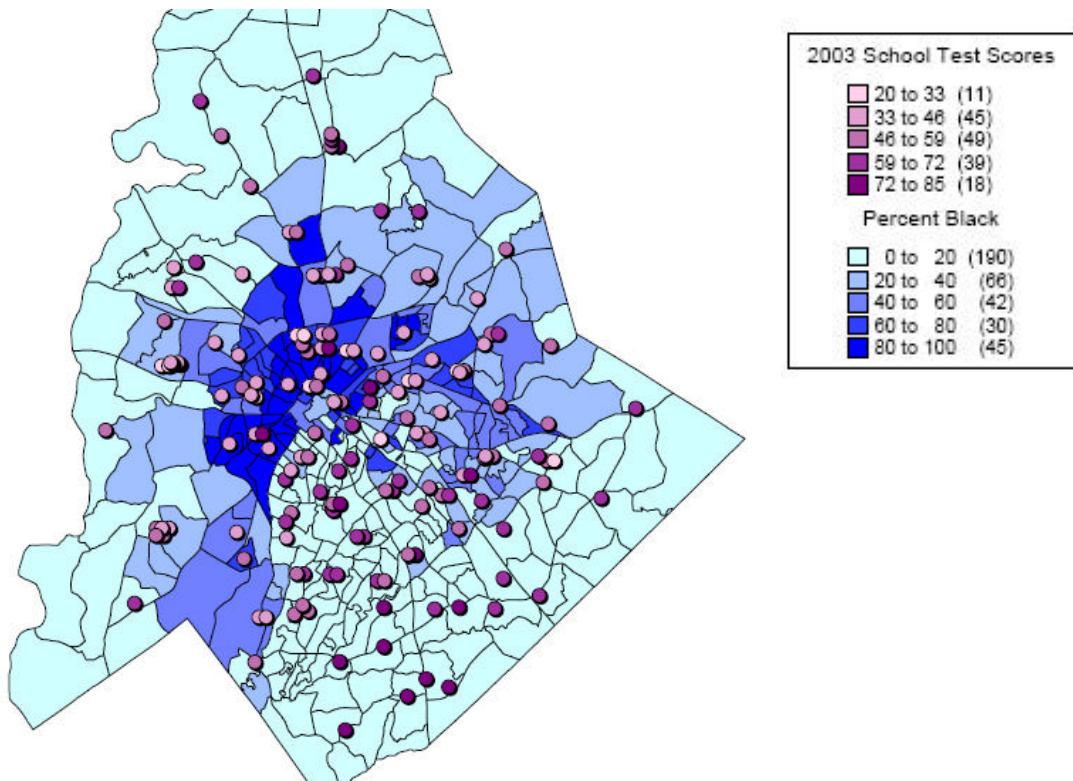
Weiher, Gregory .R. and Kent L. Tedin, “Does Choice Lead to Racially Distinctive Schools? Charter Schools and Household Preferences” *Journal of Policy Analysis and Management* (2002) Vol. 21, No. 1, pp. 79-92.

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

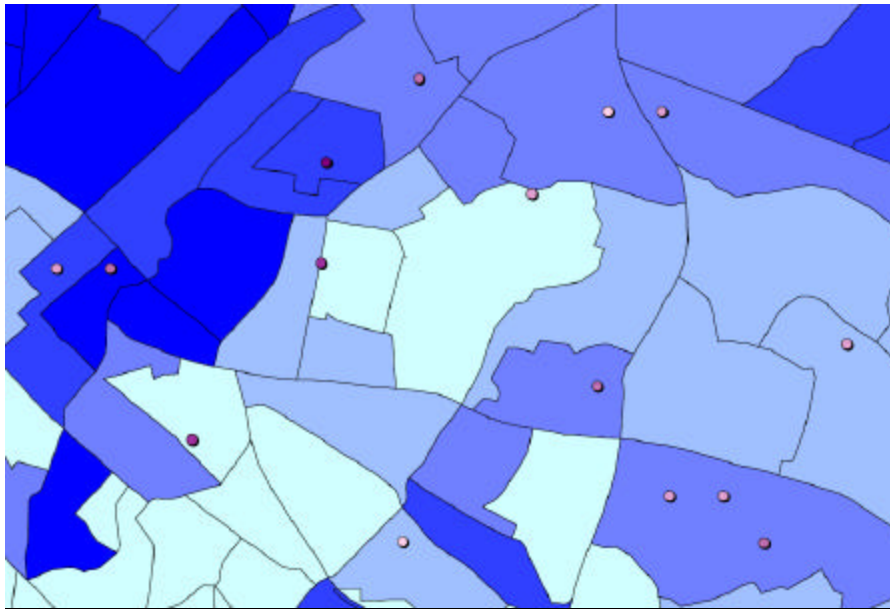
**Table I: Summary Statistics of Students and Choices**

	Not Receiving Lunch Subsidies		Receiving Lunch Subsidies	
	White	Black	White	Black
<b><i>Student Characteristics</i></b>				
<b>Average Test score</b>	0.6384	-0.0905	-0.0851	-0.6128
<b>(St. Dev.)</b>	(0.8249)	(0.8395)	(0.8480)	(0.7996)
<b>Neighborhood Income</b>	73,812	50,635	52,734	36,459
<b>(St. Dev.)</b>	(25,866)	(21,506)	(22,329)	(16,241)
<b><i>Choice Characteristics</i></b>				
<b>Percent Listed 1st Choice Only</b>	0.5123	0.2768	0.3311	0.2065
<b>Percent Listed 2 Choices</b>	0.1985	0.1778	0.2057	0.1664
<b>Percent Listed 3 Choices</b>	0.2892	0.5454	0.4631	0.6271
<b>Percent Chose Home School 1st</b>	0.6443	0.4251	0.514	0.3827
<b><i>Student-Choice Characteristics</i></b>				
<b>Home School Average Test Score</b>	0.2131	-0.1864	-0.1711	-0.3919
<b>(St. Dev.)</b>	(0.4035)	(0.3613)	(0.3739)	(0.3247)
<b>Distance to Nearest School in the Top Quartile</b>	2.5664	2.6616	2.4523	2.1272
<b>(St. Dev.)</b>	(1.6134)	(1.4828)	(1.4359)	(1.2000)

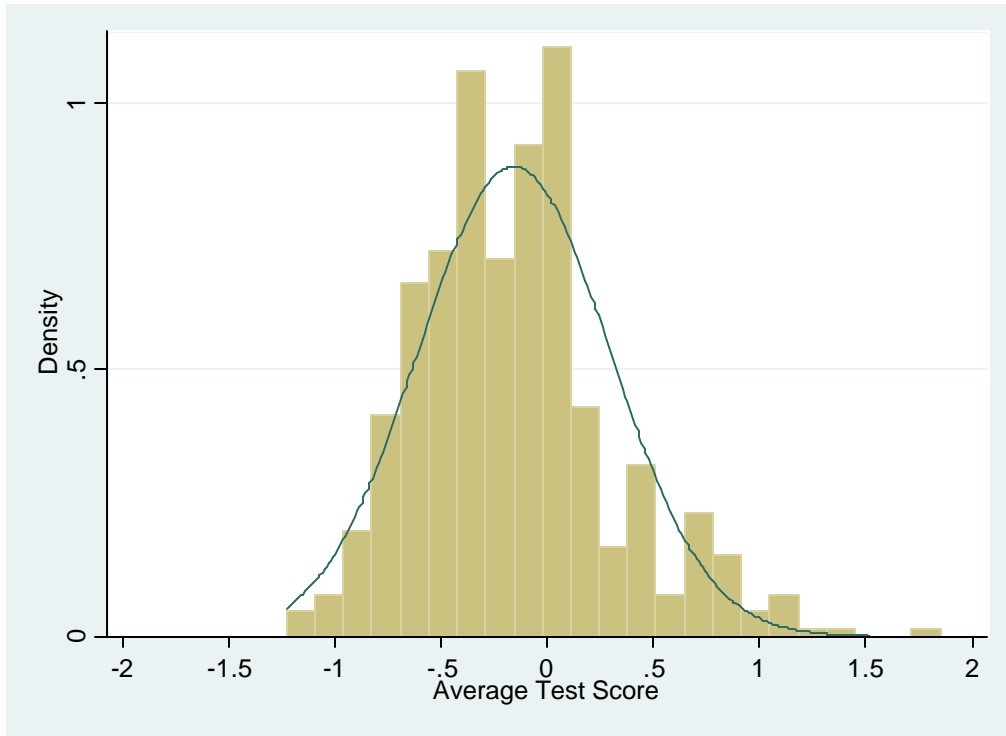
**Figure 1a: Thematic Map of Charlotte Mecklenburg County with Census Block Groups by Race and School Location by Average Test Score**



**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 II: Explanatory Variable Definitions**

<b>Variable</b>	<b>Description</b>
<b>Distance</b>	Driving distance from student <i>i</i> to school <i>j</i> calculated using MapInfo with Census Tiger Line files.
<b>School Score</b>	Average of the student-level standardized scale score for students in school <i>j</i> on math and reading End of Grade exams for the 2002-2003 school year. This is the average of the test score variable described below across all students in school <i>j</i> .
<b>Test Score</b>	The sum of student <i>i</i> 's scale score on End of Grade math and reading exams in baseline year 2001-2002 standardized by the mean and standard deviation of district-wide scores for students in his or her grade.
<b>Income</b>	The median household income reported in the 2000 Census for households of student <i>i</i> 's race in student <i>i</i> 's block group. Income is demeaned by the county-wide average of approximately \$51,000 and is reported in thousands of dollars.
<b>Percent Black</b>	The percent of students in school <i>j</i> who are black according to 2002-2003 school year administrative data.

**Table III: Explanatory Variable Summary Statistics**

<b>Summary Statistics Using First Choice Data</b>						
<b>Variable</b>	<b>Obs.</b>	<b>Mean</b>	<b>Std. Dev.</b>	<b>Min</b>	<b>Max</b>	
<b>Distance</b>	2434113	13.0071	6.7254	0.0010	42.4069	
<b>Last-year School</b>	2434113	0.0150	0.1214	0.0000	1.0000	
<b>School Score</b>	2434113	-0.1087	0.4487	-0.9537	1.9478	
<b>Test score</b>	2434113	0.0567	0.9886	-2.9113	3.0255	
<b>Test score*School-Score</b>	2434113	-0.0037	0.4579	-2.6651	5.8931	
<b>Income</b>	2434113	5.1226	27.5669	-48.5010	149.0010	
<b>Income*School-Score</b>	2434113	-0.5517	12.9342	-142.1051	229.5352	
<b>Percent Black</b>	2434113	0.5252	0.2507	0.0584	0.9801	

**Table IV: Estimates from Mixed Logit Model**

<i>Variable</i>	<i>Preference Parameter</i>	<b>Parameter Estimates*</b>			
		<b>Not Receiving Lunch Subsidies</b>		<b>Receiving Lunch Subsidies</b>	
		<i>White</i>	<i>Black</i>	<i>White</i>	<i>Black</i>
<b>Distance**</b>	<b>Mean (normal)</b>	-1.1749	-1.4572	-1.2631	-1.5596
	<b>Std. Dev. (normal)</b>	0.5145	0.5329	0.7632	0.7336
	<b>Mean (lognormal)</b>	-0.3526	-0.2684	-0.3784	-0.2751
	<b>Std. Dev. (lognormal)</b>	0.0684	0.0413	0.1273	0.0639
<b>Last-year School</b>	<b>Mean</b>	3.7941	3.3837	3.5016	2.8495
	<b>Std. Dev.</b>	2.4977	2.7896	3.4651	3.3825
<b>Home School</b>	<b>Mean</b>	2.1300	1.7373	1.9816	1.7710
	<b>Std. Dev.</b>	0.5130	0.6799	0.8248	0.7752
<b>Choice Zone</b>	<b>Mean</b>	1.1909	1.2484	1.9203	1.6132
	<b>Std. Dev.</b>	0.8285	1.2418	1.5083	1.2442
<b>School Score</b>	<b>Mean</b>	1.1732	1.8035	0.3671	0.9396
	<b>Std. Dev.</b>	0.5674	0.2688	0.6175	0.2706
<b>Test score * School Score</b>	<b>Mean</b>	0.5558	0.5734	0.2924	0.4995
	<b>Std. Dev.</b>	--	--	--	--
<b>Income*School Score</b>	<b>Mean</b>	0.0151	0.0126	--	--
	<b>Std. Dev.</b>	--	--	--	--
<b>Percent Black</b>	<b>Mean</b>	3.3068	5.1340	1.9268	3.1409
	<b>Std. Dev.</b>	2.6417	1.6447	2.0795	0.8745
<b>Percent Black Squared</b>	<b>Mean</b>	-5.4580	-3.6790	-3.5385	-2.3005
	<b>Std. Dev.</b>	--	--	--	--
<b>Implied Mean Preferred % Black</b>		0.3029	0.6977	0.2723	0.6827
	<b>Std. Dev.</b>	0.2420	0.2235	0.2938	0.1901
<b>Estimated Correlation Coefficients:</b>					
	<b>Corr(Distance, School Score)</b>	0.4939	-0.1055	0.3379	-0.6355
	<b>Corr(Distance, Home School)</b>	-0.0788	0.0007	-0.2623	-0.1122
	<b>Corr(School Score, Home School)</b>	-0.7888	-0.6016	-0.8411	-0.5895

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

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

**Table V: Calculations Illustrating Trade-offs Between Home School and School Scores, Varying Test score, Income, and Idiosyncratic Preferences**

			Ratio of Preference for Home School to Preference for School Score			
Home School Preference	Baseline Test score	Standardized Income	White No Lunch Subsidies	Black No Lunch Subsidies	White Lunch Subsidies	Black Lunch Subsidies
Average	Score=0	Income=0	1.8155	0.9633	5.3975	1.8849
Average	Score=2	Income=0	0.9323	0.5889	2.0818	0.8321
Average	Score=2	Income=100	0.5618	0.4126	--	--
1 St.dev. Below	Score=0	Income=0	0.9977	0.5380	1.3049	0.9060
1 St.dev. Below	Score=2	Income=0	0.5918	0.3398	0.7863	0.4746
1 St.dev. Below	Score=2	Income=100	0.3815	0.2418	--	--
2 St.dev. Below	Score=0	Income=0	0.5337	0.1775	0.2362	0.1753
2 St.dev. Below	Score=2	Income=0	0.3472	0.1153	0.1668	0.0977
2 St.dev. Below	Score=2	Income=100	0.2356	0.0832	--	--

**Table VI: Summary Statistics Comparing Redistricted and Non Redistricted Students**

	Overall		No Lunch Subsidies				Lunch Subsidies			
	No Redist.	Redistricted	White		Black		White		Black	
	No Redist.	Redistricted	No Redist.	Redistricted	No Redist.	Redistricted	No Redist.	Redistricted	No Redist.	Redistricted
<b>%White, Non-Lunch</b>	0.5419	0.2892								
<b>%White, Lunch</b>	0.0674	0.0461								
<b>%Non-White, Non-Lunch</b>	0.1569	0.1903								
<b>%Non-White, Lunch</b>	0.2338	0.4744								
<b>Median Income</b>	61,311	48,267	73,325	71,564	52,132	48,523	41,770	33,647	41,770	33,647
<b>Average Z-Score</b>	0.2215	-0.1870	0.6486	0.5747	-0.0344	-0.1540	-0.5144	-0.6720	-0.5144	-0.6720
<b>Percent Chose Home 1st</b>	0.6921	0.3105	0.7693	0.4070	0.5913	0.2667	0.5687	0.2734	0.5687	0.2734
<b>Percent Chose Last Year</b>	0.6609	0.4042	0.7196	0.4692	0.6141	0.4264	0.5557	0.3507	0.5557	0.3507
<b>School Percent Made 3 Choices</b>	0.3872	0.5814	0.2551	0.3727	0.5100	0.6057	0.6025	0.7019	0.6025	0.7019
<b>Score of New Home School</b>	0.0119	-0.2410	0.2065	0.1224	-0.1511	-0.2552	-0.2884	-0.4557	-0.2884	-0.4557
<b>Score of Old Home School</b>	0.0119	-0.1675	0.2065	-0.0778	-0.1511	-0.2459	-0.2884	-0.1827	-0.2884	-0.1827
<b>Average Score Difference: Old -New</b>	0.0000	-0.0733	0.0000	0.2025	0.0000	-0.0115	0.0000	-0.2713	0.0000	-0.2713

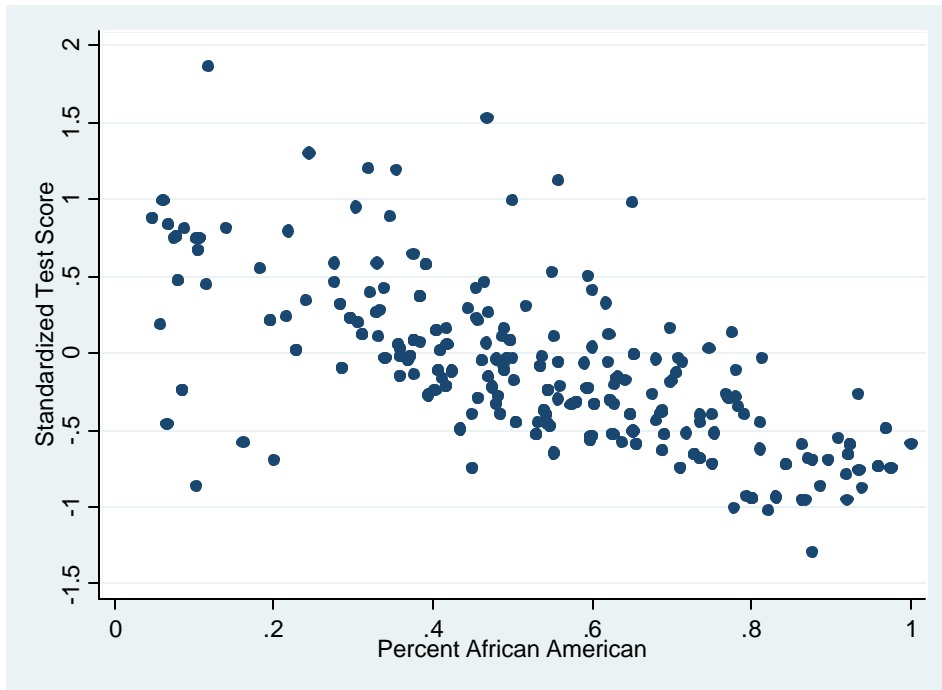
**Table VII: Mixed Logit Estimates for Redistricted Sub-sample of Students**

<i>Variable</i>	<i>Preference Parameter</i>	<b>Parameter Estimates*</b>			
		<b>No Lunch Subsidies</b>		<b>Lunch Subsidies</b>	
		<i>White</i>	<i>Black</i>	<i>White</i>	<i>Black</i>
<b>Distance**</b>	<b>Mean (normal)</b>	-1.0623	-1.4276	-1.2945	-1.5501
	<b>Std. Dev. (normal)</b>	0.4278	0.5186	0.5329	0.7173
	<b>Mean (lognormal)</b>	-0.3788	-0.2744	-0.3158	-0.2745
	<b>Std. Dev. (lognormal)</b>	0.0643	0.0418	0.0572	0.0618
<b>Last-year School</b>	<b>Mean</b>	3.3870	3.2727	3.4607	2.9128
	<b>Std. Dev.</b>	2.4192	2.7605	3.0653	3.3166
<b>Home School</b>	<b>Mean</b>	2.3134	1.8607	2.1120	1.8723
	<b>Std. Dev.</b>	0.2970	0.2790	1.4359	1.1570
<b>Choice Zone</b>	<b>Mean</b>	1.0569	1.3178	1.8295	1.6459
	<b>Std. Dev.</b>	0.7367	1.0511	1.2499	1.0220
<b>School Score</b>	<b>Mean</b>	0.9842	1.8750	0.2635	0.8725
	<b>Std. Dev.</b>	0.5708	0.1762	0.4202	0.2568
<b>Test score * School Score</b>	<b>Mean</b>	0.5371	0.4270	0.0928	0.3071
	<b>Std. Dev.</b>	--	--	--	--
<b>Income*School Score</b>	<b>Mean</b>	0.0218	0.0181	--	--
	<b>Std. Dev.</b>	--	--	--	--
<b>Percent Black</b>	<b>Mean</b>	3.2151	5.0013	0.9597	2.3609
	<b>Std. Dev.</b>	2.7598	1.7264	1.7844	1.0622
<b>Percent Black Squared</b>	<b>Mean</b>	-5.4452	-3.6580	-2.7623	-1.7100
	<b>Std. Dev.</b>	--	--	--	--
<b>Implied Mean Preferred % Black</b>		0.2952	0.6836	0.1737	0.6903
	<b>Std. Dev.</b>	0.2534	0.2360	0.3230	0.3106
<b>Estimated Correlation Coefficients:</b>					
	<b>Corr(Distance, School Score)</b>	0.1613	-0.5445	-0.3659	-0.7767
	<b>Corr(Distance, Home School)</b>	0.2464	-0.3001	-0.1858	-0.1999
	<b>Corr(School Score, Home School)</b>	-0.8672	-0.6014	-0.1025	-0.2383

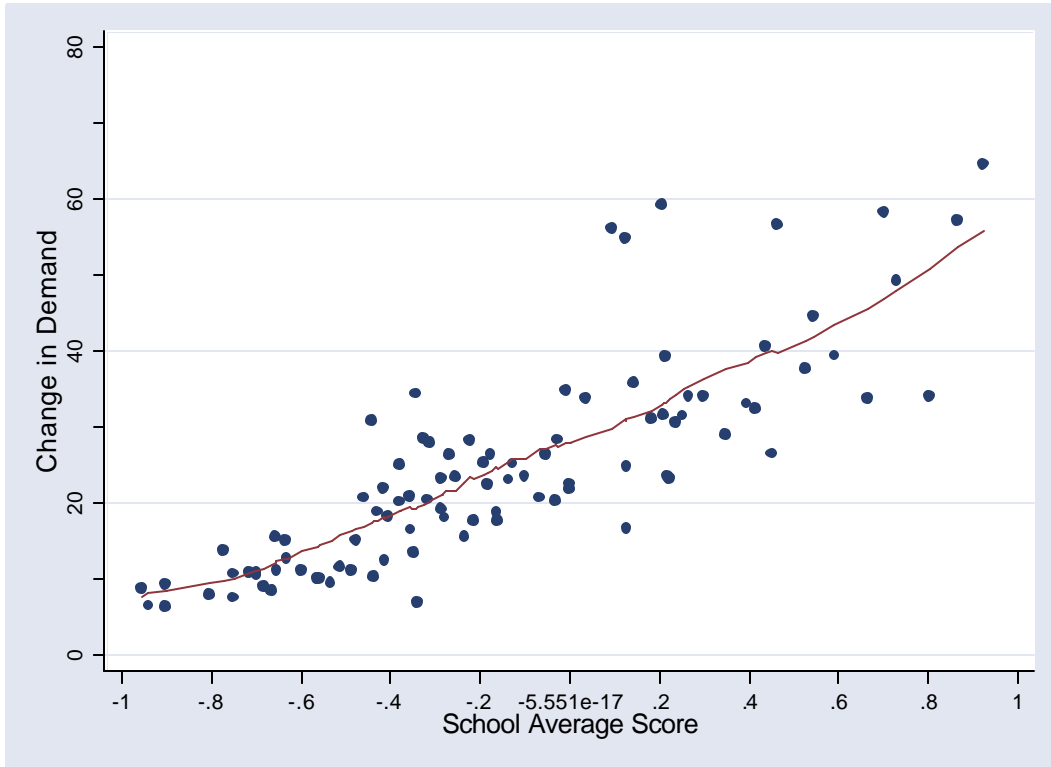
\* All estimates are significant at the 1% level or higher

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

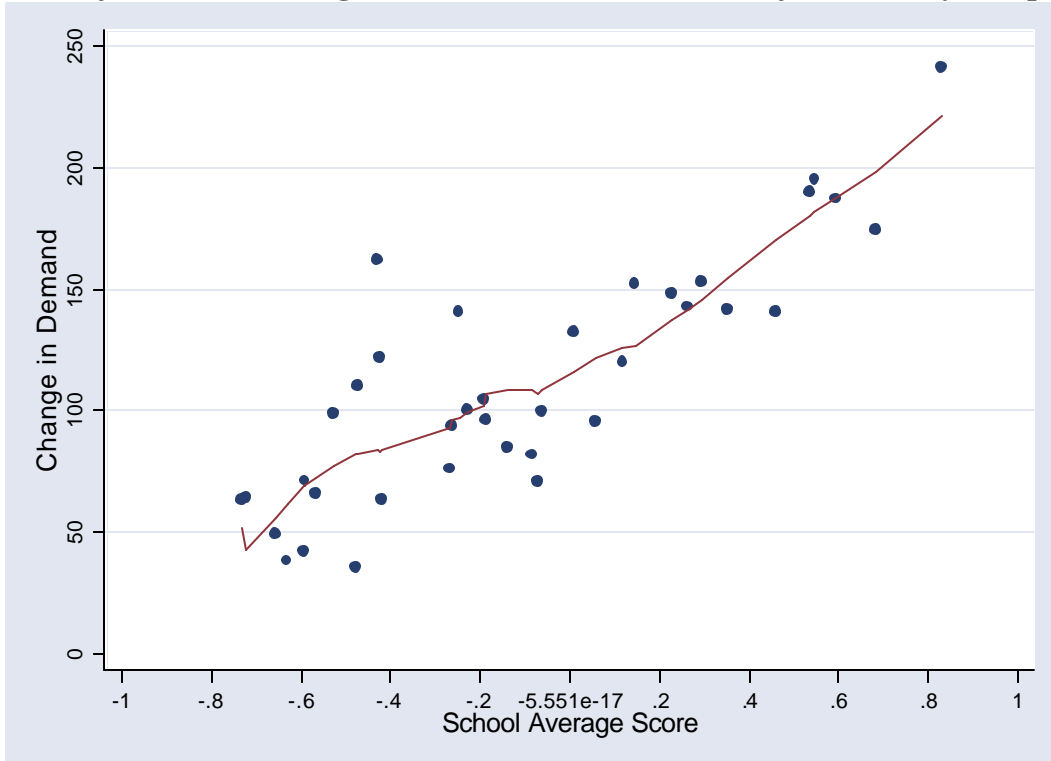
**Figures 4: Scatter Plot of Average Percentile Score of Students in a School Program versus Percent Minority in the School Program**



**Figure 5a: Elementary Schools: Simulated Change in Number of Students Choosing School  $j$  when the Average Standardized Score at School  $j$  increase by 0.33 points.**



**Figure 5b: Middle Schools: Simulated Change in Number of Students Choosing School  $j$  when the Average Standardized Score at School  $j$  increase by 0.33 points.**



**Figure 6: Percent of the Additional Students who Choose School  $j$  in Response to a 0.33 point Increase in Standardized Percentile Score at School  $j$  who qualify for Free Lunch.**



**Figure 7: Average 2002 St.Dev Scale Score for the Additional Students who Choose School  $j$  in Response to a 0.33 point Increase in Ave. Score at School  $j$ .**

