Public School Choice and Student Outcomes: Choices, Preferences and Heterogeneous Treatment Effects in a School Choice Lottery

by

Justine S. Hastings
Dept. of Economics
Yale University and NBER

Thomas J. Kane
School of Education
Harvard University and NBER

Douglas O. Staiger
Dept. of Economics
Dartmouth College and NBER

January 2006

The authors thank the Charlotte-Mecklenburg School district for making this project possible. Sean Hundtofte and Orkun Sahmali provided outstanding research assistance. We would also like to thank Joseph Altonji, Patrick Bayer, Steven Berry, David Card, Kenneth Chay, Philip Haile, Fabian Lange, James Powell and participants at various seminars for helpful comments and discussion. The project is funded by a grant from the Institution for Social and Policy Studies at Yale University and the US Department of Education

ABSTRACT

We use data surrounding implementation of a public school choice program with school assignment by lottery to estimate the effect of attending a 1st choice school on academic outcomes. We model heterogeneous treatment effects within a random utility model of school choice, and show that students who place a high value on academic achievement in their school-choice decision should have significant gains in their own academic outcomes as a result of winning the lottery to attend their first-choice school. Using random assignment to schools generated by the lottery, we estimate the impact of winning the school choice lottery on academic achievement, allowing the treatment effect of attending a first choice school to vary with the estimated preference for academics that generated the choice. The results indicate that on average, students do not have significant gain in test scores as a result of attending their first choice school, consistent with the results from prior studies. However we find that students who placed high value on academics experience significant gains in test scores. These findings are consistent with a utility-maximizing trade-off. Parents may trade-off expected gains to academic achievement for gains along other dimensions, such as proximity or racial composition.

I. Introduction

Public school choice programs are becoming an increasingly prevalent alternative to school assignment based on neighborhood location. Under public school choice plans, students can apply to attend schools other than their neighborhood school. By decoupling geographic location and schools, choice plans in principal allow students to select a school independently from housing location, thereby providing access to high quality schools for all students regardless of race, ethnicity, or income. In fact, the recent federal No Child Left Behind (NCLB) legislation requires districts receiving Title I funds to allow students in failing schools to choose to attend non-failing schools outside of their neighborhood. The goal is to allow disadvantaged students to benefit academically from attending higher performing schools in other neighborhoods.

By allowing parents to separately optimize over residential location and public schooling, school choice plans could lead to higher utility by simply relaxing binding constraints. Nevertheless, the effectiveness of school choice for public policy depends to a large extent on the effect it has on student achievement. If parents choose schools for both academic and non-academic reasons, it is not clear that school choice will lead to higher academic achievement. Indeed, recent studies of impacts of school choice find on average no affect of attending a first choice school on academic outcomes (Hastings, Kane and Staiger (2006), Cullen, Jacob and Levitt (2003), Mayer, et. al. (2002), Krueger and Zhu (2002)). When utility for schools is a function of academic and non-academic characteristics, school choice may increase utility but not necessarily academic achievement.

This paper investigates the relationship between preferences, choices, and student outcomes in the context of public school choice. Our approach explicitly accounts for heterogeneity in treatment effects generated by the underlying decision process that leads students to participate in school choice (Heckman, Smith and Clements, 1997; Heckman, 1997). We use data surrounding the implementation of a public school choice program with school assignment by lottery to estimate the impacts of attending a 1st choice school on academic outcomes. We allow the treatment effect of attending a first choice school to vary with the value parents placed on academics when making their choice. The underlying preferences for academics are estimated using a random utility model of school choice. We find that students who have strong preferences for academics experience significant gains in standardized test scores when randomized into their first-choice school.

Our data and research design come from the 2002 district-wide implementation of school choice in the Charlotte-Mecklenburg school district (CMS) in North Carolina. CMS introduced public school choice in the fall of 2002 after the former race-based busing plan was terminated by the courts. Under the choice plan, parents in the district submitted their top three choices of schools for their children, and the district assigned students to schools through a lottery system. The school district provided us with data on student's choices, student's lottery numbers, and student assignments, along with data on individual student's demographics and academic achievement for the year before and after the implementation of the school choice plan.

We use the multiple response data to estimate a random utility model of school choice with heterogeneous preferences for school characteristics. The mixed-logit discrete choice demand model yields parameter estimates of the preferences distributions for school characteristics in the indirect utility function. Using these estimated distributions as the prior, we calculate posterior estimates of the weight each parent placed on academics given their choices and their characteristics. The posterior estimates allow us to capture in one statistic how a student's choices differed from the average choices a student would make if facing the same situation. In other words, given the choices made by the student, the preference distribution in the population, her demographic characteristics and the choice set she faced, how strong were her preferences for academics?

We then use these preference estimates to show that students who place a high value on academic achievement have significant gains to their own academic achievement as a result of winning the lottery to attend their first-choice school. Using the random assignment to schools by lottery, we create treatment and control groups for estimating the impact of winning the school choice lottery on academic achievement. We allow the treatment effect of attending a first choice school to vary with the estimated preference for academics. The results indicate that on average, students do not have significant gain in test scores, consistent with the results from prior studies. However, once we allow the effect of attending a 1st choice school to vary with the estimated value placed on academics when making choices, we find that students who placed high value on academics experience significant gains in test scores. In particular students with estimated values of academics in the top quartile experienced rises in end of grade test scores of approximately 4-6 percentile points. On the other hand, students with the lowest value of academics actually experienced statistically significant declines in standardized test scores.

These findings are consistent with a utility-maximizing trade-off. If utility over schools is a function of various school attributes, and not solely academics, parents may trade-off expected gains to academic achievement for gains in utility along other dimensions, such as proximity or school racial composition. In this case, utility is maximized, but we may only see academic gains for students whose utility is strongly increasing in academics relative to other school attributes. In this setting we also may expect to see declines in academic achievement for students who chose primarily for non-academic attributes – trading of academic losses for larger utility gains on other dimensions. In a public school choice setting where parents have heterogeneous preferences over multiple school attributes, parents get what they want: those that want strong academics and increased achievement for their children are able to get it.

This paper proceeds in five main sections. The first section lays the background for the data and estimation by describing key details of the CMS School Choice Plan. The second section outlines the relationship between expected academic outcomes and preferences in a school choice plan, where parents choose schools based on expected academic achievement and other school characteristics, and students are then granted admission to schools by lottery. In the third section we generate estimates of the preferences for academic achievement. We then incorporate these preference estimates into our final estimation of the effect of attending a first choice school on academic outcomes. The final section concludes.

II. Background: The CMS School Choice Plan

A. School Choices

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 assignment plan. In December of 2001, the school board voted to approve a new district-wide public school choice plan.

In the spring of 2002, parents were asked to submit their top three choices of school programs for each child. Each student was assigned a "home school" in their neighborhood, typically the closest school to them, and was guaranteed admission to this school if they were not admitted to any of their top three choices. Students were similarly guaranteed admission to continue in magnet programs in which they were enrolled in spring 2002. Admission to non-guaranteed schools was granted based on lottery assignment as described in the next section.

The implementation of the school choice program resulted in a large redistricting of home school assignments. Prior to choice school assignment zones were discontinuous, incorporating isolated neighborhoods to achieve racial balance. When these zones were changed to 'home-school' zones, approximately 50 percent of parcels lost property rights to the school they had rights to under busing. 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. This provides an interesting environment in which to examine parental preferences and the impact of attending a first choice school on academic achievement.

The district received choice response forms for 95% of students. However, since students were guaranteed a slot in their home school, many parents filled out only one choice – their home school. 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 reports the number of choices submitted by each parent by race and free lunch eligibility. Among white free-lunch-ineligible students, about half (51%) listed only one choice on their forms. Black and free-lunch-eligible students were much more likely to fill out all three choices. Moreover, among those who were ineligible for the free lunch program, non-white students were nearly twice as likely to list all three choices relative to white students (54 percent versus 29 percent).

There are at least two reasons why white students who are not free-lunch eligible were more likely to list only a single choice. First, the average quality of their home schools is significantly higher. Table I shows the average test score, measured in student-level standard deviation units, for home schools by race and free lunch eligibility. The average scores for home schools of white and free-lunch-ineligible students are 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 on both academic quality and proximity. However, the last row in Table I shows that students in all four race-lunch-recipient categories had top a quartile school within approximately the same average distance from their home. Due to historic school placement for bussing, very good schools were within a reasonable distance for students of all socio-economic groups.

Table I also reports the fraction of students in each demographic group that listed their home school as their first choice. According to Table I, 64% of white, lunch-ineligible students chose their home school first, while only 51% listed only a first choice. This implies about a fifth of the white, lunch-eligible parents whose top choice was their home school actually provided additional rankings. About half of the black, lunch-eligible children whose top choice was their home school, provided some additional listings. Whatever their reasons for doing so, the availability of multiple choices from those who listed their home school first will further aid in the identification of the preference parameters. For the estimation using the randomization, we will use only students who did not choose their neighborhood school, and thus were assigned based on the lottery system.

B. Lottery Assignments

Admission of students to non-home choices was limited by grade-specific capacities set by the district. The district allowed significant increases in school enrollment size at high-demand schools in the first year of the choice program in an expressed effort to give each child one of their top three choices.[1]

Approximately one third of the schools in the district were oversubscribed. The district implemented a lottery system for determining enrollments in those oversubscribed schools. Under the lottery system, students choosing non-home schools were first assigned to priority groups and student admission was then determined by a lottery number. The priority groups for district schools were arranged in lexicographic order based on the following priorities:

Priority 1: Student who had attended the school in the prior year. (Students were subdivided into 3 priority groups depending upon their grade level, with students in terminal grades—grades 5, 8 and 12—given highest priority.)

Priority 2: Free-lunch eligible student applying to school where less than half the students were free-lunch eligible.

Priority 3: Student applying to a school within their choice zone.

Under the lottery system, students listing a given school as their first choice were sorted by priority group and a randomly assigned lottery number.[2] Slots remaining after home school students first choices were accounted for were assigned in order of priority group and random number.[3] If a school was not filled by those who had listed it as a first choice, the lottery would repeat the process with those listing the school as a second choice, using the same priority groups as above. However, for many oversubscribed schools, the available spaces were filled up by the time the second choice priority groups came up.

---

[1] For this reason and others, we do not find evidence of strategic hedging in the schools that parents listed. Hastings, Kane and Staiger (2005) test for strategic hedging and do not find evidence that parents with poor performing home schools chose lower-performing first choice schools with potentially higher odds of admission in an effort to hedge against being denied admission to any of the chosen schools.

[2] The random number was assigned by a computer using an algorithm that we verified with CMS computer programmers.

[3] Once any sibling was admitted to a school, other siblings could choose to attend the school. We dropped those who were admitted to a school because of a sibling preference.

Students who were not assigned one of their top choices were placed on a waiting list. About 19% of students winning the lottery to attend their first choice schools subsequently attended a different school, with 13% choosing to attend their home school instead and another 6% choosing to attend a different school entirely, with most of these students changing address. When slots became available, students were taken off the wait list based on their lottery number alone, without regard for their priority group.

III. The Marginal Impact of Choice

Our goal is to exploit the randomization in school admission based on lottery number to estimate the effect that attending one's first-choice school will have on student outcomes. More precisely, we wish to estimate the "treatment-on-the-treated" parameter: In the population of students that chose an over-subscribed school as their first choice, what was the effect of attending the first-choice school on student outcomes? Even though random assignment allows identification of the average treatment effect, we may be interested in more than the average. If heterogeneous preferences and the resulting choices lead to heterogeneous treatment effects, incorporating the factors driving decisions may yield further economic insights. For example, student who highly value academics may experience gains, and students who are willing to trade-off academics for gains on other dimensions may experience losses. These heterogeneous treatment effects may average to zero, but have very different economic and policy implications than a homogeneous treatment effect of zero.

A number of recent studies of school choice in the United States have used some form of randomization to estimate an average treatment effect, and have generally found little evidence of improved academic outcomes. Cullen, Jacob and Levitt (2003) exploited lottery outcomes to study the effect of a public school choice program in Chicago, and found no discernible impacts on various measures of academic achievement for those winning the lotteries. In general, an insignificant result may not mean that schools do not have an impact on outcomes. It may simply mean that parents were optimizing some objective other than academic achievement in their choice of school. Using the unique multiple response data we will be able to examine the mechanisms driving choice, and the implications this has for heterogeneous treatment effects.

From analysis of the data on student level choices, it is clear that parents have very heterogeneous preferences over school characteristics. Figure 1 shows that approximately 20% of students chose schools that had lower test scores than the school they had guaranteed admission to. This suggests that some parents valued these schools for reasons other than average academic achievement. In addition, among those with the same elementary home school for 2002-03, parents on average listed 10 different elementary schools as their first choice.[4] Such a diversity of choices suggests that heterogeneous preferences may play a key role in school selection, and may therefore generate differential gains in academic achievement depending on the utility trade-offs parents face.

In general, suppose that parents choose schools for both the expected academic gain for their child, but also for other reasons, such as proximity or racial composition. Consider the following utility function of parent *i* for school *j*

(1) $U_{ij} = \beta_i A_{ij} + V_{ij}$

where A_{ij} is the expected academic achievement of student *i* if she attends school *j*, V_{ij} is the utility for student *i* from attending school j along non-academic dimensions, and β_i is the weight that parent *i* places on academic achievement relative to non-academic dimensions. The utility gain from attending the first choice over the alternative school is:

(2) $\Delta U = \beta_i \Delta A + \Delta V$

where delta denotes the difference in variables between school alternatives *k* and *j*. A student will choose an alternative school over their home school only if the utility gain is positive, *i.e.* $\Delta U > 0$. Among students choosing an alternative school over their home school, the expected academic gain of a student randomized into their 1st choice school is given by:[5]

(3) $E(\Delta A \mid \beta_i \Delta A + \Delta V > 0)$

---

[4] This statistic excludes heterogeneity in choices generated solely by heterogeneity in prior-year school assignment under the bussing system. If we include choices driven by preferences for prior-year schools by students with different prior-year schools under bussing, but the same new home-school assignment area under choice, this statistic increases to 14.6.

[5] As noted earlier, the lottery was run as a 'first-choice maximizer'. Because of this, most students who did not win the lottery for their first choice school were assigned to their home school.

In this simple framework, students with high β_i have a positive expected treatment effect (gain in academic achievement from attending the first-choice school). In fact, as β_i gets very large, the expected treatment effect alone determines choice and, therefore, must be positive for all students who choose an alternative school. For a student with low β_i (near zero), the expected treatment effect is ambiguous. If ΔA and ΔV are independent and ΔA is mean zero, then the expected treatment effect is zero, i.e. $E(\Delta A \mid \Delta V > 0) = 0$. If ΔA and ΔV are negatively correlated, as is likely to be the case for proximity and scores or percent African American and scores, then the treatment effect will be on average negative for students placing a near zero weight on academic outcomes. That is, test scores of students choosing for a school characteristic that is negatively correlated with academics will tend to fall if they are admitted to their first choice school. Hence, this basic framework generates the prediction that the expected treatment effect is positively correlated with the preference for academic achievement. This correlation is stronger if parents face trade-offs – if expected academic achievement is negatively correlated with other valued school characteristics.[6]

IV. Estimating the Preferences for School Characteristics:

In order to test for heterogeneous treatment effects, we first need to estimate the underlying preferences that drive choices. This estimation proceeds in two steps. First, we estimate a random utility model of demand for schools, allowing preferences for school characteristics to vary across individuals in the population. Using these preference distribution estimates as the prior, we calculate a posterior estimate of the weight each student placed on academic achievement given the choices they made and the choice set they faced. This section describes how we constructed estimates of the weight each student placed on academic achievement. The next section uses the student assignment lottery to estimate the treatment effect of attending a first choice school, and tests

---

[6] With additional assumptions on ΔA and ΔV, one can derive stronger empirical implications from the model regarding how the treatment effect will vary. For example, if $A_{ij} = X_j\beta_i + v_{ij}$ and $V_{ij} = Z_{ij}\gamma_i + \omega_{ij}$, with v_{ij} i.i.d. normal and ω_{ij} i.i.d. extreme value, and there are no common variables in X and Z, then one can estimate $E(\Delta A|\Delta U>0)$ directly from the random utility model. We estimated models of this form and found that they performed poorly in terms of predicting the magnitude of the treatment effect, suggesting that either our assumptions were too restrictive or the individual-choice level parameter estimates required in the selection term were poorly identified given the available data.

whether the treatment effect on a student's academic performance increases with the estimated weight each student places on academic achievement.

A. Model

Let U_{ij} be the expected utility of individual *i* from attending school *j*. Assume that utility is a linear function of the academic achievement of *i* at school *j*, A_{ij}, and other school-student characteristics, Z_{ij}, such as distance from home, busing availability, and racial composition. Let ω_{ij} represent an unobserved idiosyncratic preference of student *i* for school *j*.

(4) $$U_{ij} = \beta_i A_{ij} + \gamma_i Z_{ij} + \omega_{ij}$$

Furthermore, suppose that

(5) $$A_{ij} = S_j + \upsilon_{ij}$$

The expected academic outcome for student i attending school j is the average test score at school *j*, S_j, plus υ_{ij}, a mean zero deviation that is known to the parent. We can re-write the indirect utility function as:

(6) $$U_{ij} = \beta_i S_j + \gamma_i Z_{ij} + \varepsilon_{ij}$$

where $\varepsilon_{ij} = \beta_i \upsilon_{ij} + \omega_{ij}$. Assuming that ε_{ij} follows an independent extreme value distribution, we will get the typical logit functional form for the probability of choosing school *j*. Note that estimation involves normalizing the variance of ε_{ij}. Since Var(ε_{ij}) is an increasing function of β_i, normalization will reduce the estimate of β_i for high-β_i types. However, while this will act to understate the estimated variation in β_i in the final model, it does not affect the relative rankings of individuals with respect to β_i – which is the information we use to estimate heterogeneous treatment effects.

We estimate a mixed-logit model of demand, allowing for the preferences for school characteristics to vary across individuals (Hastings, Kane and Staiger, (2005); Train (2003)). Preference heterogeneity enters in two ways: through interactions of preferences and observable student characteristics, and through idiosyncratic preferences (random coefficients). We allow the mean preference for a school's standardized score to vary with students' baseline test scores and family income. We allow all preference distributions to vary by race and free-lunch recipient status as well. Idiosyncratic variation in preferences is governed by distributional assumptions. We will assume that preference parameters are drawn from a joint normal distribution.

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

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

$$= \int \prod_{c=1}^{3} \frac{e^{X_{ij^c}\beta}}{\sum_{k \in J_i^c} e^{X_{ik}\beta}} f\left(\beta \mid \mu, \theta\right) d\beta \qquad (7)$$

We assume that $\beta \sim f\left(\beta \mid \mu_\beta, \theta\right)$, where $f\left(\cdot\right)$ is a joint-normal mixing distribution, μ denotes the mean, and θ represents the variance parameters. The term inside the integrand represents the probability of observing the three ranked choices conditional on the preference coefficients (β): this is the product of three logit probabilities evaluated at β_i, corresponding to the probability of making each choice from among the remaining options.[7] This conditional probability is integrated over the distribution of β to yield the unconditional probability of observing the ranked choices. Estimation was by the method of maximum simulated likelihood, using 100 draws of β from $f\left(\cdot\right)$ for each individual in the data set. The results were not sensitive to the number of draws used. We assume that all random parameters are drawn from a normal or log normal distributions, and allow for correlation among some of the main preference parameters as reported in the tables.

B. Specification

The explanatory variables included key observable school characteristics and trade-offs that parents consider when selecting a school. For proximity measures we included the travel distance (measured in miles along roads) from each student to each school, an indicator if the student was eligible for busing to the school, and an indicator if

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

the school was designated as the student's neighborhood school. We included an indicator variable for the school the child attended in the prior year to capture preferences for staying at the same school, which may be particularly important for terminal-grade students. 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. 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).[8] Table II lists the dependent variables in the indirect utility function and describes how they were constructed.

We allow the mean preference for academic achievement (the coefficient on school test scores) to vary linearly with the student's standardized baseline test score (from the spring of the prior year, 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-meaned with the countywide median of $51,000). Preferences for distance are constrained to be negative, following a lognormal distribution. We allow preferences for proximity and academic quality to be correlated. All other preference distributions are assumed to be independently and normally distributed. We estimate the parameters of the preference distribution separately by race and lunch-subsidy status. This permits the preference distributions and logit-normalization to vary across the four socio-economic categories.

C. Identification and Results from Mixed Logit Estimation

Before discussing the results, it is helpful to note some key aspects of the data that will aid in identification of preference parameters. First, the large scale redistricting that occurred with the introduction of school choice helps identify preference parameters separately from residential sorting. In addition, multiple choices listed by those selecting their home school first further separates preferences for school characteristics from

[8] We use the average test scores at the end of the first year of choice instead of those at the end of the year before school choice was implemented. We compared these two specifications and found that the post-choice test scores better explained parent's choices over schools as a function of academic quality.

residential sorting by simulating the unavailability of the neighborhood school. Without redistricting and the multiple-choice responses, residential sorting would potentially confound the preferences for proximity with preferences for other desired school attributes. [9]

Multiple choices also help identification of the idiosyncratic variation in preferences (variance of the random coefficients). Intuitively, when only a single (1st) choice is observed for every individual, it is difficult to be sure whether an unexpected choice was the result of an unusual error term (ε_{ij}) or unusual preferences by the individual (β_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. In addition, because CMS is a fairly integrated district, with a sizable non-white middle-class, a sizeable low-income white population, and historic placement of schools for racial integration, there is a weaker correlation between average test scores, percent African American, and neighborhood location of top-tier schools. This provides differential variation in proximity, academic achievement and school racial composition in the choice set for students of all socio-economic groups.

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.

Table IV presents the results from the mixed logit demand estimation by race and lunch-recipient status. All of the point estimates were precisely estimated and statistically different from zero at less than the 1 percent level. We report the estimates for the means, standard deviations, and correlation coefficients (where appropriate) for the preference distributions. The discussion of results is focused around the parameters most relevant for our final estimation of the effect of attending a first choice school on academic

[9] For a comparison of preference estimates for the redistricted sub-sample versus the full sample, please see Hastings, Kane and Staiger (2005).

achievement. For a further discussion of the results and their implications for student sorting and competition on quality in public school choice, please see Hastings, Kane, and Staiger (2005).

The first four rows of coefficients in Table IV report the preferences for school test scores by race and lunch-recipient status. The first row of coefficients reports the mean preference for school scores for the average student. It is positive for all four demographic groups, implying that school test scores have a positive effect on choosing a school for the average student. For a student with average baseline test scores and average income, the mean preference for school scores is larger for non-white students (1.80) than for white students (1.17) among students not receiving lunch subsidies. Preference for school scores among students receiving lunch subsidies are lower for both whites and nonwhites, but the difference between whites and nonwhites is similar.

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-subsidy status, we included interactions between school test scores, student's baseline test scores, and neighborhood income level.[10] The third and fourth rows of parameter estimates report the coefficient on the interaction of school scores with income and the student's baseline score respectively. Recall that both neighborhood income and the student's baseline score are "de-meaned", 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 coefficients on the income interaction imply that mean preferences for academics are increasing with income. The magnitudes of these parameters are roughly consistent with the differences in the mean preferences for test scores between lunch-recipients and non-lunch recipients within race. Similarly, the mean preference for school scores is increasing in the student's baseline test score. The coefficient on the interaction between the student's baseline 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 effect of a student's baseline score on the preference for

[10] 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. In initial specifications using a conditional logit, income interactions with the preference for school scores were generally insignificant for the lunch-recipient segments.

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 $25,000) is associated with a 0.3-0.4 increase in the mean preference for school test scores.

The coefficients on the interactions of income and baseline score with school scores demonstrate that there is considerable observable heterogeneity in preferences for academics. Parameter estimates for the standard deviation in idiosyncratic preferences for academics are reported in Row 2. While 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, the estimated standard deviation in idiosyncratic preferences for school test scores is also around 0.3 for non-whites and 0.65 for whites. Hence, there is substantial unobserved heterogeneity in preferences for test scores. Taken together, the coefficients imply that academics are on average a significant determinant of school choice. Furthermore, the substantial variation across students in the weight placed on academics suggests that we may expect to see strong school choice selection on academic outcomes for some students and not for others. The fact that much of the heterogeneity in preferences is unobservable implies that the traditional approach of allowing the treatment effect to vary with observable characteristics, such as race or lunch status, may not completely capture heterogeneous treatment effects by preferences for academics.

The parameter estimates for the remaining coefficients indicate that parents face important trade-offs between academic and non-academic factors when choosing schools. Rows 5 and 6 report the parameter estimates for the lognormal distribution of preferences for distance. Rows 7 and 8 report the mean preference and standard deviation for the neighborhood (or 'home') school.[11] Parents dislike distance and prefer their neighborhood school. These coefficients indicate that parents on average must trade-off utility for proximity in order to gain utility from expected academic outcomes. For most

[11] Hastings, Kane and Staiger (2005) discuss the interpretation of the neighborhood school. They test if this coefficient represents a non-linearity in the preference for proximity or if it is potentially consistent with a default effect. They provide evidence that the preference for the neighborhood school is a neighborhood preference that is not generated by default behavior.

students, attending a high-achieving school will require them to choose a school that is farther than their home school and a school that is not their home school. Hence there is a negative correlation between school characteristics that measure proximity and those that capture academic achievement. This implies that parents of all races must, on average, trade-off utility for academic gains against utility gains for proximity.

In addition to trading-off proximity for academics, African American parents must trade-off utility for the racial composition of peers. The preference coefficients on percent black imply that the average African American parent prefers schools where approximately 70% of the student body population is black. However, the percent black at a school is negatively correlated with average test scores (correlation is around -0.65). Recall that the district as a whole is approximately 45% African American. The negative correlation between test scores and racial composition implies that African American parents must value academic achievement much more than their white counterparts in order to induce them to choose a higher performing school that also has, on average, fewer African American students. Given the coefficients for the quadratic term in racial preferences, the loss in utility for black families is highest when percent black is low (less than 40%), which is precisely the range in which school average test scores are highest.

Overall, the preference estimates imply that parents face significant trade-offs between expected academic outcomes and other school characteristics that they value. Furthermore, there is significant variation in the value placed on academics and the magnitude of utility trade-offs parents have to make. We use these estimates of the preference distribution with the choice data to generate estimates of individual-level preferences for academic achievement. These in turn will be used to determine how utility trade-offs and heterogeneous preferences affect student-level gains in test scores as a result of attending a first-choice school.

D. Posterior Estimates of Individual Preferences for Academics

Given the preference distribution estimates, we use the multiple choice data to estimate what preferences for each student must have been given the choices they made and the choice set they faced. These estimates will be used to test for heterogeneous treatment effects in the next section. Using the estimated preference distributions in Table

IV, we calculate posterior estimates of the weight each student placed on school scores in the following way:

$$(8) \qquad E(\beta^A{}_i \mid y_i, X_{ij}) = \frac{\int \beta P\left(y_i \mid X_{ij}, \beta\right) f(\beta) d\beta}{P\left(y_i \mid X_{ij}\right)}$$

Where y_i denotes the choices the student made, X_{ij} denotes the student and school characteristics that enter the indirect utility function, and β_i^A represents the weight the student placed on school test scores (including the estimated effect of income and student baseline scores). This equation simply calculates the expected value of student *i*'s preference for academics given the choices she made, the characteristics of the schools given her location and characteristics. Equation 8 was then simulated for each student using 1000 draws from the estimated preference distributions in Table IV.

Intuitively, the estimated weights that each parent placed on school scores were generated from the trade-offs each student made when listing a school in each subsequent choice. These estimates are identified by the multiple choices observed for each student. If one observes multiple choices for this student, and sees the student consistently bypass closer, lower-performing schools (for example) in order to chose high-performing schools, it becomes most likely that the observed choices, given the choice set, could only have been generated by a high draw on β_i^A. The posterior estimate calculates the most likely value of β_i^A the student must have had to generate the observed choices, relative to the choices a student with average preferences would have made if facing the same X_{ij}.[12]

Thus estimating β_i^A allows us to succinctly and flexibly incorporate all of the relevant choice information for each student into one statistic – the estimated value the student places on a school's academic performance. We incorporate these posterior estimates of the weight placed on academics into our estimation of the treatment effect of attending a first choice school.

[12] See Train (2003) p. 270 for Monte Carlo Simulations of the accuracy of individual-level parameter estimates and the number of observed choice situations.

V. Estimating the Impact of Attending a 1st Choice School on Academic Achievement

In this section we estimate the causal relationship of attending ones 1st choice school on academic achievement, allowing the treatment effect to vary with posterior estimates of the weight each student placed on academics in her school choice. We use the randomization of students by lottery to their 1st choice school to generate exogenous treatment and control groups. In order to estimate the causal relationship using randomization by lottery, we focus on the subset of students choosing schools that were over-subscribed. We then limit our sample to the marginal priority groups within those schools for whom lottery number alone determined initial admission. Throughout most of the analysis, we will ignore members of priority groups in which all students were either admitted or denied admission—since the assignment of lottery numbers had no impact on their admission status. In some schools, the marginal priority group will consist of those who attended the school the year before, or free-lunch eligible students, or students from the choice zone. The marginal priority group may also be different for different grade levels in a school.

Within the marginal priority groups, we would like to estimate the impact of attending a first-choice school on academic achievement, allowing the effect to vary with the estimated individual preferences for academics. Since not all of those who won the lotteries actually chose to attend their first choice school, and some of those who lost the lotteries were subsequently admitted off the waiting lists, we use winning a lottery as an instrumental variable to estimate the impact of attending one's first choice school in following regression:

$$Y_{ij} = X_i\alpha + \gamma_1 Attended1stChoice_{ij} + \gamma_2 Attended1stChoice_{ij} * \hat{\beta}_i^A + \delta_j + \varepsilon_{ij} \quad (10)$$

Winning the lottery and winning the lottery interacted with $\hat{\beta}_i^A$ serve as instruments. Note that all of the information used to derive the preference weights was observed prior to randomization. Since $\hat{\beta}_i^A$ depends only on baseline data that is independent of whether the student won the lottery, its interaction with winning the lottery is a valid instrument once one has conditioned on baseline data. We include as regressors: $\hat{\beta}_i^A$, gender, race/ethnicity, free lunch status, home school dummy variables, baseline test scores, income, absences, suspensions, and grade retentions. Random assignment by lottery implies that the impact of winning the lottery, γ_1, is consistently estimated even without these control variables, but the additional control variables greatly improve precision.

The fixed effects, δ_j, are included for each school and grade, to account for the fact that the probabilities of winning the lottery varied across lotteries. We report robust standard errors, allowing for correlations in outcomes among students with the same first-choice school (which may include more than one grade with a lottery). As long as winning the lottery has an impact on student outcomes only through the likelihood that one attends a first choice school, then the IV estimates of γ_1 and γ_2 using the lottery outcome and its interaction with $\hat{\beta}_i^A$ as instruments, will be consistent estimates of the impact of attending one's first choice on various outcomes.[13]

In equation (10) the effect of attending one's first choice school is $\gamma_1 + \gamma_2 \hat{\beta}_i^A$. If the dependent variable is the student's own test score, we expect $\gamma_2>0$, implying that students who place more weight on test scores experience a larger treatment effect. The parameter γ_1 gives the treatment effect for a student that places no weight on test scores in their school choice decision, and could in principal be negative as such a student would trade off other school attributes for lower test scores.

A. Lottery Data and Characteristics of the Randomized Subpopulation

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 over 95% of all the students enrolling that fall. After dropping students who were not in grades 4-8, who had special disabilities needs, and students who were admitted because of siblings, we were left with a sample 37,115. Of these, 22,872 listed their guaranteed home school (n=19,669) or magnet continuation school (n=3,203) and, therefore, were not subject to randomization. Another 7,583 students were in groups sufficiently high on the priority list that they were not subject to the randomization. There were 3,065

[13] After the initial lotteries some students were taken off the waitlist according to lottery number. Adding the waitlisted students to our sample (in addition to the marginal priority groups), we estimated specifications similar to (2) and (3) above, using as instruments both whether or not a student won the lottery and the randomly assigned lottery number interacted with being placed on the waitlist. The results were quite similar to the results we report.

students in marginal priority groups, described above as those priority groups within the schools where slots were allocated on the basis of a random number. Finally, there were 3,595 students in priority groups that were sufficiently low on the priority list that all members of the priority group were denied admission and placed on the waitlist.

Our outcome measures include data on student absences, suspensions, and standardized test scores for all students in the district for the years surrounding the implementation of the choice program. Because students in kindergarten through 2nd grade do not take the state exams, and because high school students only take the end-of-course exams in the subjects they choose, we had reading scores only for students in grades 3 through 9 and math scores for students in grades 3 through 8.[14] We standardized the test scores by grade level and year to have mean zero and a standard deviation of one. In addition the testing data in North Carolina also include student self-reports on the number of hours of home work they did each week.

B. Empirical Results: Summary Statistics

We focus on the 3,065 sample members in the marginal priority groups that were subject to the randomization. Table V compares descriptive statistics on the baseline characteristics for these students and for students in the district as a whole. Students in the marginal priority group were more likely to be African American, and more likely to receive lunch-subsidies, reflecting their higher probabilities of choosing non-guaranteed schools and the priority boosts given to these students. Students in the randomized group also tended to have lower test scores, higher absences, and more suspensions, and have home schools with lower average test scores and higher percent minority. It is important to keep these differences in mind, since we are only able to estimate the impact of the school choice program for the population of students in the randomized group.

In order to verify the validity of the randomization of lottery numbers, we examine the baseline characteristics of lottery winners and losers within the randomized group. Table VI reports these baseline characteristics for our estimation sample. Our

[14] For grades 3 through 8, we used math and reading scores on North Carolina end-of-grade exams. For the grade 9 reading score, we used the student scores on the test given at the end of English I, since over 90 percent of freshmen took the exam. The equivalent math test, Algebra 1, is taken by 8th, 9th, and 10th graders depending on their level of advancement through math courses. Hence, a math score measure is not as easily identified as the reading score measure for 9th graders.

estimation sample excludes 181 students who were in marginal priority groups but missing needed baseline characteristics (such as address, which was used in the choice model). The table reports unadjusted differences, as well as differences from an OLS regression including fixed effects for the school program and grade for which the lottery is being conducted. Before adjusting for lottery block fixed effects, there are some differences in baseline characteristics between lottery winners and losers (although none are statistically significant). However, these differences were largely due to a correlation between the characteristics of lottery participants and the lottery odds. After including a fixed effect for each school program and grade, all such differences were smaller and were not significantly different from zero.

C. Impact of Winning Lottery on the Characteristics of School Attended

Before presenting the effects of winning the lottery on student outcomes, it is important to show the effects of winning the lottery on the characteristics of the school attended as a first stage analysis. Table VII reports the results of winning the lottery on the characteristics of the school attended, based on OLS estimates of the following equation:

(11) $$Y_{ij} = X_i\beta + \gamma WonLottery_{ij} + \delta_j + \varepsilon_{ij}$$

In these regressions, we control for student baseline characteristics, and home school and choice-grade fixed effects. The regression results give the average impact of winning the lottery on the characteristics of the school attended ended, Y_{ij}. The first row of estimates in Table VII shows that lottery winners were 53 percentage points more likely to attend their first choice school than the lottery losers. This is the first stage regression for the instrumental variables regression of the impact on test scores of attending a first choice school. This estimate is not equal to 100 percent for two reasons: first, some of those who were given the opportunity to attend their first choice did not do so and, second, some of those who were originally waitlisted at their first choice were subsequently called off the waitlist. Overall, approximately 75% of the winners attended their first choice and 25% percent of the lottery losers did.

The second row of estimates in Table VII show the effect of winning the lottery on whether or not the student was enrolled in CMS in the 2002-2003 school year. This estimate gives the differential attrition rate between lottery winners and losers. The coefficient is small in size (-0.018) and not significantly different than zero, indicating that there was no significant differential attrition by the end of the 2002-2003 school year. Rows three and four indicate that winning the lottery was associated with approximately one-tenth of a student-level standard deviation increase in the average combined reading and math scores at the attended school. In addition, winning the lottery implied that students attended a school with significantly lower concentration of free-lunch recipients.

F. The Effect of Attending a First-Choice School on Student Outcomes

As stated earlier, in order to estimate the marginal impact of allowing parents to switch to their first choice school, we used an indicator of whether or not a student won their lottery as an instrument for attending one's first choice school to estimate:

$$(12) \quad Y_{ij} = X_i\alpha + \gamma_1 Attended1stChoice_{ij} + \gamma_2 Attended1stChoice_{ij} * \hat{\beta}_i^A + \delta_j + \varepsilon_{ij}$$

Estimates of a homogeneous treatment effect (equation 12 with γ_2) for various student-level outcome measures, Y_i, are reported in Table VIII. [15]

The estimates in Table VIII are broken down by academic and non-academic outcomes. For non-academic outcomes we include the impact of winning the lottery on absences, suspensions, retentions, and homework time. Among these outcome measures, the average treatment effect is significant and negative for retention rates. Winning the lottery to attend a first choice school causes a dramatic reduction in retentions – a 2.3 percentage point decrease off of an average base of 2.2%. We do not find a significant impact on absences or suspensions, however. In addition, we find that students who are randomized into their first choice school report spending more time on homework. The outcome measure is an indicator if the student reports spending more than 3 hours per week on homework on self-reported surveys given to students with the end of grade exams. Even though students who attend their first choice school report a significant increase in homework hours, we find no measurable average effect on standardized test

[15] Some readers may prefer to see the reduced form impact of winning the lottery on various student outcomes. Recall from Table V that lottery winners were roughly 50 percentage points more likely to attend their first choice school than lottery losers. To obtain a rough estimate of the reduced form impact of winning the lottery, simply divide the estimates of γ_2 and its standard error in Table VI by 2.

scores. The final row of estimates in Table VIII shows no significant impact of attending a first choice school on standardized test scores.[16] The point estimate is nearly zero, but there is a relatively large standard error. The results are consistent with the current literature, and while they exploit randomization into first choice schools to create credible counterfactual comparisons, the average treatment effect may mask important heterogeneity arising from how utility maximizing parents select a first choice school.

G. Heterogeneous Treatment Effects on Test Score Outcomes

Table IX incorporates the weights placed on academic achievement when choosing a school into the estimated impact of attending a first choice school on standardized test scores. The coefficients imply that the effect of attending one's first choice school on a student's test scores is significantly increasing with the weight that a student placed on test scores in choosing a school. The estimated weight placed on academics displays significant heterogeneity across students, ranging from just below zero at the 1st percentile to just over 4 at the 99th percentile, with a mean of 1.3 and a standard deviation of 0.8. The regression estimates imply that a one standard deviation increase in the weight that an individual places on school test scores raises the treatment effect on the student's own test score by 0.062 standard deviations. For students who place no weight on test scores in their school choice, the coefficient on attending one's first-choice school implies a negative (although not significant) treatment effect – their test scores fall by 0.10 standard deviations if they attend their first-choice school. These estimates imply a small negative impact (-0.003 standard deviation score gain) of attending a first-choice school on test scores for an average student with a $\hat{\beta}_i^A$ of 1.3, and a large positive effect on test scores (about 0.10-0.20) for a students in the top decile of the $\hat{\beta}_i^A$ distribution. To understand the economic significance of these estimates, a 0.1 standard deviation increase in a student's test score results is equivalent to a 3-4 percentile rank gain in test scores. Child development psychologists suggest that a 5 percentile rank gain in a student's test score translates into a significant cognitive gain in academic aptitude. Alternatively, estimates of the impact that test scores have on future

[16] Regression estimates show the same effect on math and reading scores when run separately, so we use the combined score to improve precision.

earnings suggest that a 0.1 standard deviation in increase in test scores is worth $10,000 to $20,000 in net present value of future earnings (Kane and Staiger, 2002).

These estimates are consistent with our general prediction that heterogeneous treatment effects should be positively related to the weight placed on academic achievement derived from a random utility model of school choice. An alternative and less structural approach, often used in the literature, is to allow for heterogeneous treatment effects to vary by observable characteristics, such as race, gender and free-lunch status. These observable measures may be crude proxies for underlying preferences, and as such be related to treatment effects. But by developing a demand model and estimating preferences, we are able to create a single index, $\hat{\beta}_i^A$, that incorporates a complex set of data that determine choices. It incorporates not only the student's characteristics, but also the characteristics of her choice set as well as the preference distribution in the population of students.

Table X shows that the average value of $\hat{\beta}_i^A$ captures observed variation in the average treatment effect across subsets of students defined by their race, income, free-lunch recipient status, and baseline test score. We also estimate treatment effects for subsets of students who differ in whether their first choice school was an academic magnet school or whether their first choice school had average test scores above the median school in the district. These characteristics of the first-choice school are rough measures of the value a student placed on academics in selecting their school. Cullen, Jacob and Levitt (2003) present results by this cut, yet do not find that it generates significant estimates of the average treatment effect of attending a first choice school.

Overall, Table X shows that the difference across these subsets of students in their average preference for academics ($\hat{\beta}_i^A$) is strongly related to the estimated treatment effect of attending a first-choice school on the student's combined test score. For example, the average weight placed on test scores is above the overall average of 1.33 for students who are white, have income above the median, are not eligible for free lunch, have above average baseline scores, or whose first-choice school is either an academic magnet or has above median test scores. In all of these subsets of students, we estimate positive effects of attending the first-choice school on academic achievement – although in only a few of the categories are the effects statistically significant (white and above median income). In all the remaining subsets of students, the average weight placed on test scores is below the overall average of 1.33, and the estimated treatment effect is negative (although never statistically significant).

The advantage of relying on estimates of $\hat{\beta}_i^A$ to identify heterogeneous treatment effects rather than ad hoc distinctions such as race and income is two-fold. First, summarizing the differences across many subsets of students with a single index increases the precision with which we can identify heterogeneous treatment effects, because the index incorporates all the information on the factors that determined school selection. They incorporate not only the score of the school, but also the choice set the parent faced, distinguishing between a student who picked a good first choice school because it happened to be close from a student who consistently picked top academic schools despite their distance because of a high underlying value for academics. In contrast, making many ad hoc comparisons ignores this type of information and will tend to obscure the underlying relationship with multiple tests, often leading to treatment effects that are jointly insignificant despite a few individually significant coefficients. The second advantage of relying on estimates of $\hat{\beta}_i^A$ is that to the extent that they capture the underlying structure of preferences, they can be used to evaluate the impact of school choice outside of the estimation sample, where the association between preferences for academics and proxies such as race and income may be quite different. Using the $\hat{\beta}_i^A$ not only helps identify who benefits academically from school choice, but it also tells us why.

Finally, while students with high $\hat{\beta}_i^A$ should have a positive expected treatment effect (gain in academic achievement from attending the first-choice school), the treatment effect for a student with low $\hat{\beta}_i^A$ (near zero) is theoretically ambiguous. Our prediction is that the expected treatment effect will be positively correlated with the preference for academic achievement, and this correlation will be stronger if parents face trade-offs – if expected academic achievement is negatively correlated with other valued school characteristics. Since the percent black at a school is negatively correlated with average test scores in CMS schools (correlation is around -0.65), the racial composition of a school is an important trade-off that many African American parents face. We estimate (from the mixed-logit results) that the average African American parent prefers schools where approximately 70% of the student population is black. Parents that prefer a school with a high proportion of students African American must value academic achievement more in order to induce them to choose a higher performing school that also has, on average, fewer African American students. Thus, all students with strong academic preferences (high $\hat{\beta}_i^A$) will have a positive gain in academic achievement from attending the first choice school, but among students with weak academic preferences (low $\hat{\beta}_i^A$) we might expect a negative treatment effect among students that prefer a school with a high proportion African American. In other words, the interaction effect between $\hat{\beta}_i^A$ and winning the school choice lottery should have a negative intercept and a steeper slope for students who have strong preferences for predominantly African American schools.

Table XI presents the results from specifications identical to those in Table IX, but estimated separately for students who prefer a school that is less than 55 percent black (primarily white students) and students who prefer a school that is more than 55 percent black (primarily non-white students). Posterior estimates of student-level preferences for school racial composition were calculated in the same way as the $\hat{\beta}_i^A$'s were. The average treatment effect is positive for students who prefer a predominantly white school, and there is no significant interaction with the weight that the student places on test scores in their school choice. For these students, both high and low $\hat{\beta}_i^A$ students experience academic gains from attending their first choice school. In contrast, for students who prefer a predominantly black school there is a significant interaction between their estimated preference for academics and the treatment effect. High $\hat{\beta}_i^A$ students experience academic gains from attending their first choice school that are similar to students who prefer a predominantly white school. In contrast, low $\hat{\beta}_i^A$ students who prefer a predominantly black school experience a negative effect on academic performance from attending their first choice school. This evidence suggests that the relationship between preferences and treatment effects may depend importantly on the trade-offs that parents face given their preferences and their school choice options. These results also highlight the potential importance of the underlying decision-making process to understanding the heterogeneous impacts that public school choice has on student academic outcomes.

VI. Conclusion

When given the choice to attend a public school other than the home school to which they have been assigned, the parents of 49 percent of the students in Charlotte took the opportunity and listed a school other than their assigned school as a first choice. In this paper, we evaluate the impact of switching schools on various academic and non-academic outcomes. On average, among those applying to the oversubscribed schools, winning the lottery had no discernable impact on students' own reading and math scores overall, even though lottery winners attended schools with higher math and reading scores than did lottery losers. Winning the lottery had modest impacts on other outcomes, however, such as increasing homework time and reducing grade retentions.

However, parents seem to choose schools for many different reasons. Indeed, one quarter of parents who were willing to switch chose schools with lower mean test scores than their assigned schools. Overall, the results presented in this paper imply that the effect of attending one's first choice school on academic outcomes is significantly increasing with the value that a student placed on test scores in choosing a school. Among students placing a high weight on school test scores, there was a positive effect of attending their first-choice school on their own test score. In contrast, for students who placed a low weight on school test scores we found negative effects of attending their first-choice school on their own test score. This is consistent with utility-maximizing trade-offs. When making choices over schools, parents face trade-offs between academic and non-academic characteristics: school scores versus proximity, school scores versus racial composition. Parent's placing high weights on academics will be willing to sacrifice on other dimensions to choose higher-scoring schools. When they do so their children gain on academic outcomes. Parents who don't weigh academics so highly will be willing to trade off expected gains in scores for other school characteristics. Their utility might be maximized at a school where their child's scores might decline.

In school choice, parents appear to get what they want. If parents want improved academic outcomes, they are able to get them. If parents value other school attributes, and are willing to trade off academic gains for utility gains on other dimensions, school choice will allow them to make decisions that may maximize utility but not academic achievement. More broadly, the success of school choice programs depends critically on how schools respond to parental demand along these academic and non-academic dimensions (Hoxby, (1999), Hastings, Kane and Staiger (2005)). Ultimately, the trade-offs that parents are willing to make along these dimensions in choosing a school for their children will determine whether school choice programs will be successful at improving academic achievement.

References

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

Hastings, Justine S., Thomas Kane and Douglas Staiger, "Gender, Performance and Preferences: Do Girls and Boys Respond Differently to School Environment? Evidence from School Assignment by Randomized Lottery," forthcoming *American Economic Review,* May 2006 (Papers and Proceedings).

Hastings, Justine S., Thomas Kane, and Douglas Staiger, "Parental Preferences and School Competition: Evidence from a Public School Choice Program," *National Bureau of Economic Research Working Paper* No.11805, November 2005.

Heckman, James, "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used In Making Program Evaluations," Journal of Human Resources, 32(3):441-462, 1997.

Heckman, James J., Jeffrey Smith, and Nancy Clements, "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts," Review of Economic Studies, 64:487-535, 1997.

Hoxby, Caroline M., "The Effects of School Choice on Curriculum and Atmosphere," in Susan Mayer and Paul Peterson, eds., *Earning and Learning: How Schools Matter*. Washington, D.C.: Brookings Institution Press,1999.

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.

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.

Revelt, David and Kenneth E. Train, "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *Review of Economics and Statistics*, Vol. 80 (4): 647-657, Nov. 1998.

Rouse, Cecilia. E. 1998. "Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics* 113(2):553-602.

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

Figure 1: Distribution of Difference in Average Standardized School Score Between Student's First Choice School and Home School.

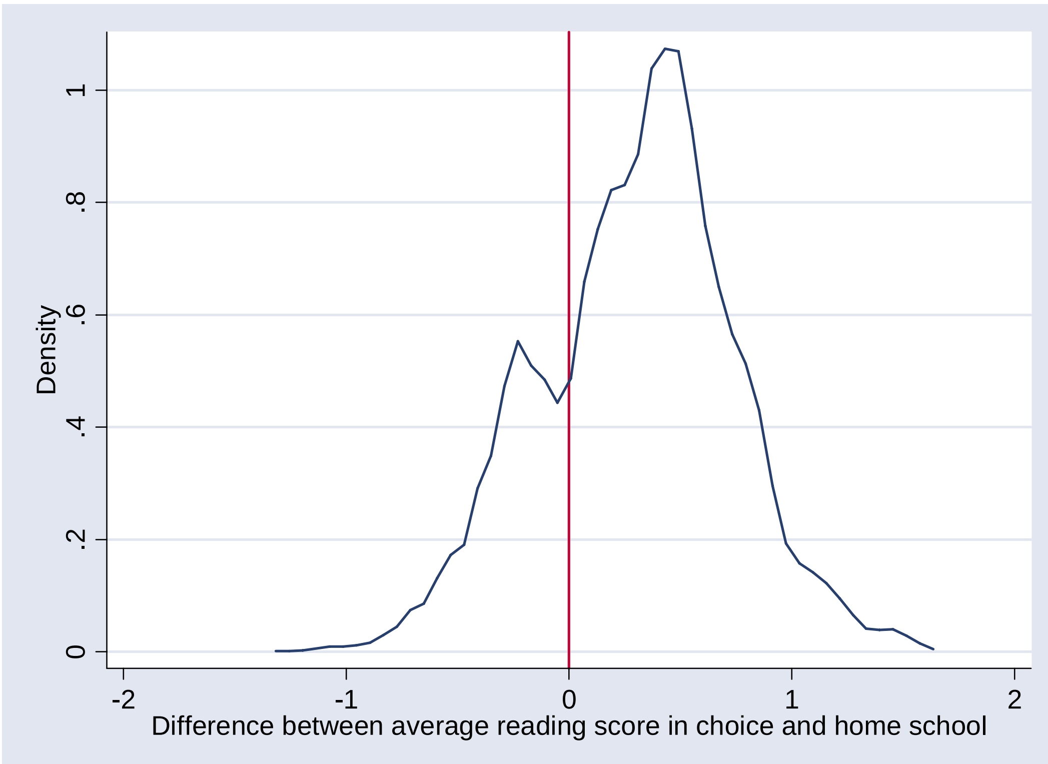

Table I: Summary Statistics of Students and Choices

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

Table II: Explanatory Variable Definitions

Variable	Description
Distance	Driving distance from student *i* to school *j* calculated using MapInfo with Census Tiger Line files.
School Score	Average of the student-level standardized scale score for students in school *j* 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 *j*.
Test Score	The sum of student *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.
Income	The median household income reported in the 2000 Census for households of student i's race in student i's block group. Income is demeaned by the county-wide average of approximately $51,000 and is reported in thousands of dollars.
Percent Black	The percent of students in school *j* who are black according to 2002-2003 school year administrative data.

Table III: Explanatory Variable Summary Statistics

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

Table IV: Estimates from Mixed Logit Model

		Parameter Estimates*			
		Not Receiving Lunch Subsidies		Receiving Lunch Subsidies	
Variable	**Parameter**	**White**	**Black**	**White**	**Black**
Preferences for Scores					
School Score	**Mean**	1.1732	1.8035	0.3671	0.9396
	Std. Dev.	0.5674	0.2688	0.6175	0.2706
Income*School Score	**Mean**	0.0151	0.0126	--	--
	Std. Dev.	--	--	--	--
Baseline own score * School Score	**Mean**	0.5558	0.5734	0.2924	0.4995
	Std. Dev.	--	--	--	--
Preferences for Proximity					
Distance**	**Mean**	-0.3526	-0.2684	-0.3784	-0.2751
	Std. Dev.	0.0684	0.0413	0.1273	0.0639
Home School	**Mean**	2.1300	1.7373	1.9816	1.7710
	Std. Dev.	0.5130	0.6799	0.8248	0.7752
Preferences for Race					
Percent Black	**Mean**	3.3068	5.1340	1.9268	3.1409
	Std. Dev.	2.6417	1.6447	2.0795	0.8745
Percent Black Sqaured	**Mean**	-5.4580	-3.6790	-3.5385	-2.3005
	Std. Dev.	--	--	--	--
Implied Mean Preferred % Black		0.3029	0.6977	0.2723	0.6827
	Std. Dev.	0.2420	0.2235	0.2938	0.1901
Other Preferences					
Last-year School	**Mean**	3.7941	3.3837	3.5016	2.8495
	Std. Dev.	2.4977	2.7896	3.4651	3.3825
Choice Zone (busing)	**Mean**	1.1909	1.2484	1.9203	1.6132
	Std. Dev.	0.8285	1.2418	1.5083	1.2442
Estimated Correlation Coefficients:					
Corr(Distance, School Score)		0.4939	-0.1055	0.3379	-0.6355
Corr(Distance, Home School)		-0.0788	0.0007	-0.2623	-0.1122
Corr(School Score, Home School)		-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: Comparison of Student Characteristics

	All Students	Chose Guaranteed School	Chose Non-guaranteed School		
			Admitted	Randomized	Waitlisted
Student demographics					
Black	44.3%	34.6%	62.5%	59.7%	54.8%
Free or reduced lunch	39.2%	31.3%	60.3%	51.3%	34.3%
Student's prior year performance					
Reading test score (SD units)	0.02	0.15	-0.26	-0.09	-0.11
Math test score (SD units)	0.02	0.16	-0.26	-0.12	-0.15
Absent 18 or more days	8.5%	6.8%	11.7%	10.8%	10.7%
Retained	1.5%	1.2%	2.0%	1.9%	1.9%
Suspended	12.2%	9.3%	17.7%	16.5%	15.4%
Choice school characteristics					
Average combined scores	0.05	0.09	-0.09	0.08	0.10
Percent free or reduced lunch	40.6%	38.6%	50.9%	36.6%	35.6%
Percent black or hispanic	49.4%	46.2%	59.8%	50.0%	47.0%
Home school characteristics					
Average combined scores	-0.08	0.03	-0.28	-0.23	-0.27
Percent free or reduced lunch	47.0%	40.7%	59.3%	53.3%	56.0%
Percent black or hispanic	53.6%	47.1%	65.3%	61.6%	63.8%
School assignment					
Assigned to 1st choice	85.4%	100.0%	100.0%	40.4%	0.0%
Assigned to guaranteed school	72.5%	100.0%	0.0%	44.6%	74.5%
School attendance 02-03					
Attended 1st choice	78.7%	92.1%	81.6%	45.4%	16.2%
Attended home school	58.8%	79.4%	9.7%	35.0%	51.3%
Number of students	37115	22872	7583	3065	3595

Notes: Data from Charlotte-Mecklenberg Schools (CMS). Sample includes all students in grades 4-8 who applied to a regular or magnet school as their 1st choice for the 2002-2003 school year and were enrolled in CMS in the 2001-2002 school year. Students guaranteed placement because of siblings and in ESL are excluded.

Table VI: Baseline Characteristics by Treatment and Control Group

Variable	Admitted	Waitlisted	Difference	Adjusted Difference
Student demographics				
Black	0.614	0.585	0.030	0.011
			(0.067)	(0.022)
Free or reduced lunch	0.467	0.531	-0.064	-0.015
			(0.078)	(0.012)
Median Income ($1000s, block group for own race)	48.4	49.4	-1.0	-0.7
			(3.6)	(0.7)
Student's prior year performance				
Reading test score	-0.127	-0.069	-0.058	-0.025
			(0.110)	(0.031)
Math test score	-0.135	-0.113	0.023	0.025
			(0.106)	(0.030)
Absent 18 or more days	0.097	0.106	-0.009	-0.007
			(0.013)	(0.016)
Suspended	0.152	0.162	-0.010	-0.022
			(0.028)	(0.015)
Retained	0.019	0.018	0.001	0.001
			(0.005)	(0.006)
Home school characteristics				
Average combined score	-0.241	-0.213	-0.028	0.003
			(0.051)	(0.013)
Percent free or reduced lunch	0.543	0.524	0.019	0.001
			(0.034)	(0.007)
Percent black	0.625	0.607	0.018	-0.003
			(0.036)	(0.007)
Number of students	1175	1709	2884	2884

Notes: Sample limited to students in randomized priority groups with complete baseline data. Difference is between students admitted (won the lottery) and waitlisted (did not win the lottery). Each adjusted difference is from a separate regression of the given baseline characteristic on whether the student was randomly assigned to her first-choice school, controlling for lottery fixed effects. Standard errors adjust for clustering at the level of the first-choice school. Asterisks indicate significance (*=.05, **=.01, ***=.001).

Table VII: The Impact of Being Randomly Assigned to 1st Choice School on Characteristics of School Attending at End of 2002-2003 School Year

Characteristic of School Attending	Mean	Estimated Impact
First choice school	0.460	0.533***
		(0.054)
Not attending CMS in 2002-2003 (Attrition)	0.098	-0.018
		(0.011)
School average combined score	-0.073	0.129**
		(0.040)
Percent free or reduced lunch	0.463	-0.070***
		(0.019)
Percent black or Hispanic	0.576	-0.049
		(0.026)
Total observations		2884

Note: Each entry in the table is from a separate regression of the given characteristic of the school a student was attending at the end of the year on whether the student was randomly assigned to her first choice school, controlling for lottery fixed effects, home school fixed effects, and the baseline covariates listed in Table VI. Sample includes only students in the randomized priority group with complete baseline data. Standard errors adjust for clustering at the level of the first choice school. Asterisks indicate significance (*=.05, **=.01, ***=.001).

Table VIII. Instrumental Variables Estimates of the Impact of Attending 1st Choice School on Student Outcomes in 2002-2003

Student Outcome	Mean	Average Treatment Effect
Non-academic Measures		
Absent 18 or more days	0.135	-0.001
		(0.023)
Suspended	0.201	0.012
		(0.032)
Retained	0.022	-0.023*
		(0.009)
> 3 hrs. homework per week	0.303	0.122*
		(0.050)
Academic Performance		
Combined test score	-0.086	-0.005
		(0.050)

Note: Each entry in the table is from a separate IV regression of the given student outcome on whether the student was attending her first choice school, using random assignment to the first choice school as an instrument. These regressions control for lottery fixed effects, home school fixed effects, and the baseline covariates listed in table VI. Sample includes 2884 students in the randomized priority group with complete baseline data. Sample sizes for homework (N=2554) and combined test score (N=2581) are smaller due to missing data on the dependent variable for some students. Standard errors adjust for clustering at the level of the first choice school. Asterisks indicate significance (*=.05, **=.01, ***=.001).

Table IX: IV Estimates of the Impact of Attending 1st Choice School with Heterogeneous Treatment by Weight Placed on Academics in Choice Decision

	Combined Score	Combined Score
Attended 1st-choice school	-0.005 (0.050)	-0.105 (0.074)
Weight * attended 1st-choice school		0.077* (0.031)
P-value for interaction with *Weight*		0.016
Joint p-value on reported coefficients	0.924	0.031
Observations	2581	2581

Notes: Robust standard errors in parentheses. Asterisks indicate significance (*=.05, **=.01, ***=.001).

Table X. Comparing the Average Weight Students Place on Test Scores in School Choice to Subgroup Estimates of the Effect of Attending 1st-Choice School on Academic Achievement.

Sample	Average *Weight*	IV Estimate of Effect of Attending 1st-Choice School on Combined Test Score	Number of Students
All Students	1.33	-0.005	2581
		(0.050)	
Race:			
Non-White	1.10	-0.067	1790
		(0.058)	
White	1.83	0.172*	791
		(0.073)	
Income:			
Below Median	0.99	-0.100	1601
		(0.058)	
Above Median	1.88	0.130*	980
		(0.063)	
Free Lunch Eligibility			
Eligible	0.73	-0.061	1296
		(0.078)	
Not Eligible	1.94	0.070	1285
		(0.043)	
Baseline Test Score			
Below Average	0.81	-0.040	1386
		(0.055)	
Above Average	1.94	0.066	1195
		(0.064)	
1st-Choice School Academic Magnet			
Not Academic Magnet	1.17	-0.021	2155
		(0.055)	
Academic Magnet	2.16	0.107	426
		(0.089)	
1st-Choice School Combined Score			
Below Median	1.05	-0.036	1337
		(0.080)	
Above Median	1.62	0.047	1244
		(0.043)	

Note: Each row of the table reports estimates for a different student sub-sample, as indicated. The first column reports the average weight that the students place on test scores (*Weight*) in the school choice decision. The second column reports IV estimates of the impact of attending the first choice school on the combined student test score, using random assignment to the first choice school as an instrument. Regressions control for lottery fixed effects, home school fixed effects, and the baseline covariates listed in table VI. Sample includes only students in the randomized priority group with complete baseline data. Standard errors adjust for clustering at the level of the first choice school. Asterisks indicate significance (*=.05, **=.01, ***=.001).

Table XI: IV Estimates of the Impact of Attending 1st Choice School with Heterogeneous Treatment by Weight Placed on Academics in Choice Decision, Estimated Separately by Student Preference for Racial Mix at School

Dependent Variable: Combined Score	**Students Who Prefer School Less Than 55% Black**		**Students Who Prefer School at Least 55% Black**	
Attended 1st-choice school	0.115 (0.058)	0.186 (0.158)	-0.054 (0.059)	-0.164* (0.078)
Weight * attended 1st-choice school		-0.041 (0.065)		0.098* (0.041)
P-value for interaction with *Weight*		0.533		0.019
Joint p-value on reported coefficients	0.052	0.097	0.250	0.053
Observations	870	870	1711	1711

Notes: Each column in the table is from a separate IV regression. The dependent variable is a student's combined standardized test score in the spring of 2003. Each specification reports the coefficients on attending the first choice school and it interaction with the weight that the student places on test scores (*Weight*) in the school choice decision, using random assignment to the first-choice school and its interaction with *Weight* as instruments. All specifications control for lottery fixed effects, home school fixed effects, the baseline covariates listed in Table VI, and a direct control for the student's *Weight* estimate. Sample includes only students in the randomized priority group with complete baseline data. Student preference for racial composition in the school is each student’s posterior estimate of the value that maximizes their quadratic utility in %black at the school. Standard errors adjust for clustering at the level of the first choice school. Asterisks indicate significance (*=.05, **=.01, ***=.001).