Heterogeneous Preferences and the Efficacy of Public School Choice

by

Justine S. Hastings
Department of Economics
Yale University and NBER

Thomas J. Kane
Harvard Graduate School of Education
and NBER

Douglas O. Staiger
Department of Economics
Dartmouth College and NBER

ABSTRACT

Public school choice plans are intended to increase equity and quality in education by offering students at under-performing schools immediate academic gains from attending a higher-achieving school and by creating demand-side pressure on under-performing schools to improve. We develop a model to show how these gains from choice depend on heterogeneity in how parents choose schools, and test the model's predictions using data on parental choices and lottery assignments to schools from a school choice plan in Charlotte, North Carolina. We estimate an exploded-mixed-logit model of demand for schools and find that higher-SES parents are more likely to choose higher-performing schools, while minority families must trade off preferences for high-performing schools against preferences for a predominantly minority school. These differences in choice behavior lead to low demand-side pressure for improvement at schools serving low-SES and minority families relative to those serving high-SES families. In addition, we find that children of parents placing high weight on academic achievement experience test score gains from moving to their preferred schools, while others experience no gain or declines in test scores. This is particularly true for minority families who must sacrifice attending a predominantly minority school to select a high-test-score school. Our results imply that public school choice may widen rather than narrow the gap in achievement.

This paper was previously circulated as Hastings, Kane, and Staiger (2007a) and (2007b). This paper combines and replaces those two papers. 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 thank the editor, three anonymous referees, Joseph Altonji, Patrick Bayer, Steven Berry, David Card, Philip Haile, Carolyn Hoxby, Fabian Lange, Derek Neal, Sharon Oster, Miguel Villas-Boas, Sofia Villas-Boas, and numerous seminar participants for helpful comments and suggestions at various stages of the project. C. Sean Hundtofte and Jeffrey Weinstein provided 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.

1. Introduction

Several urban public school districts are currently experimenting with public school choice plans, and the federal No Child Left Behind Act of 2001 includes a choice provision allowing parents of children in failing schools to send their children to non-failing schools outside of their neighborhood. The goal of these school-choice plans is two-fold. First, school choice allows disadvantaged students the immediate opportunity to benefit academically from attending a higher-performing school. Second, school choice is intended to increase pressure on failing schools to improve through the threat of losing students, thereby improving educational equity.

Both of these potential benefits from public school choice implicitly assume that all parents value academics highly and will choose schools accordingly once residential restrictions are lifted. However, parents may have a variety of reasons for choosing schools, and differences in preferences and school choices may generate both heterogeneous short-run effects of attending a first choice school on academic outcomes (Heckman (1997), Heckman, Smith, and Clements (1997), Heckman, Urzua, and Vytlacil (2006)) and disparate demand-side pressure for academic improvement across low-achieving and high-achieving schools (e.g., vertical differentiation in a differentiated products market, Anderson, de Palma, and Thisse (1992)). For example, if disadvantaged families place less weight on academics when choosing schools, then schools serving these families will face little pressure to improve, and children of these families may not reap academic gains from changing schools. If this is the case, then school choice may have the smallest immediate impact on the families it is intended to help, and cause greater educational stratification over time. While the prior literature has postulated the range of potential impacts school choice could have on the achievement gap (Hanushek (1981), Hoxby (2002, 2003)), there is very little empirical evidence identifying heterogeneity in parental choices and the link between choice behavior and academic outcomes.

In this paper, we use a unique dataset and policy experiment from the Charlotte-Mecklenburg School Public School District (CMS) in North Carolina to examine heterogeneity in parental school choice behavior and its implications for public school choice. We use data on parents' multiple-ranked choices for schools during the implementation of district-wide school choice in 2002 to estimate an exploded-mixed-logit demand model for schools, allowing heterogeneity to influence choice behavior through both observable and unobservable family characteristics. We find that the weights parents place on key school characteristics are very heterogeneous, with high-income parents of high-achieving students placing the largest weights on test scores when selecting schools. In addition, we show that parents of each race prefer schools where their own race is the clear majority, implying that minority parents face much larger tradeoffs between academics and social preferences when choosing schools.

We then use our demand estimates to examine the ability of school choice to increase pressure on low-performing schools to improve through the threat of losing students. To do this, we simulate the demand response that would result if a school were to boost average test scores, holding all else constant. We find that demand-side pressure is largest for high-performing schools that serve very elastic families and minimal for low-performing schools that serve inelastic, disadvantaged families. These results suggest that public school choice, without additional incentive mechanisms, may lead to greater educational stratification of schools serving high- and low-income families, rather than providing a competitive tide that lifts all boats.

Finally, we examine how heterogeneity in choice behavior affects the immediate gains in test scores from attending a first choice school for children from advantaged versus disadvantaged backgrounds. CMS used lotteries to assign students to oversubscribed schools, allowing us to identify the causal effect of attending a first choice school on gains in own academic achievement. Theory suggests that the academic gains from attending a first choice school will be positive for parents who place a high weight on academics but potentially negative for parents who place a low weight on academics and high weights on other school characteristics that are negatively correlated with academics. Using the exploded-mixed-logit model to estimate the weight that each parent placed on test scores when choosing a school, we find that students of parents with weights at the 95th percentile experienced significant rises in End of Grade test scores of approximately 0.1 student-level standard deviations, while students of parents placing little value on academics actually experienced declines in achievement. These declines were strongest for African American families that placed little weight on academics, since they generally prefer schools with more minority students (which in CMS tend to be schools with lower academic performance). These results imply that the immediate impact of attending a first choice school broadens (instead of narrows) the gap in performance across high- and low-SES families exercising choice, and highlight how the school choice tradeoffs minority parents face between academic achievement and social match may reinforce the achievement gap.

This paper makes a number of contributions. First, it examines the potential efficacy of public school choice by estimating the underlying determinants of parents' choices. We are able to identify heterogeneity in choice parameters in our model because we have rich micro data with multiple-ranked choices (Berry, Levinsohn, and Pakes (2004)), variation in choice set characteristics from large-scale redistricting with the introduction of the school choice plan, and a diverse underlying student-body population. While prior papers have examined aspects of parental choice using a variety of measures from survey responses to residential location decisions (Armor and Peiser (1998), Vanourek et al. (1998), Greene et al. (1997), Kleitz et al. (2000), Schneider et al. (1998), Glazerman (1997), Nechyba (1999, 2000, 2003), Epple and Romano (1998, 2002), Bayer, Ferreira and McMillan (2004)), they have not estimated heterogeneity in parental choices nor linked this with demand-side pressure to improve academics and gains from attending a first choice school.

Second, our approach allows us to examine if school choice will provide stronger demand-side pressure for low-performing schools to improve or if it will result in greater stratification across schools serving low- and high-SES families. The prior literature has focused on estimating competitive pressure and its impact on academic achievement, using aggregated data and measures such as HHIs (Herfindahl-Hirschman Indices) of local districts as proxies for competition (Borland and Howsen (1992), Hoxby (2000), Hanushek and Rivkin (2003), Belfield and Levin (2002)). This implicitly assumes that all schools (or districts) face homogeneous demand elasticity and competitive pressure with respect to average test scores (Farrell and Shapiro (1990)). Expanding on this literature, Rothstein (2006) investigated the possibility that parents may choose schools based on factors other than academic achievement, but maintained a similar HHI framework and aggregated analysis for measuring competition. Modeling consumer (parental) choice for differentiated products (schools) allows different types of schools to face different demand elasticities, and thus we can test how public school choice may affect demand-side incentives to improve quality across low- and high-achieving schools.

Finally, we provide an economic explanation for why certain subgroups of students benefit from exercising choice. Several recent papers have estimated the average treatment effect of winning a lottery to attend a first choice school, abstracting away from the underlying factors that led parents to select these schools. Cullen, Jacob, and Levitt (2006) examine high school lotteries in Chicago and find no significant average impact on test scores from attending a chosen school, leading them to conclude that measurable school inputs have little causal impact on student outcomes. Others studies find positive outcomes for some subgroups or for students applying to different types of schools, but the results vary across studies (Ballou (2007), Betts et al. (2006), Hastings, Kane and Staiger (2006)). By connecting expected treatment effects with choice behavior and the tradeoffs parents face when choosing schools, we provide a framework for understanding why subgroup impacts may vary across studies and across student groups as preferences, tradeoffs, and choice sets vary.

This paper proceeds in four sections. The next section provides background on the CMS school choice plan. Section 3 presents our empirical model and outlines testable implications for the efficacy of public school choice. Section 4 presents our empirical results, including a description of the data, estimates from the exploded-mixed-logit model, demand simulations, and estimates of the gains from attending a first choice school. The final section concludes with some thoughts on the implications these results have for the design of school choice.

2. The CMS School Choice Plan

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 it to dismantle the race-based student 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 of their children. Each student was assigned a “home-school” in her neighborhood, often the closest school to her, 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. Parents could choose any school in the district. However, transportation was only provided to schools in a student’s quadrant of the district (the district was split into four quadrants called

“Choice Zones”). The district allowed significant increases in enrollment in many schools in the first year of the school choice program in an expressed effort to give each parent one of her top three choices. In the spring of 2002, the district received choice applications from approximately 105,000 of 110,000 students. Admission to oversubscribed schools was determined by a lottery system as described below.

Once the district was declared “unitary” and the court order requiring race-based busing was terminated, CMS could no longer draw boundaries based on the racial composition of a neighborhood. As a result, the former school assignment zones, which often paired non-contiguous black and white neighborhoods, were dramatically redrawn. 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 busing system. At the middle school and high school levels, this number was 52 and 35 percent, respectively. Therefore, the 2002-2003 home-school for many students was often not the school they would have been assigned at the time they chose their residence. This dramatic change in school assignment zones, the simultaneous introduction of a sweeping school choice plan, and the assignment of students to high demand schools by lottery provides a unique opportunity to examine the implications of parental choice behavior for achievement of disadvantaged students in a public school choice plan.

2.1 Heterogeneous Choices

Table I provides an overview of student characteristics and parents’ choices in our sample, broken down by race and lunch subsidy status.Throughout the analysis, we focus on students entering grades four through eight because of the lack of test scores for other grades (details on the data and estimation sample are provided in Section 4). Whites and Blacks each comprised approximately 45 percent of the student-body population. Approximately 10 percent of white students received federal lunch subsidies, while just over 60 percent of African Americans did. Minority and lunch-subsidy recipients had, on average, lower achievement (test scores were from Spring 2002 and standardized to have a mean of zero and a standard deviation of one across all students within each grade) and came from lower income neighborhoods, but there was substantial variation in test scores and income within group as well. Neighborhood income is measured as the median income for families living in a student’s census block group of the student’s same race according to the 2000 Census.

In addition, minority and lunch-subsidy students had lower-performing home-schools as measured by average student test scores.[1] However, because CMS had previously bused students of all races to schools in different neighborhoods, as well as to "midpoint" schools between neighborhoods, most students faced a diverse set of school choices within reasonable distance from their homes. On average, elementary school students had sixteen school choice options within a 12 mile driving radius from their home, and middle school students had six. On average the home-school was about three miles away from the student's house (a bit further for white non-lunch recipient students who are more likely to reside in outer suburbs). However, the next closest non-home-school option was on average only a slightly further drive. Table I shows that students had a considerable range of school test scores at school choice options within six miles (approximately twice the distance to the home-school). Across the district, school test scores range from approximately negative one to one, so students of all socioeconomic backgrounds had schools ranging from the lower to the upper quartile within reasonable proximity. Although test scores at the top scoring school within six miles are on average higher for non-poor white families, the standard deviation is large, and the variation in school average test scores for proximate schools is substantial for all socioeconomic groups. This variation along with multiple choices will help identify the underlying determinants of parents' school choices.[2]

Since parents were guaranteed a slot in their default school, many parents listed only one or two schools on their choice forms. Table I indicates that parents of students who were white and ineligible for lunch subsidies were approximately twice as likely as non-white-lunch-recipient student to choose their home-school as their first choice, which is consistent with the difference in average home-school test score across the two groups. Overall, parents of white, lunch-ineligible students were also much more likely to list only one choice than parents in the other subgroups. Parents of free-lunch eligible non-white children were most likely to list three choices. Many parents listed subsequent choices after listing their home-school option, despite the fact that admission to the home-school was guaranteed. Of parents who chose the home-school first or second, 27.4, 43.0, 39.2 and 50.9 percent, across the four race and income categories respectively, listed subsequent choices beyond the guaranteed option. Listing additional choices took little time (particularly if parents had already considered alternative

[1] School average test scores were calculated as the average of the combined standardized 2001-2002 test score for students attending the school in the 2002-2003 school year. See Section 4 for details.
[2] For maps of school options and student characteristics in CMS see Hastings, Kane and Staiger (2007a).

schools in making their choice), and doing so when the form allowed it was a reasonable precaution when parents lacked complete understanding of or confidence in the process.[3] The availability of multiple choices from those parents who listed their home-school first or second adds further choice set variation and aids in the identification of demand parameters, although estimates were similar between such parents and the sample as a whole.

Despite the proximity of high-performing schools to neighborhoods of various socioeconomic backgrounds, there was little unanimity in parents' choices of schools. Even within the same elementary home-school assignment for 2002-2003, parents listed, on average, 10.4 different first choice elementary schools.[4] It is clear that parents have very heterogeneous preferences over school characteristics. Figure 1 shows that approximately 20 percent of parents chose schools that had lower test scores than the school they had guaranteed admission to, suggesting that school characteristics that were potentially negatively correlated with average test scores were the strongest determinants of choice for some families. This suggests that heterogeneous preferences may play a key role in school selection and academic outcomes in public school choice.[5]

2.2 Lottery Assignments

The district implemented a lottery system for determining admissions to oversubscribed schools. Approximately one third of the schools in the district were oversubscribed. 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 three priority groups depending upon their grade level, with students in terminal grades—grades five, eight, and 12—given highest priority).

[3] Non-white and free-lunch parents may also have been less likely to use the automated phone system to choose, which ceased prompting for additional choices after the home -school was chosen.
[4] This number accounts for differences in choices listed driven by differences in the prior-year's school.
[5] Hastings and Weinstein (2008) find less heterogeneity, with most parents choosing schools with higher test scores, when they receive simplified information on school test scores. Our model of parental choice in Section 3 allows heterogeneity in preferences to be interpreted as heterogeneity in information costs in collecting and/or interpreting measures school academic quality.

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

Priority 3: Parent applying to a school within her choice zone

Parents listing a given school as their first choice were sorted by priority group and a randomly assigned lottery number.[6] Any slots remaining after home-school students were accommodated were assigned in order of priority group and random number.[7] 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.

Children of parents not assigned to one of their top choices were placed on a waiting list. About 19 percent of students winning the lottery to attend their parents' first choice schools subsequently attended a different school, with 13 percent attending their home-schools instead and another six percent attending a different school entirely, with most of these students changing address. When slots became available, students were taken off the waitlist based on their lottery numbers alone, without regard for their priority groups.

2.3 Potential for Strategic Choice

The lottery mechanism used by CMS was a "first-choice-maximizer", in which slots were first assigned to all those listing a given school as a first choice before moving to those listing the school as a second or third choice. In such a mechanism, parents with poor home-school options may have an incentive to misstate their preferences – not listing their most preferred school if it had a low probability of admission (Glazerman and Meyer (1995), Abdulkadiroğlu and Sönmez (2003), Abdulkadiroğlu et al. (2006)). Instead, they may have hedged their bets by listing a less

[6] The random number was assigned by a computer using an algorithm that we verified with CMS computer programmers. We also verified the randomization using the lottery and student assignment data.

[7] The choice program allowed for sibling admissions in the following way. If a parent had two children in the same grade level (i.e. both in middle school), and listed the same school as the first choice for both children, then once one child was admitted the second child was also admitted. No preference was given to siblings in different grade levels, e.g. a child choosing an elementary school that feeds into a sibling's first choice middle school. When we turn to estimate the effects of attending a first choice school on academic outcomes, we will only use students for whom lottery number alone determined admission, which will consequently exclude the few students who were admitted to their first choice school based on a sibling's lottery number.

preferred option with a higher probability of admission in order to avoid being assigned to their home-school. Such strategic behavior would imply that parental choices would not reflect true preference orderings for schools – to the extent that parents are *not* listing their preferred match due to strategic hedging.

However, there were a number of reasons why such strategic behavior was probably rare in the first year of the choice plan that we are studying. First, this was a new school choice plan, and parents did not know the details of how the lottery system would be operated. The handful of district officials who knew the lottery details were not allowed to communicate them to parents, and parents were never given their actual lottery numbers.[8] The district also told parents that they would make every attempt to give each student admission to one of their chosen schools and *instructed them to list what they wanted*.[9] Even in long-standing limited choice plans where districts instruct parents to choose strategically, there is evidence that many parents do not act strategically (Pathak and Sönmez (2008)). 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 underperforming home-schools would be given top priority. This would counteract the incentive for these students to hedge their choices, as outlined above.

If there were widespread strategic behavior by parents, we would expect those with low-quality default schools to hedge their bets and list less desirable schools for which they might have a higher probability of admission. Hastings, Kane and Staiger (2007a) use the redistricting to test if parents with *exogenous* changes to the quality of their home-school had lower preferences for high-quality schools, as would be predicted if parents were behaving strategically, and they find no evidence that this was the case. We discuss this evidence in more

[8] Parents were given a choice booklet to aid with their school choice decision, which stated that "Students will be chosen through a lottery process in the following order of priority:" and then listed the priorities outlined above. This was the only description of the lottery given.

[9] Specifically, the school choice booklet instructed parents to read the school descriptions, decide what their first, second and third choices were, and complete the choice application worksheet. It also stated that every student must apply or they would not be guaranteed a seat at their home-school. They were given a help line number to call if they needed assistance. The staff at CMS emphasized to parents that they should list the schools they want, and the district would do its best to give each parent one of their top three choices.

detail in Section 4 and in an appendix. Overall, we did not find evidence indicating that strategic behavior played a significant role in this first year of school choice.[10]

3. Model of Parental School Choice and Academic Achievement

This section develops a general framework linking parental choice behavior with demand-side pressure to improve academics and gains from attending a first choice school. We begin by laying out a general model of the demand for schools and student achievement, and then choice of schools, and then turn to estimation of the model.

3.1 Demand for Schools and Student Achievement

Our model of the demand for schools is based on a standard random utility framework. Let U_{ij} be the expected utility of parent *i* from attending school *j*. We assume that parents choose the school that maximizes utility, where utility is a linear combination of the student's predicted academic achievement at that school, A_{ij}, and non-academic factors, such as distance from home and school racial composition, V_{ij}.

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

β_i^A represents the weight student *i*'s parents place on academic achievement, which may vary idiosyncratically and with observable student characteristics (such as income, race, and baseline academic achievement). The weight parents place on academic achievement may vary for two reasons. First, some parents may simply place an inherently high value on academic achievement. Second, even if all parents place high importance on academic achievement, some parents may lack information or face high decision making costs, leading them to place lower expressed weights on academic achievement when determining their expected utility and selecting a school (Borghans, Duckworth, Heckman and Weel (2009), Della Vigna (2009), Hastings and Weinstein (2008)). These two sources of heterogeneity cannot be separately

[10] 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.

identified in our analysis because they result in observationally equivalent choice behavior and student outcomes.[11]

We do not observe academic achievement at each school directly. Instead, we assume that the academic achievement of student *i* at school *j* is proportional to a latent measure of academic quality at each school (A_{ij}^*) that is the sum of the school average test score, S_j, and an idiosyncratic match between the school and the student, γ_{ij}.

(2) $$A_{ij} = \kappa_i A_{ij}^* = \kappa_i \left(S_j + \gamma_{ij} \right)$$

The parameter κ_i captures sensitivity of student *i's* academic performance to school academic quality: students with higher κ_i obtain more academic benefit from attending a school with high academic quality. κ_i may also vary idiosyncratically and with observable student characteristics such as income, race, and baseline academic achievement.

Non-academic factors that affect utility of student *i*'s parents from sending their child to school *j* are assumed to be a linear function of other observable school characteristics such as proximity and racial composition, X_{ij}, the importance parent *i* places on those characteristics, β_i^X, and an idiosyncratic component, ω_{ij}.

(3) $$V_{ij} = X_{ij} \beta_i^X + \omega_{ij}$$

Given these assumptions, we can rewrite equation (1) to state utility in terms of observable school characteristics as:

(4) $$U_{ij} = S_j \beta_i^S + X_{ij} \beta_i^X + \varepsilon_{ij}$$

Where, $\beta_i^S = \beta_i^A \kappa_i$ and $\varepsilon_{ij} = \beta_i^S \gamma_{ij} + \omega_{ij}$. Equation (4) determines parental preference rankings over schools based on a combination of observable school characteristics and the idiosyncratic weights that parents place on these characteristics. Summing the top ranked school across all

[11] This is a general problem with revealed choice and inference on underlying preferences. Borghans, Duckworth, Heckman and Weel (2009) provide an insightful discussion of how heterogeneous preferences or heterogeneous perceived choice sets can generate the same expressed choices, and DellaVigna (2008) draws links to the marketing literature that has focused on how measured 'preferences' might change with salience, focus, suggestion and framing. To separately identify the information effect, Hastings and Weinstein (2008) provided simplified information on school characteristics to randomly selected families. Without such an intervention, these two underlying reasons for estimated heterogeneity in preferences are observationally equivalent. We have kept our theoretical and empirical framework general enough to allow for either underlying interpretation.

parents yields the market demand curve for each school as a function of the school's academic and non-academic characteristics.

In this general framework, a parent will place more weight on school test scores (β_i^S is large) either because their child's academic achievement is more sensitive to the school they attend (κ_i is large), or because their parent placed more weight on academic achievement when selecting schools (β_i^A is large, reflecting inherent preferences or better information, as discussed above). In either case, one would expect parents who placed a high weight on school test scores to be more likely to choose schools that improved their child's academic achievement. Thus, among those parents who exercise choice, the expected gain in academic achievement should be increasing in the weight that the parents placed on school test scores when selecting a school.

To see this more formally, let the student's potential test score at every school (y_{ij}) depend on the school contribution (A_{ij}), student characteristics (z_i) and idiosyncratic performance that is not known at the time of choice (u_{ij}):

(5) $$y_{ij} = A_{ij} + z_i\delta + u_{ij}$$

Thus, A_{ij} captures how expected academic achievement varies across schools for each student. Each student is assigned a default school (the home-school) but may choose an alternative school (the first choice school) if they prefer another school over the default school. Let y_{i0} and y_{i1} represent the test scores of student i at the default (j=0) and alternative (j=1) school, and define Δ to be the difference operator so that, for example, $\Delta y_i = y_{i1} - y_{i0}$. Since utility must be higher at the preferred school ($\Delta U_i > 0$), the expected gain in academic achievement among those parents who prefer an alternative school is given by:

(6) $$E(\Delta y_i \mid \Delta U_i) = E(\Delta A_i \mid \beta_i^A \Delta A_i + \Delta V_i > 0)$$

Equation (6) defines a "treatment-on-treated" parameter that captures one of the key potential benefits of school choice: the academic impact of switching schools among those who would chose to switch if we allow school choice. Note that as β_i^A grows very large, the expected achievement gain alone determines choice and, therefore, must be positive for all students who choose an alternative school. However, for a student with low β_i^A (near zero), non-academic factors determine choice and the expected achievement gain is ambiguous – it could even be negative if ΔA is negatively correlated with ΔV, i.e. if non-academic school characteristics that are important to parents lead them to choose schools with weak academic performance. For example, the parent of a minority student who wants her child to attend a school with same-race peers may give up gains in her child's academic achievement if the weight she places on academics is low. Hence, this basic framework generates the prediction that the expected treatment effect is positive for all students placing a large weight on academic achievement. Among students placing less weight on academic achievement, the expected treatment effect will depend on the tradeoffs that parents face; it could even be negative if expected academic achievement is sufficiently negatively correlated with other valued school characteristics.

Using Equation (2), we can restate equation (6) in terms of the sensitivity of the child's academic achievement to latent school quality (κ_i):

(6') $$E\left(\Delta y_i \mid \Delta U_i\right) = E\left(\kappa_i \Delta A_i^* \mid \beta_i^A \kappa_i \Delta A_i^* + \Delta V_i > 0\right)$$

Note that κ_i has a similar impact on the expected achievement gain as β_i^A: as κ_i gets very large, the expected achievement gain alone determines choice and, therefore, must be positive for all students who choose an alternative school. However, the expected achievement gain goes to zero when $\kappa_i = 0$, since this implies that the student's achievement is not sensitive to school quality.

Taken together, our model implies that there will be an unambiguously positive impact of attending a first choice school on academic achievement only for parents who placed a high weight on student test scores ($\beta_i^S = \beta_i^A \kappa_i$ is large). This impact will tend toward zero for students of parents who place little weight on their child's test scores because their child's achievement is not sensitive to the school they attend. But if placing a low weight on school test scores reflects weak parental preferences for achievement, then the impact of attending a first choice school could be either positive or negative, depending on the tradeoffs parents face between expected academic achievement and other valued school characteristics.

3.2 Estimation of the Model

Estimation of the model proceeds in two steps. In the first step we estimate an exploded-mixed-logit model of demand for schools, and use these estimates to simulate the extent to which schools face demand-side pressure to improve test scores. In the second step we use lottery assignments to oversubscribed schools to estimate the impact of attending a first choice school on standardized test scores, allowing the impact to vary explicitly with the estimated weight parents placed on school test scores when selecting their schools.

3.2.1 Estimating an Exploded-Mixed-Logit Model of Demand for Schools

We use the parental choices, along with data on each student and school, to estimate how parents weigh different school characteristics and how this varies in the population. We estimate an exploded-mixed-logit model for parental choice data (McFadden and Train (2000), Train (2003)). "Exploded" refers to a multinomial logit model that incorporates multiple-ranked choices for each person (not just the first choice), and "mixed" refers to logit models with random coefficients. By introducing individual heterogeneity in the logit coefficients, the mixed logit model allows for flexible substitution patterns – generating credible estimates of demand elasticities. The mixed logit can approximate any random utility model, given appropriate mixing distributions and explanatory variables (Dagsvik (1994), McFadden and Train (2000)).

Specifically, let U_{ij} be the expected utility of parent *i* from having their child attend school *j* as defined in equation (4). Parent *i* chooses the school *j* that maximizes his or her utility over all possible schools in the choice set. For the first choice, the parent chooses over the set of all available schools (denoted J_i^1), so that:

$$d_{ij}^1 = 1 \text{ if and only if } U_{ij} > U_{ik} \ \forall k \in J_i^1$$

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

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

We assume that that ε_{ij} in equation (4) is distributed *i.i.d.* extreme value and that the idiosyncratic portions of preferences are drawn from a multivariate normal mixing distribution ($\beta \sim f(\beta \mid \mu_\beta, \theta)$, where μ and θ denote the mean and variance parameters), yielding a traditional exploded-mixed-logit model.

Given these assumptions, the probability that parent *i* chooses schools (j^1,j^2, j^3) is given by:

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

$$(7) \qquad = \int \prod_{c=1}^{3} \frac{\exp(S_j \beta_i^S + X_{ij} \beta_i^X)}{\sum_{k \in J_i^c} \exp(S_k \beta_i^S + X_{ik} \beta_i^X)} f(\beta|\mu, \theta) d\beta$$

These probabilities form the log-likelihood function: [12]

$$(8) \qquad LL(X, \mu, \theta) = \sum_{i=1}^{N} \sum_{j=1}^{J_1} \sum_{k=1}^{J_2} \sum_{l=1}^{J_3} d_{ij}^1 d_{ik}^2 d_{il}^3 \ln(P_i(j,k,l))$$

Our baseline specification used 2002 average test scores of the students in the school in 2003 as the proxy for academic strength of the school. To capture non-academic factors commonly thought to influence school choice, we included driving distance from the student to the school, the fraction African American at the school and its square (allowing preferences for fraction African American to peak at some value), and whether the school was in the student's Choice Zone, was the student's prior-year school, or was the student's designated home-school.

We estimated the model separately by race and lunch subsidy status, thus allowing for heterogeneity in choice behavior through full interactions with these two key socioeconomic variables. We allowed mean preferences to vary with whether the student was an elementary or middle school student, and allowed mean preferences for school test scores to vary with the student's prior year test score performance and measures of neighborhood income. We assumed that the random parameters follow a joint normal distribution, where the preference for distance was drawn from a negative lognormal distribution (so that all people dislike commuting) and idiosyncratic preferences for school test scores and measures of proximity (distance and home-school) can covary freely.[13] Since equation (8) does not have a closed form solution, simulation methods were used to generate draws of β from $f(\cdot)$ to numerically integrate over the distribution of β. Estimation was by the method of maximum simulated likelihood, using 300 draws of β from $f(\cdot)$ for each individual in the data set. The results were not sensitive to increasing the number of draws used.

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

[13] We focus on covariance between idiosyncratic preferences for test scores, home-school and driving distance since these covariances could significantly impact demand-side pressure for schools to improve quality. Allowing for a 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. For further discussion of robustness checks we performed, please see Appendix A.

3.2.2 Identification of the Exploded-Mixed-Logit Model

Several aspects of the CMS school choice data help to identify the parameters in our demand model. First, the large scale redistricting that occurred with the introduction of school choice helps to identify values placed on distance separately from residential sorting. Residential sorting could overstate the importance of proximity and neighborhood schools if parents had previously located near to their preferred schools. However, the former school assignment zones often required parents to live far from their preferred school, and with large scale redistricting many parents unexpectedly found themselves assigned to new neighborhood schools. Thus, in the first year of the choice plan, home-school assignment and distance to a school were not strongly linked to preferences through prior residential sorting. Multiple choices listed by those selecting their home-schools first further separates preferences for school characteristics from residential sorting by simulating the unavailability of the neighborhood school.

Second, historic placement of schools for busing in CMS provides wide variation in school characteristics for families in all socioeconomic groups, dampening collinearity problems that may be present in other settings. For example, as was seen in Table I, students from all socioeconomic groups had high scoring schools within reasonable proximity. Third, approximately 95 percent of parents submitted choices for the choice plan. Thus we have data for nearly the entire student population—whereas most work using school choice data has been dependent on limited and potentially non-representative subgroups of students.

Fourth, the multiple-ranked responses provided by each parent creates 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 values parents place on school characteristics from observed substitution patterns for each individual – a stronger source of variation for identification than cross-sectional changes in the choice set based on geographic location (Train (2003), Berry, Levinsohn, and Pakes (2004)). Intuitively, when only a single (first) 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 an unusually high weight placed by the individual on some aspect of the choice (β_i). 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.

3.2.3 Simulating Demand-side Pressure to Improve Test Scores

We use these demand estimates to simulate the extent to which school choice will place demand-side pressure on underachieving schools to improve through the threat of losing students. To do this we simulate the expected change in demand, Q_j, for each school were it to raise the average test scores of its students, holding all else constant. This demand response is the partial derivative of the probability that student *i*'s parents choose a school *j* as their first choice school with respect to the average test score performance at school *j*, summed over all students:

(9) $$\frac{\partial Q_j}{\partial S_j} = \sum_{i=1}^{N} \frac{\partial \hat{P}_{ij}}{\partial S_j}$$

where the probability that parent *i* selects school *j*, $\hat{P}_{ij}$,is given by the exploded-mixed-logit demand specification integrated over the random draws from the estimated distribution of underlying utility weights.

3.2.4 Estimating the Impact of Attending a First Choice School.

To estimate the impact of attending a first choice school on standardized test scores, we use lottery assignments of students to oversubscribed schools and allow the impact to vary explicitly with the estimated weight parents placed on academic achievement when selecting their schools. We exploit lottery random assignment by analyzing the subset of students choosing non-home-schools that were oversubscribed and limit our sample to students in marginal priority groups within those schools – priority groups for which lottery number alone determined initial admission. We ignore members of priority groups in which all students were either admitted or denied admission, since lottery numbers had no impact on their admission status.

Our estimation strategy is motivated by equations (5) and (6). Let D_i = 1 if student *i* attends her alternative (first choice) school, and D_i = 0 otherwise (which, for most of these students, will mean that they attend their home-school). Then the observed test score for each student is given by $y_i = D_i y_{i1} + (1 - D_i) y_{i0}$. Using the definition of y_{ij} from equation (5) yields:

(10) $$y_i = A_{i0} + z_i\delta + D_i(\Delta A_i) + (u_{i0} + D_i \Delta u_i)$$

Or in conventional regression notation:

(10') $$y_i = \delta_0 + z_i\delta + D_i\gamma + \upsilon_i$$

Where $\delta_0 = A_{i0}$ and $\gamma = \Delta A_i$ represent random coefficients, and $\upsilon_i = u_{i0} + D_i \Delta u_i$.

Our goal is to estimate the mean of γ among students who prefer an alternative school, which corresponds to $E(\Delta y_i \mid \Delta U_i)$ – the treatment-on-treated parameter defined by equation (6). Estimating equation (10') by OLS among the sample of students who listed a non-home-school as their first choice (not all of whom eventually attended their first choice school) will only yield unbiased estimates if D is independent of both y_{i0} and Δy_i in this sample. This is unlikely to be true if parents choose schools in part based on expected academic performance.

Therefore, we use lottery assignment to schools as an instrument for D. In the sample of students who choose a non-home-school as their first choice, lottery number is a valid instrument because it is the primary determinant of whether the student attends that school, and is independent of y_{i0} and Δy_i . Because the lottery assignment does not perfectly assign students to schools, the IV estimate will be a Local Average Treatment Effect (Angrist and Imbens (1995), Angrist, Imbens and Rubin(1996)) corresponding to the impact of attending a first choice school among those students whose school attended was influenced by the lottery outcome (the "compliers").

We estimate the following empirical version of equation (10'), letting the treatment effect depend on the estimated weight parents placed on test scores ($\hat{\beta}_i^S$) when selecting their schools:

$$(11) \qquad y_i = \delta_0 + z_i\delta + D_i\gamma_1 + \left(D_i\hat{\beta}_i^S\right)\gamma_2 + \upsilon_i$$

where winning the lottery and winning the lottery interacted with $\hat{\beta}_i^S$ were instruments for whether the student attended her first choice school and its interaction with $\hat{\beta}_i^S$. We control for baseline characteristics (including $\hat{\beta}_i^S$), home-school fixed effects (to capture variation in δ_0) and lottery fixed effects (since winning the lottery is random within, but not between, lotteries). $\hat{\beta}_i^S$ is a posterior estimate of the weight each parent placed on school tests scores, calculated from our demand model using Bayes' rule, as follows (Revelt and Train (1998), Train (2003)):

$$(12) \qquad E\left(\beta_i^S \mid d_i, X_{ij}, \mu, \theta\right) = \frac{\int \beta_i^S P\left(d_i \mid X_{ij}, \beta\right) f\left(\beta \mid \mu, \theta\right) d\beta}{P\left(d_i \mid X_{ij}, \mu, \theta\right)}$$

where d_i denotes the choices the parent made. This equation is the expected value of the weight parent *i* placed on school test scores given his or her characteristics, the choices he or she made, the characteristics of his or her choice set, and the distribution of demand parameters in the population. We calculate this posterior for each student in our randomized lottery admission group using 1,000 draws from the estimated demand parameter distribution.[14]

Note that $\hat{\beta}_i^S$ incorporates all information about a parent's choices and the tradeoffs she faced into one index summarizing the importance of school test scores in determining that parent's school choice. In addition, it depends only on baseline data that is independent of whether the student won the lottery, so its interaction with winning the lottery is therefore a valid instrument once one has conditioned on baseline data. Finally, note that coefficient estimates for terms involving $\hat{\beta}_i^S$ are not attenuated by the usual measurement error bias because the measurement error ($\beta_i^S - \hat{\beta}_i^S$) is uncorrelated with the posterior estimate $\hat{\beta}_i^S$ by construction (Hyslop and Imbens (2001)).

4. Data and Empirical Results

4.1 Data

We obtained secure access to administrative data for all students in CMS for the year before and after the implementation of school choice. Throughout the analysis, we focus on students entering grades four through eight because of the lack of test scores (either baseline or outcome) for other grades. Students take End of Grade tests in grades three through eight, so that students entering grade four are the first to have a baseline test available.

For each student, we have the choice forms submitted to CMS, allowing each parent to specify up to three choices for her child's school. In addition to the parental choices our data contain student characteristics for the years before and after school choice, including geocoded residential location, race, gender, lunch subsidy recipient status, and school assignment. Student test scores are for North Carolina End of Grade Exams in math and reading, and were standardized to be mean zero and standard deviation one in each grade and year. These tests were developed specifically by North Carolina to measure student progress, and were aligned to measure progress on a single underlying factor across all grades (Sanford (1996)). We use these data to construct key covariates in the demand for schools, such as driving distance from each

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

student to each school, an indicator for busing availability, an indicator for the prior-year's school, measures of student-level income, student baseline academic achievement, school-level academic achievement, and school-level racial composition. The variables used in our model are described in detail in Table II.

The final estimation sample used to estimate the exploded-mixed-logit model included 36,887 students entering grades four through eight. The means and standard deviations of these variables across the 2.4 million school, student, and choice rank combinations used to estimate the model are reported in columns 1 and 2 of Table III, and the mean and standard deviation of the average characteristic across students is reported for the 36,887 students are reported in columns 3 and 4.

4.2 Empirical Results

Table IV presents the results from the exploded-mixed-logit demand estimation in four subsamples, defined by race and lunch recipient status. We report the estimates for the mean of each logit coefficient, along with the standard deviations and correlations (where appropriate) for the random parameter distributions. The parameters determining the importance of school test scores in school choice are reported at the top of Table IV, followed by the parameters that govern the distribution of preferences for non-academic factors. We allowed the mean of each logit coefficient to vary by grade level (elementary versus middle school student), and we further allowed the mean weight that a parent placed on school test scores to depend on the student's own baseline test score and neighborhood income (i.e., interaction terms with school test scores). Baseline test scores and neighborhood income (in thousands) are constructed as deviations from the average, so that the coefficient on school scores represents the weight placed on test scores for a student with average test scores and neighborhood income (when the interactions are zero). The standard errors for each of the estimates are reported in Appendix Table B.I. All of the parameters were precisely estimated and statistically significant.

Broadly speaking, the estimates are consistent with expectations. Focusing first on the means of the coefficients, parents of both elementary and middle school students in all four subsamples were less likely to choose a school far away, and more likely to choose a school that had high test scores, was their home-school[15], had a majority of students their own race, was their school in the prior year, or was a school in their choice zone (assuring school bus transportation). The average middle school parent placed relatively less weight on non-academic factors (distance, attending last year's school, and attending a school in the choice zone) and relatively more weight on academic factors (school test scores) than did the average elementary school parent, implying that parents of middle school students were more responsive to test score differences all else equal (as one might expect).

Among both elementary and middle school parents, the estimates in Table IV suggest substantial heterogeneity in the weight placed on academic versus non-academic factors. The positive interactions with school score indicate that the average weight placed on school score increased with both neighborhood income and the student's baseline test score for parents of students in all four subgroups. Holding neighborhood income and baseline test scores constant, non-whites placed more weight on school scores than whites, while parents of students not receiving lunch subsidies placed more weight on school scores than those receiving lunch subsidies. Relative to parents of white students, parents of non-white students preferred schools in which a larger fraction of students were black, but otherwise tended to place less weight on the non-academic factors.

In addition to the variation related to observable characteristics, idiosyncratic preferences contributed substantial variation to the weight placed on both academic and non-academic factors, with the estimated standard deviation for most of the random coefficients ranging between one-quarter and one-half of the mean for each coefficient. The variation arising from idiosyncratic preferences for school scores was similar in magnitude to the variation arising from neighborhood income and baseline scores. A one standard deviation increase in the random coefficient raised the weight placed on school scores by 0.2 to 0.7 across the four subsamples, while a one standard deviation increase in neighborhood income (about $25 thousand) raised the weight placed on school scores by 0.1 to 0.4, and a one standard deviation increase in student baseline test score raised the weight placed on school scores by 0.2 to 0.6. For all four subsamples, the random coefficient on school scores was negatively correlated with the random coefficient on home-school and positively correlated with the random coefficient on distance.

[15] The home-school preference could be consistent with a default effect, rather than a preference for a neighborhood school, since the home-school was listed as the default option. Hastings et al. (2007a) provide evidence suggesting that preference for a home-school represents a neighborhood preference rather than default behavior.

Hence parents who consider options outside of their home-schools are more likely looking for high test scores when deciding which schools to pick. For the average parent, selecting a high-achieving school will require her to choose a school that is farther than her home-school and a school that is not the home-school, leading to tradeoffs between academic gains and proximity.

One way to interpret the magnitude of the coefficients and the amount of heterogeneity reported in Table IV is in relative terms. For example, the ratio of the mean school score coefficient to the mean distance coefficient for white, non-lunch, elementary students implies that the average parent would be willing to travel an additional 3.8 miles to attend a school with one standard deviation test scores. Within this subsample, variation in neighborhood income, student baseline score, and the random coefficients for school score and distance each generate a standard deviation across parents of about one to two miles in the willingness to travel to attend a high score school. Parallel calculations suggest a similar amount of variation between the subsamples, and within each subsample, in the tradeoffs parents are willing to make. Another way to interpret the magnitude of the coefficients is to calculate what they imply for the probability of choosing an alternative school, relative to the home-school (the odds ratio). The mean school score coefficient for white, non-lunch, elementary students implies that a one standard deviation increase in school scores at the alternative school will increase the odds of choosing that school by 331 percent ($e^{1.46}$-1). The increase in the odds of choosing the alternative school would be considerably larger for a parent whose coefficient on school scores was one standard deviation above the mean ($e^{1.46+0.402}$-1 = 544 percent). Looking ahead, this 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 preferences for academics.

In addition to trading-off proximity for academics, African American parents must tradeoff academic gains against the racial composition of peers. The coefficients on percent black and its square imply that the average African American parent prefers majority black schools, while the average white parent prefers schools that are majority white. The racial preferences are stronger for elementary school students than for middle school students in three of the subsamples, although it may be that racial preferences near endpoints are difficult to estimate precisely since there are few schools (particularly middle schools) that are near 100 percent of either race. The coefficients in Table IV imply that the average parent of an African American student is 10-70 percent less likely to choose a school that is 35 percent black versus a school that is 80 percent black, all else equal, whereas the average parent of a white student is two to three times more likely to select the low-minority school. In CMS, the percent black at a school is negatively correlated with average test scores, implying that African American parents must value academic achievement 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.

4.3 Robustness Checks for Demand Model

We conducted several robustness checks of the exploded-mixed-logit model.[16] First, we estimated the model using alternative measures of a school's academic performance. Given the major policy change under the choice plan, parents may have had a difficult time predicting expected academic performance at schools in their choice set, and may have used indicators of school performance other than the measure we used in our baseline model. Our baseline specification used average 2002 test scores of the students in the school in 2003. But we also estimated models using average 2003 test scores of the students in the school in 2003, average 2002 scores for students in each school in 2002, and a "value added" measure of each school's impact on academic achievement in the prior year (estimated as the school fixed effect from a regression of End of Grade test score on prior year test score and other student baseline characteristics, such as race, gender, and subsidized lunch status). The parameter estimates across the three measures of average test scores were quite similar, but the model using our base specification had the highest likelihood value, and in that sense fit the choice data best. We found that the model using value added did not fit the observed choice data well at all. This should not be surprising, since school value added estimates were not available to parents in 2002.

Second, due to the way the lottery was run, parents may have had an incentive to misrepresent their true preferences. If they understood the allocation mechanism, a parent with an undesirable home-school might want to hedge against being assigned to the home-school. They would do so by picking less desirable schools than they actually prefer – trading off

[16] See Hastings, Kane and Staiger (2007a) for details. The key findings are reproduced in Appendix A for convenience.

desirability for increased chance of being admitted. However, as discussed earlier, it was not at all clear that parents had the information or experience in the first year of choice to understand how to exploit the incentives of the allocation mechanism. We tested for the presence of strategic behavior in the first year of choice by exploiting the redrawing of school boundaries. Many of those who lived in the same contiguous school assignment zone in 2001-2002 were given different home-school assignments in 2002-2003. Hence, among those with the same school assignments who lived in the same neighborhood in 2001-2002, some students experienced positive and others negative shocks to the quality of their guaranteed school. If strategy was a major component of parental choices, we would expect to see very different choice behavior for those with negative versus positive shocks to the quality of their home-school. We did not, however, find significant differences in the exploded-logit parameters (particularly for the coefficients relating to school test scores) across families experiencing positive and negative shocks to their guaranteed school within the subsample of students who lived in 2001-2002 assignment zones that were split by redistricting. This suggests that strategy was not a major factor driving our demand estimates in this first year of the public school choice plan.

Finally, we used the exogenous reassignment of nearly half of the students in the district to test the extent to which residential sorting affected our demand estimates. Residential sorting could lead us to overstate preferences for proximity if parents had already sorted to live next to the schools they preferred. To test for the potential effects of residential sorting on our estimates, we re-estimated our model for the subsample of students who were reassigned (whose school assignments under the busing plan in 2001-2002 were different from their home-school in 2002-2003). The estimated parameters were qualitatively and quantitatively similar in the reassigned subsample compared to those for the full sample, suggesting that endogenous residential location is not a major source of bias in this data. The similarity of the results is not surprising because our model is using the information in multiple choices to identify preferences. A substantial fraction of parents who listed their home-school also listed subsequent choices. For these parents, multiple choices simulate reassignment whether or not they were actually reassigned.

4.4 Implications for Demand-side Pressure to Improve Academic Achievement

The exploded-mixed-logit estimates can be used to simulate the degree to which demand-side pressure for academic improvement varied across low-and high-performing schools under public school choice. To examine the extent to which public school choice affects the demand for each school, we took each school individually, added 0.34 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.[17] The simulated change in demand tells us how responsive each school's demand curve is with respect to their own test score outcomes.[18]

Figure 2 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). Because of the difference in size, we plot the results separately for elementary (Figure 2a) and middle (Figure 2b) schools. 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 exploded-mixed-logit model. Parents with high preferences for school scores, and thus low preferences for their neighborhood schools, are sensitive to changes in school scores and are willing to consider a broader set of schools beyond their home-school. These parents are likely to only consider high scoring schools for their children and are willing to change schools in response to an increase in score at another high scoring school, even one that is located further away. This pattern is substantially stronger for middle schools, reflecting the fact that parents of middle school students tend to place higher weights on test scores and lower (less negative) weights on distance than elementary school parents do. The results imply that the demand-side incentives to focus on student performance are larger for higher-performing schools, since these schools compete more intensely on academic quality for the quality-elastic segment of the population, and this disparity in demand-side pressure is particularly strong for middle schools when compared to elementary schools.

Figure 3 plots differences in average baseline test scores (in 2002) between the marginal students (those who are drawn in by the .34 average student-level standard deviation score increase) and students who previously enrolled in each school. The incentive for any school to

[17] This is approximately equivalent to a 10 point increase in the average percentile score for students attending that school.
[18] As a check on the model fit, we compared the predicted demand based on observed school characteristics with the actual number of students who listed each school choice first and found that the predicted and actual demand was correlated 0.91 across schools. A regression line of actual demand on predicted demand had a coefficient of 0.954 that is not statistically different from one.

improve its performance would be dampened if, in doing so, they were swamped by lower-performing students, who would bring down mean performance and potentially be more costly to educate. The fact that most points lie above the 45 degree line implies that the marginal students, on average, were higher performing than the students already enrolled. Again, this reflects the heterogeneity that was estimated in the exploded-mixed logit model, with higher performing students placing the highest weight on school test scores in choosing a school.

The key features of the simulations reported in Figures 2 and 3 appear to be driven primarily by the estimated heterogeneity in preferences, rather than other details of the specification. We found similar results in all the alternative specifications we have estimated that allowed for heterogeneity in preferences.[19] The results imply that instead of increasing pressure on low-performing schools to improve through the threat of losing students, high-performing schools will face the most pressure to improve because they serve high-performing students whose parents are willing to switch schools in response to small changes in test scores. Thus, school choice is likely to generate changes in demand that act to widen rather than narrow the gap between high- and low-performing schools. We next turn to examine the implications of heterogeneous choice behavior for the immediate impact of allowing parents to send their children to the school of their choice.

4.5 Heterogeneous Treatment Effect from Attending a First Choice School

As outlined in Section 3, heterogeneous choice behavior across families may have immediate impacts on academic achievement for students of parents exercising choice. To estimate the effect of attending a first choice school and how it relates to the importance placed on academics when selecting a school, we analyzed the subset of students choosing schools that were oversubscribed and limited our sample to students in marginal priority groups within those schools – priority groups for which lottery number alone determined initial admission. We ignore members of priority groups in which all students were either admitted or denied admission –

[19] For example, estimates from an exploded-logit model without random coefficients (but otherwise flexibly specified as in Table 4) yielded similar simulation results and similar (though less precise) estimates of how the impact of attending a first choice school varied with the weight placed on school test scores. Because we used individual-level data with multiple responses and wide variation in the choice set, we could specify the model very flexibly and identify significant heterogeneity in preferences even without the random coefficients through interactions with observable characteristics.

since the assignment of lottery numbers had no impact on their admission status. This allows us to use the random admission of students into a school, conditional on the school they chose, as an instrument for attending a first choice school.[20]

We began with a sample of 37,115 students entering grades four through eight. 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 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.

Table V compares baseline characteristics of students in the marginal priority group to other students in the district. Overall, students in the marginal priority group appear to be fairly representative of students who chose a non-guaranteed school. The only notable differences are that students in the marginal priority group were more likely to be eligible for lunch subsidies than students in the waitlisted group (reflecting the priority given to eligible students), and they applied to schools with higher test scores than did students in the admitted group (reflecting capacity constraints at schools with higher test scores). Not surprisingly, students choosing non-guaranteed schools differed from students who chose a guaranteed school: they had home-schools with lower test scores and higher proportions of minority and lunch-eligible students, and were more likely to be minority, poor, and doing poorly in school themselves. Thus, the marginal priority group should provide a reasonable estimate of the impact of attending one's first choice school for a typical student who chose a non-guaranteed school (i.e., treatment on the treated). In addition, the final row in Table V shows the mean weight placed on school scores ($\hat{\beta}_i^S$) for students in the marginal priority group versus other students in the district, calculated by

[20] In some schools, the marginal priority group will consist of those who attended the school the year before, free-or reduced-lunch eligible students, or students from the choice zone. The marginal priority group may also be different for different grade levels in a school.

equation (8). The mean $\hat{\beta}_i^S$ for students in the randomized group is very similar to that for students in the district as a whole.[21]

Table VI verifies the lottery randomization and examines the impact winning the lottery had on characteristics of the attended school in the 2002-2003 school year. For the experiment to be valid, baseline characteristics should be balanced across lottery winners and losers, differential attrition should be minimal, and winning the lottery should significantly increase the probability of attending a first choice school. The adjusted difference reported in Table VI is the estimated impact of winning the lottery on each of these outcomes, from a regression that included school lottery fixed effects to focus on differences within lottery and account for the fact that the probabilities of winning the lottery varied across lotteries (Rouse 1998). Our 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 first panel reports the adjusted difference between lottery winners and losers for various baseline characteristics. None of the coefficients are significant, confirming that winning the lottery was independent of baseline characteristics, as would be expected if it was randomly assigned within the marginal priority groups. The second panel of Table VI tests for differences in attrition, using presence in CMS at the end of the 2002-2003 school year as the dependent variable. The results show no evidence of differential attrition across lottery winners and losers.[22] The third panel of Table VI shows the impact of winning the lottery on the characteristics of the school attended at the end of the 2002-2003 school year. Lottery winners were 53 percentage points more likely to attend the first choice school than the lottery losers. This estimate is not equal to 100 percent for two reasons: first, some of those who were given the opportunity to attend the first choice did not do so, and second, some of those who were originally waitlisted at the first choice school were subsequently called off the waitlist. Students who won the lottery attended schools with approximately one-tenth of a student-level standard deviation higher test scores. Thus winning the lottery appears to be a valid instrument for

[21] The low mean weight for students in the admitted group is consistent with the fact that these students were typically choosing lower-scoring and capacity unconstrained schools, as can be seen from the differences in sample means for the Average Combined Scores of the chosen school (row 8 in Table V).
[22] Average attrition rates were fairly low at 9.8 percent and consistent with estimates of inter-county mobility rates from the Census.

attending a first choice school and significantly increased the test score of the school students attend.

To test the empirical implications of the model derived Section 3, we estimated the impact of attending a first choice school on a student's test score using equation (11):

$$(11)\qquad y_i = \delta_0 + z_i\delta + D_i\gamma_1 + \left(D_i\hat{\beta}_i^S\right)\gamma_2 + \upsilon_i$$

where winning the lottery and its interaction with $\hat{\beta}_i^S$ (the estimate weight each parent placed on school test scores) were instruments for attending a first choice school (D_i) and its interaction with $\hat{\beta}_i^S$. We control for baseline characteristics (to capture z_i), home-school fixed effects (to capture variation in δ_0) and lottery fixed effects (since winning the lottery is random within, but not between lotteries). Baseline characteristics included $\hat{\beta}_i^S$, measures of student performance from the prior year (math and reading score, had more than 18 absences, was suspended, was retained) and measures of student demographics (neighborhood income, African American, other minority, female, received lunch subsidy).[23] We did not correct the standard errors of our IV regressions for the presence of a generated regressor, because the estimation error in $\hat{\beta}_i^S$ (due to estimation error in the underlying parameters in the exploded-mixed-logit model) was less than 1 percent.

Table VII presents the results from estimating (11) by instrumental variables, using student combined average test score at the end of the 2002-2003 school year as the dependent variable. The key prediction from Section 3 was that the coefficient on the interaction should be positive ($\gamma_2 > 0$). The first two columns present results for all students, with and without the interaction between attending a first choice school and $\hat{\beta}_i^S$. The first column shows no significant average impact of attending a first choice school on own test score, similar to average treatment effects in prior studies (Cullen, Jacob, and Levitt (2006)). The point estimate is very close to zero (-0.005) and has a large standard error (0.050). The second column shows that the weight parents placed on school scores ($\hat{\beta}_i^S$) when selecting schools is a significant and positive determinant of how attending a first choice school affects test score outcomes. The regression estimates imply

[23] We do not report the coefficients on the baseline covariates, but the largest determinants of test scores was baseline test scores. Female was the only other student-level demographic characteristics that was consistently significant, and was positively associated with test scores.

that a one standard deviation increase (0.81) in the weight that an individual places on school test scores raises the treatment effect on the student's own test score by 0.066 standard deviations. For parents who placed no weight on test scores in their school choices, the coefficient on attending the first choice school implies a negative (although not significant) treatment effect – the students' test scores fall by 0.143 standard deviations if they attend the first choice school. These estimates imply a near zero impact (0.002 standard deviation score gain) of attending a first choice school on test scores for an average student with a $\hat{\beta}_i^S$ of 1.79, and a large positive effect on test scores (about 0.10) for students at the 95th percentile of the $\hat{\beta}_i^S$ distribution. A 0.1 standard deviation increase in a student's test score results is equivalent to a three to four percentile rank gain in test scores. Estimates of the impact that test scores have on future 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)).[24]

The second general prediction from our model is that if $\hat{\beta}_i^S$ is low because parents place low weight on a school's contribution to their child's achievement (as opposed to because school test scores are unrelated to their child's achievement), then the impact of attending a first choice school will depend on how other valued school characteristics are related to the academic quality of the school. To test this prediction, the last four columns of Table VII show estimates for two subgroups of parents: those for whom the estimated preferred percent minority at a school is less than 50 percent (primarily white students) and those for whom it is greater than 50 percent (primarily black students).[25] Since school test scores are negatively correlated with percent minority in a school, parents who prefer a school with a high proportion of African American students will tend to choose lower performing schools unless they place a high weight on school scores. In contrast, parents that prefer schools with a low proportion of African American students will tend to choose high performing schools regardless of the weight they place on school scores. In other words, the interaction effect between $\hat{\beta}_i^S$ and attending a first choice

[24] There is some evidence that students may experience a drop in test scores from switching schools, particularly moving from elementary to middle schools. (Cook et al. (2006), Rockoff and Lockwood (2009)). We estimate our main regression specification separately for rising-grade and non-rising-grade students (rising sixth graders versus all others), and find some evidence of an adverse mean impact of matriculating to middle school but overall the pattern of results are similar across the two subgroups.

[25] Posterior estimates of preferences for school racial composition were calculated in the same way as the $\hat{\beta}_i^S$ s.

school should have a negative intercept and a steeper slope for students who have strong preferences for predominantly African American schools.

Columns 3 and 4 show that the average treatment effect is positive for students whose parents prefer a predominantly white school, and there is no significant interaction with the weight that the parent places on test scores in their school choice. Among students of parents who prefer predominantly white schools, both relatively high- and low-$\hat{\beta}_i^S$ students experienced academic gains from attending the first choice school. In contrast, columns 5 and 6 indicate that students of parents who prefer a predominantly black school have a significant interaction between the estimated preference for academics and the treatment effect. High-$\hat{\beta}_i^S$ students experienced academic gains from attending the first choice school that are similar to students whose parents prefer a predominantly white school. In contrast, low-$\hat{\beta}_i^S$ students with parents that prefer a predominantly black school experience a negative effect on academic performance from attending the first choice school.

These results suggest that heterogeneity in the relative importance parents place on academics when choosing schools and the tradeoffs they face in their choice sets are both important determinants of the immediate gains from school choice. They imply that public school choice may have a negative impact on the academic achievement of the students it is most intended to help – minority students from disadvantaged backgrounds. The results in Table VII suggest that, on average, school choice widens rather than narrows the gap in achievement between low- and high-SES families exercising choice.

Not only do measures of parental choice behavior explain heterogeneous impacts from attending a first choice school, but they also highlight why subgroup impacts examined in prior studies vary through correlation between demographics and choice behavior. Table VIII shows estimates of the average treatment effect on student test scores in various subgroups of students defined on the basis of student demographics or characteristics of the school chosen. Prior studies have used these subgroups of students who on a priori grounds may have different underlying reasons for choosing schools that may be correlated with the expected treatment effect. Estimates for most of the subgroups remain insignificant. However, the estimated treatment effect is positive and significant for two of the subgroups (whites and students with above median income) and there is an apparent pattern of positive treatment effects for higher

SES students and students applying to higher-scoring schools. The pattern of subgroup impacts is strongly related to the average weight ($\hat{\beta}_i^S$) that parents place on school test scores. Column 2 of Table VIII reports the mean of $\hat{\beta}_i^S$ for students in each of the subgroups, and Figure 4 plots this mean against the subgroup estimates from column 1. The strong positive correlation between the two (correlation=0.89) suggests that differences in impacts across subgroups may be generated by differences in the underlying determinants of choice. Moreover, the large variation in $\hat{\beta}_i^S$ within each subgroup reported in column 3 of Table VIII suggests that any of these simple subgroup impacts will capture only a fraction of the heterogeneity in outcomes if differences in the weights parents place on academics are driving heterogeneous treatment effects.

This evidence highlights three advantages of using demand estimates to identify heterogeneous treatment effects versus subgroup estimation based on observables, such as race and income. First, using a single index, rather than estimating differences in impacts for an arbitrary number of subgroups, increases the precision with which we can identify heterogeneous treatment effects by exploiting all of the within- and between-subgroup variation in preferences. Second, the $\hat{\beta}_i^S$ incorporate information on parental rankings and the choice set they faced, distinguishing between students who pick a school because it is convenient versus students who repeatedly pick less convenient schools for academics reasons. Third, the demand estimates give us an economic interpretation of subgroup impacts, allowing us to evaluate the impact of school choice outside of the estimation sample, and to potentially redesign school choice plans to address differences in the underlying drivers of parental choice (Hastings and Weinstein (2008)).

5. Conclusion

One of the most important goals of public school choice plans is to increase academic performance for disadvantaged students by allowing them to attend higher-performing schools and by creating pressure on failing schools to improve through the threat of losing students. Both of these goals require that parents value academics and choose schools accordingly when offered the opportunity to do so. This paper departs from the prior empirical literature on public school choice by combining rich choice data, lottery assignments, and student outcomes to examine the implications that heterogeneous choice behavior had for equity and quality of educational opportunities for disadvantaged families.

We find that heterogeneity in the weights parents place on school test scores drive both demand-side pressure for schools to improve and immediate academic gains from attending an alternative school. Because low-SES families place less emphasis on academics when choosing schools, the school choice plan led to high demand response for high-performing schools serving high-SES families and little demand response for low-performing schools serving local, low-SES families. Thus, the demand response under public school choice would tend to increase education stratification in Charlotte rather than produce a competitive tide that lifts all boats. At the same time, allowing parents to choose an alternative school led to worse academic outcomes for many children of disadvantaged minority families who had to trade off academic strength of the school to attend a predominantly black school, but better academic outcomes for white children who faced no such tradeoff. Therefore, school choice widened rather than narrowed the gap in achievement between low- and high-SES families exercising choice.

Finally, understanding the economic underpinnings of heterogeneous treatment effects and heterogeneous pressures to improve academics can point us towards simple changes in public school choice design that may increase efficacy in providing greater equity and quality in public education. In general, the impact of school choice on academic outcomes will depend on both the willingness of parents to make tradeoffs and the extent to which the available school choices require such tradeoffs to be made. In other school districts, where parents face less stark tradeoffs or focus more on academic differences between schools, the outcomes of choice could be very different. Moreover, interventions such as changing the information set provided to low-income families (Hastings and Weinstein (2008)) or the strategic placement of schools may increase demand-elasticity among low-SES families and minimize the tradeoffs they face, allowing them to share in the potential benefits from public school choice.

<u>*References*</u>

Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez (2006): "Changing the Boston School Choice Mechanism: Strategy-proofness as Equal Access," Working Paper 11965, NBER.

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

Anderson, Simon P., André de Palma, and Jacques-François Thisse (1992): *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.

Angrist, Joshua and Guido Imbens (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association,* 90, 431-442.

Angrist, Joshua, Guido Imbens and Donald Rubin (1996): "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association,* 91, 444-472.

Armor, David J. and Peiser, B.M. (1998): "Interdistrict Choice in Massachusetts," in *Learning from School Choice,* ed. by Paul Peterson and Bryan Hassel. Washington, DC: Brookings Institution Press, 157-185.

Ballou, Dale (2007): "Magnet Schools and Peers: Effects on Student Achievement," Working Paper, Vanderbilt University.

Bayer, Patrick, Fernando Ferreira and Robert McMillan (2004): "Tiebout Sorting, Social Multipliers and the Demand for School Quality," Working Paper 10871, NBER.

Belfield, Clive R., and Henry M. Levin (2002): "The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States," *Review of Educational Research*, 72, 279-341.

Berry, Steven, James Levinsohn, and Ariel Pakes (2004): "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112, 68-105.

Betts, Julian R., Lorien A. Rice, Andrew C. Zau, Y. Emily Tang, and Cory R. Koedel (2006): *Does School Choice Work? Effects on Student Integration and Achievement*. San Francisco, CA: Public Policy Institute of California.

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

Borghans, Lex, Angela Lee Duckworth, James J. Heckman and Bas ter Weel (2009): "The Economics and Psychology of Personality Traits," *Journal of Human Resources*. 43: 972-1059.

Cook, Philip J., Robert MacCoun, Clara Muschkin and Jacob Vigdor (2006): "Should Sixth Grade be in Elementary or Middle School? An Analysis of Grade Configuration and Student Behavior," Working Paper 12471, NBER.

Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt (2006): "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries," *Econometrica*, 74, 1191-1230.

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

Della Vigna, Stefano (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, (Forthcoming).

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

Epple, Dennis and Richard Romano (2002): "Educational Vouchers and Cream Skimming," Working Paper 9354, NBER.

Farrell, Joseph and Carl Shapiro (1990): "Horizontal Mergers: An Equilibrium Analysis," *American Economic Review*, 80, 107-126.

Glazerman, Steven (1997): "Determinants and Consequences of Parental School Choice" Unpublished Working Paper, University of Chicago, Harris School of Public Policy.

Glazerman, Steven, and Robert H. Meyer (1995): "Public School Choice in Minneapolis" in *Midwest Approaches to School Reform: Proceedings of a Conference Held at the Federal Reserve Bank of Chicago October 26-27, 1994*, ed. by Thomas A. Downes, and William A. Testa. Chicago, IL: Federal Reserve Bank of Chicago, 110-126.

Greene, Jay P., Paul E. Peterson, and Jiangtao Du (1997): "The Effectiveness of School Choice: The Milwaukee Experiment," Harvard University Education Policy and Governance Occasional Paper 97-1.

Hanushek, Eric A. (1981): "Throwing Money at Schools," *Journal of Public Policy Analysis and Management*, 1, 19-41.

Hanushek, Eric A., and Steven G. Rivkin (2003): "Does Public School Competition Affect Teacher Quality?," in *The Economics of School Choice*, ed. by Caroline M. Hoxby. Chicago, IL: The University of Chicago Press.

Hastings, Justine S., and Jeffrey M. Weinstein (2008): "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of Economics,* 123, 1373-1414.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger (2006): "Gender and Performance: Evidence from School Assignment by Randomized Lottery," *American Economic Review Papers and Proceedings,* 96, 232-236.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger (2007a): "Parental Preferences and School Competition: Evidence from a Public School Choice Program," Working Paper No. 11805, NBER.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger (2007b): "Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery," Working Paper No. 12145, NBER.

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

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

Heckman, James J., Sergio Urzua, and Edward Vytlacil (2006): "Understanding Instrumental Variables in Models with Essential Heterogeneity," *Review of Economics and Statistics*, 88, 389-432.

Hoxby, Caroline M. (2000): "Does Competition Among Public Schools Benefit Students and Taxpayers?," *American Economic Review,* 90, 1209-1238.

Hoxby, Caroline M. (2002): "School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?," Working Paper 8873, NBER.

Hoxby, Caroline M. (2003): "School Choice and School Productivity. Could School Choice Be a Tide that Lifts All Boats?," in *The Economics of School Choice,* ed. by Caroline M. Hoxby. Chicago, IL: The University of Chicago Press, 287-342.

Hyslop, Dean R., and Guido W. Imbens (2001): "Bias from Classical and Other Forms of Measurement Error," *Journal of Business and Economic Statistics,* 19, 475-481.

Kane, Thomas J., and Douglas O. Staiger (2002): "The Promise and Pitfalls of Using Imprecise School Accountability Measures," *Journal of Economic Perspectives,* 16, 91-114.

Kleitz, B., G.R. Weiher, K. Tedin, and R. Matland (2000): "Choices, Charter Schools and Household Preferences," *Social Science Quarterly* 81, 846-854.

McFadden, Daniel, and Kenneth Train (2000): "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15, 447-470.

Nechyba, Thomas (1999): "School Finance Induced Migration Patterns: The Impact of Private School Vouchers," *Journal of Public Economic Theory*. 1, 5-50.

Nechyba, Thomas (2000): "Mobility, Targeting and Private School Voucher," *American Economic Review*, 90, 30-146.

Nechyba, Thomas (2003): "School Finance, Spatial Income Segregation and the Nature of Communities," *Journal of Urban Economics* 54, 61-88.

Pathak, Parag A., and Tayfun S*ö*nmez (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98, 1636-1652.

Peterson, Paul E., David Myers, and William G. Howell (1998): "An Evaluation of the New York City School Choice Scholarship Program: The First Year," Cambridge, MA, Mathematica Policy Research and the Harvard Program on Education Policy and Governance.

Rockoff, Jonah E. and Ben B. Lockwood (2009): "Stuck in the Middle: The Educational Impacts of Grade Configuration," Unpublished Manuscript, Columbia University.

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

Roth, Alvin E., and Ido Erev (1995): "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8, 164-212.

Rothstein, Jesse (2006): "Good Principals or Good Peers? Parental Evaluations of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions," *American Economic Review*, 94, 1333-1350.

Rouse, Cecilia E. (1998): "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program," *Quarterly Journal of Economics*, 113, 553-602.

Schneider, M., M. Marschall, P. Teske and C. Roch (1998): "School Choice and Culture Wars in the Classroom: What Different Parents Seek from Education" *Social Science Quarterly* 79, 489-501.

Sanford, Eleanor E. (1996): *North Carolina End-of-Grade Tests Technical Report #1*. Public Schools of North Carolina, State Board of Education, Department of Public Instruction.

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

Vanourek, G., B.V. Manno, and C.E. Finn (1998): “Charter Schools as Seen by Students, Teachers and Parents” in *Learning from School Choice,* ed. by Paul Peterson and Bryan Hassel. Washington, DC: Brookings Institution Press, 187-211.

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

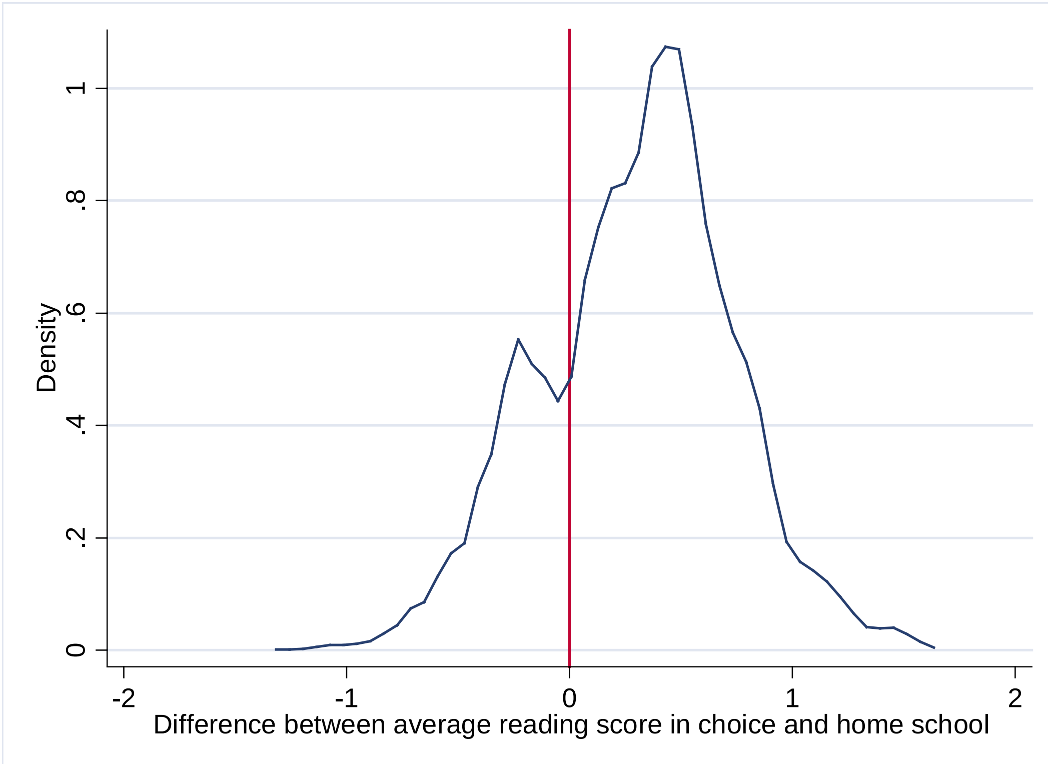

Figure 2a: Elementary Schools: Simulated Change in Number of Students Choosing School *j* When Average Standardized Score at School *j* Increases by 0.34 Points

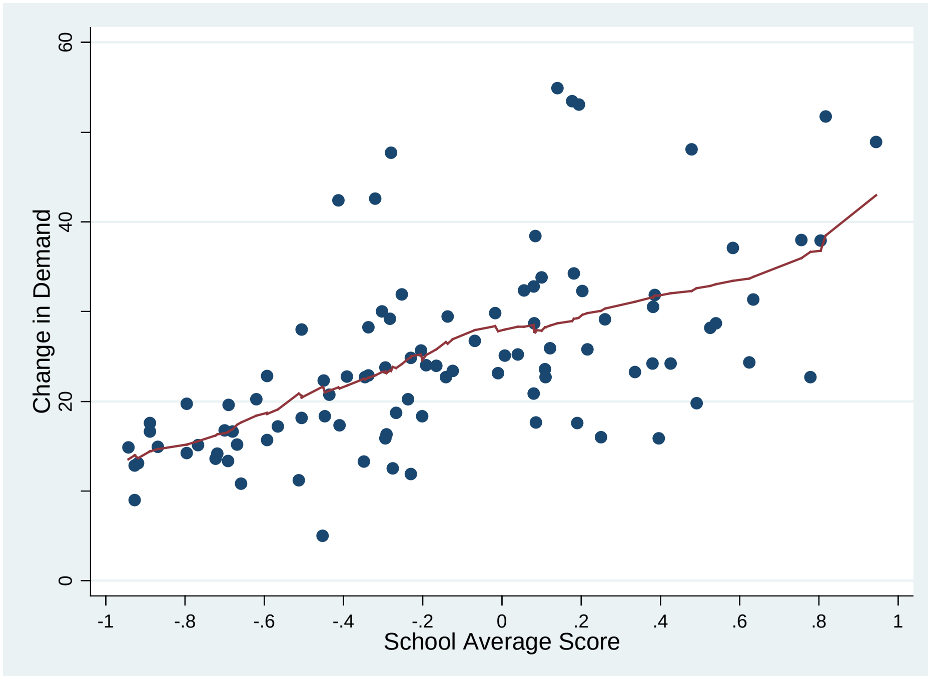

Figure 2b: Middle Schools: Simulated Change in Number of Students Choosing School *j* When Average Standardized Score at School *j* Increases by 0.334Points

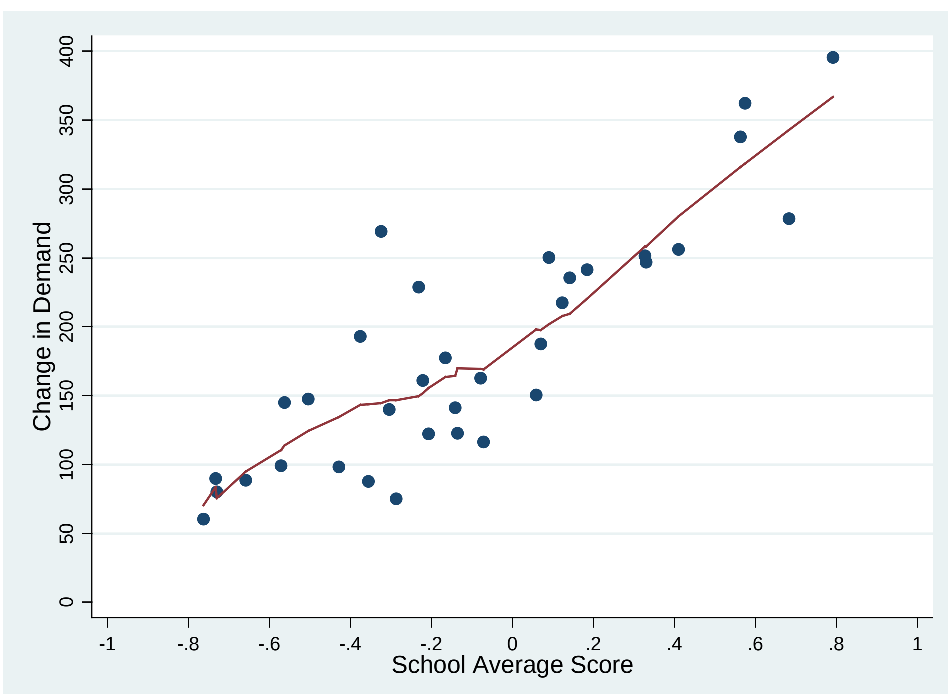

Figure 3: Average 2002 Standard Deviation Scale Score for Additional Students Who Choose School *j* in Response to 0.34 Point Increase in Average Score at School *j*

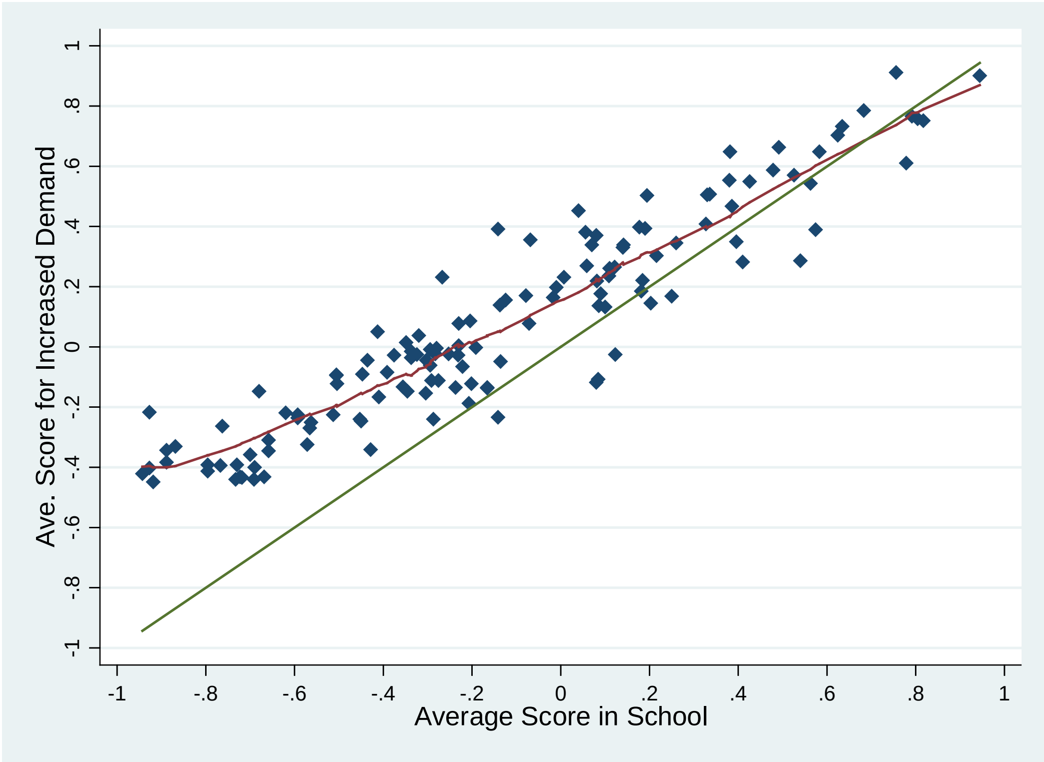

Figure 4: Subgroup Estimates of the Effect of Attending a First-Choice School

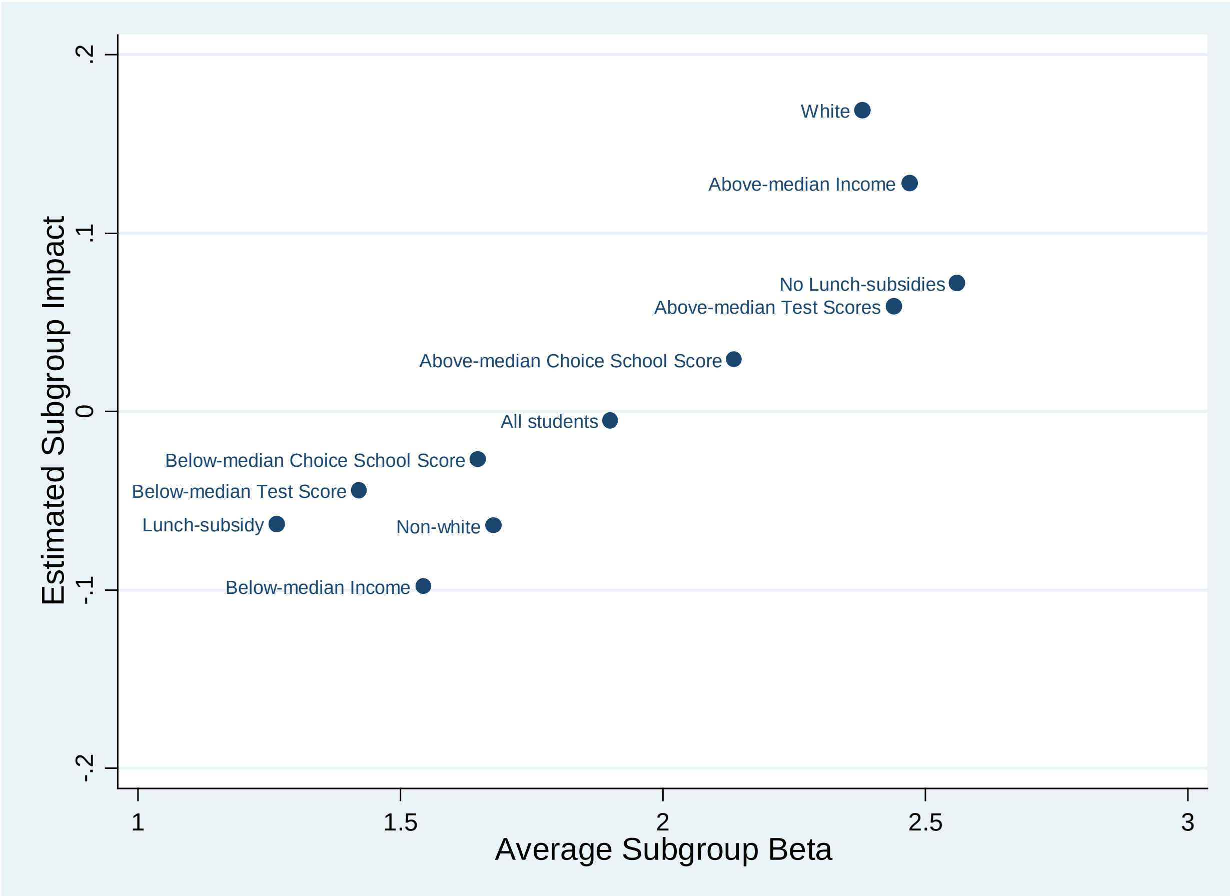

Table I: Descriptive Statistics for Students, Schools, and Parental Choices

	Not Receiving Lunch Subsidies		Receiving Lunch Subsidies	
	White	Not White	White	Not White
Student Characteristics				
Baseline test score	0.638	-0.091	-0.089	-0.613
	(0.825)	(0.839)	(0.849)	(0.799)
Neighborhood Income	$73,804	$50,612	$52,763	$36,455
	(25,869)	(21,502)	(22,556)	(16,230)
Student-School Characteristics				
Driving Distance to the 'Home-School'	3.298	2.915	2.725	2.445
	(2.085)	(1.777)	(1.759)	(1.625)
Driving Distance to Closest 'non-Home-School'	3.752	3.091	3.004	2.269
	(2.008)	(1.765)	(1.808)	(1.393)
Home School Ave. Test Score	0.239	-0.160	-0.147	-0.385
	(0.384)	(0.368)	(0.377)	(0.357)
Range of School Test Scores in 6 miles	0.670	0.817	0.827	1.018
	(0.503)	(0.438)	(0.448)	(0.367)
Highest Scoring School in 6 miles	0.530	0.288	0.317	0.314
	(0.342)	(0.317)	(0.329)	(0.268)
Choice Characteristics				
Chose Home School	0.647	0.405	0.529	0.361
Submitted 1 Choice	0.512	0.254	0.321	0.181
Submitted 2 Choices	0.202	0.186	0.203	0.149
Submitted 3 Choices	0.286	0.559	0.475	0.670
Observations	16,187	6,194	2,134	12,372

Notes: Standard deviations are given in parentheses. Baseline test score is the average of each student’s standardized test score in reading and in math on the North Carolina End of Grade Exams. Neighborhood income is the median block group income for families of the student’s own race according to the 2000 US Census. Driving distances from student to schools were calculated in MapInfo using geocoded residential locations, school locations, and Census TIGER road network files for Mecklenburg county. All school level test scores are reported in student-level standard deviation units. Each school’s score is calculated using the average baseline combined test score (2001-2002 school year) of students who attended the school in the 2002-2003 school year. White includes Asian students who constitute less than 4% of the student population. Not White includes Hispanic and multi-racial students who constitute approximately 6% of the student population. Lunch recipients are defined as students who qualified for either free- or reduced-price federal lunch subsidies.

Table II: Explanatory Variable Definitions

Variable	Description
School Test Score	Average of the 2001-2002 student-level standardized scale score for End of Grade Math and Reading exams for students in school *j* in the 2002-2003 school year. This is the average test score variable described below across all students in school *j*.
Student Baseline Score	The sum of student *i*'s scale score on End of Grade math and reading exams in the 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.
Distance	Driving distance from student *i* to school *j* calculated using MapInfo with Census 2000 TIGER/Line files.
Home School	An indicator if the school choice is designated as the student's Home School Choice.
Choice Zone School	An indicator if the school is a school in the student's choice zone.
Last Year's School	An indicator if the school choice was the student's attended school in the 2001-2002 school year. Students transitioning from Elementary to Middle school do not have a last year's school by construction.
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

Variable	Mean	St.Dev.	Mean of Student-level Mean	St. Dev. of Student-level Mean
Distance	13.015	6.755	12.973	3.852
Busing Provided	0.254	0.435	0.251	0.043
School Test Score	-0.112	0.453	-0.104	0.028
Student Score	0.056	0.989	0.054	0.988
Student Score* School Score	-0.003	0.463	-0.003	0.106
Neighborhood Income (demeaned, thousands of dollars)	5.091	27.566	5.165	27.599
Income * School Score	-0.564	13.076	-0.538	2.973
N	2,342,254		36,887	

Notes: Baseline test score is the average of each student's standardized test score in reading and in math on the North Carolina End of Grade Exams. Neighborhood income is the median block group income for families of the student's own race according to the 2000 US Census. Income used in the analysis is demeaned by the district-wide mean of $51,000 and divided by 1,000. All school level test scores are reported in student-level standard deviation units. Each school's score is calculated using the average baseline combined test score (2001-2002 school year) of students who attended the school in the 2002-2003 school year.

Table IV: Estimates from Exploded Mixed Logit Model of Demand for Schools

		Not Receiving Lunch Subsidies		Lunch Subsidy Recipient	
Variable	Parameter	White	Not-White	White	Not-White
Preferences for Test Scores					
School Score	Mean Elementary	1.461	2.103	0.102	1.444
	Mean Middle	1.888	2.714	1.157	1.398
	St.Dev.	0.402	0.204	0.664	0.321
Income* School Score	Mean Elementary	0.017	0.014	0.004	0.010
	Mean Middle	0.015	0.013	0.010	0.004
	St.Dev.	--	--	--	--
Own Baseline Score*School Score	Mean Elementary	0.378	0.379	0.152	0.379
	Mean Middle	0.587	0.451	0.309	0.243
	St.Dev.	--	--	--	--
Preferences for Proximity					
Distance	Mean Elementary	-0.385	-0.335	-0.488	-0.357
	Mean Middle	-0.355	-0.245	-0.303	-0.235
	St.Dev.	0.083	0.070	0.183	0.096
Home School	Mean Elementary	2.110	1.801	1.804	1.709
	Mean Middle	2.150	1.659	1.953	1.714
	St.Dev.	0.826	0.650	0.314	0.259
Preferences for Race					
Percent Black	Mean Elementary	3.162	5.565	1.812	4.414
	Mean Middle	4.912	6.064	3.174	3.559
	St.Dev.	3.114	1.512	2.306	1.037
Percent Black Squared	Mean Elementary	-4.582	-2.666	-3.424	-2.032
	Mean Middle	-6.523	-3.977	-4.102	-2.856
	St.Dev.	--	--	--	--
Implied Preferred % Black	Mean Elementary	0.345	1.044	0.265	1.086
	Mean Middle	0.376	0.762	0.387	0.623
	St.Dev.	0.340	0.283	0.337	0.255
Other Preferences					
Last-Year's School	Mean Elementary	4.830	4.505	4.128	4.022
	Mean Middle	3.343	2.696	3.062	1.986
	St.Dev.	2.606	2.858	3.330	3.332
In Choice Zone	Mean Elementary	1.666	1.325	2.170	1.711
	Mean Middle	1.213	1.306	1.902	1.562
	St.Dev.	1.117	1.223	1.590	1.279
Estimated Correlation Coefficients					
Corr(Distance,School Score)		0.449	0.886	0.272	0.561
Corr(Distance,Home School)		0.107	0.136	0.281	0.341
Corr(School Score,Home School)		-0.215	-0.053	-0.668	-0.585

Notes: Estimates generated by Simulated Maximum Likelihood. Estimates for means and standard deviations of the distribution of preferences for school characteristics. All estimates are significant at the 5% level or higher,. Standard errors are reported in Appendix Table B.I. Distribution of preference on distance is constrained to follow a negative lognormal distribution and parameter estimates of the lognormal are reported.

Table V: Comparison of Student Characteristics

	All Students	Chose Guaranteed School	Chose Non-Guaranteed School		
			Admitted	Randomized	Waitlisted
Student Characteristics					
Fraction Minority	50.3%	39.7%	68.7%	65.3%	60.1%
Fraction Lunch-subsidy Recipient	39.3%	30.6%	60.1%	50.6%	33.6%
Basline Standardized Test Scores	0.054	0.211	-0.244	-0.073	-0.097
Weight Placed on Academics - $\hat{\beta}_i^S$	1.882	1.991	1.557	1.901	1.983
Choice School Characteristics					
Average Test Score	0.076	0.122	-0.078	0.097	0.144
Fraction Lunch-subsidy Recipient	40.7%	38.2%	50.5%	36.7%	35.5%
Fraction African-American	43.9%	40.2%	52.9%	46.2%	42.6%
Home School Characteristics					
Average Test Score	-0.059	0.06	-0.257	-0.185	-0.226
Fraction Lunch-subsidy Recipient	47.3%	40.6%	59.2%	53.2%	55.8%
Fraction African-American	46.9%	40.6%	57.6%	53.9%	55.3%
Assignment Characteristics					
Assigned to 1st Choice School	86.2%	100.0%	100.0%	40.6%	0.0%
Assigned to Home-School	70.0%	100.0%	0.0%	44.4%	74.8%
N	36,887	22,000	8,602	2,897	3,380

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

Table VI: Baseline Characteristics by Treatment and Control Group

Variable	**Mean**	**Adjusted Difference**
Panel 1: Baseline Characteristics		
Black	0.597	0.014
		(0.021)
Free or Reduced Lunch	0.505	-0.014
		(0.012)
Median Income ($1000s) by Race	48.993	-0.700
and Block-Group in 2000 Census		(0.700)
Combined Baseline Test Score	-0.093	0.003
		(0.030)
Home School Average Combined Score	-0.122	0.20
		(0.014)
Home School Fraction Free or Reduced Lunch	0.532	0.001
		(0.007)
Home School Fraction Minority	0.614	-0.009
		(0.006)
Panel 2: Attrition		
Not Attending CMS in 2002-2003	0.098	-0.021
		(0.011)
Panel 3: School Attended in 2002-2003		
First-Choice School	0.460	0.530***
		(0.053)
Test Score of School Attended	-0.073	0.129**
		(0.040)
Number of Students		2,894

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 the 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: IV Estimates of Impact of Attending First-Choice School with Heterogeneous Treatment by Weight Placed on Academics in Choice Decision

Dependent Variable: ***Combined Score in Spring 2003***	**All Students**		**Students Who Prefer School Less Than 50% Black**		**Students Who Prefer School at Least 50% Black**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Attended First-Choice School	-0.006 (0.049)	-0.176 (0.091)	0.135* (0.066)	0.252 (0.209)	-0.050 (0.061)	-0.230 (0.116)
*Weight** attended First-Choice School		0.090** (0.032)		-0.050 (0.072)		0.103* (0.047)
P-value for Interaction with *Weight*		0.007		0.487		0.033
Joint P-Value on Reported Coefficients	0.907	0.012	0.047	0.105	0.415	0.101
Observations	2,591	2,591	720	720	1,871	1,871

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 its interaction with the weight that the parent places on test scores (*Weight*= $\hat{\beta}_i^S$) 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, a direct control for the student's *Weight* estimate, and the student baseline covariates listed in the text. 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: Subgroup Estimates of Effect of Attending a First-Choice School

Sample	IV Estimate of Effect of Attending First-Choice School on Combined Test Score	Mean $\hat{\beta}_i^S$	Standard Deviation of $\hat{\beta}_i^S$	Number of Students
	(1)	(2)	(3)	(4)
All Students	-0.006	1.901	0.831	2,591
	(0.049)			
Race:				
Non-White	-0.064	1.678	0.747	1,800
	(0.056)			
White	0.169*	2.38	0.799	791
	(0.073)			
Income:				
Below Median	-0.098	1.545	0.644	1,611
	(0.058)			
Above Median	0.128*	2.47	0.778	980
	(0.063)			
Free Lunch Eligibility				
Eligible	-0.063	1.264	0.371	1,312
	(0.078)			
Not Eligible	0.072	2.561	0.638	1,279
	(0.043)			
Baseline Test Score				
Below Average	-0.044	1.422	0.549	1,363
	(0.055)			
Above Average	0.059	2.44	0.761	1,228
	(0.065)			
First-Choice School Combined Score				
Below Median	-0.027	1.647	0.733	1,262
	(0.084)			
Above Median	0.029	2.136	0.847	1,329
	(0.040)			

Notes: Each row reports estimates for a different student subgroup, as indicated. Column 1 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 the text. 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). Column 2 reports the average weight that parents place on test scores ($\hat{\beta}_i^S$) in their school choice decision. The third column reports the standard deviation of $\hat{\beta}_i^S$ among families in each of the subgroup categories.

Appendix A: Robustness Checks

There are a number of reasons that the estimates in Table IV may not accurately represent student preferences. In this section we present a series of robustness checks that address three particularly important concerns that could lead our estimates to understate the strength of preferences for academic quality at a school. These specification checks do not find evidence that the preference estimates in Table IV are biased by any of these three important concerns. This Appendix is taken from the Robustness Section in Hastings, Kane and Staiger (2007a). Please see that paper for further details.

Alternative Measures of Academic Quality

In Table IV, we used average base-year (2002) test scores among students attending the school in 2003 as a measure of academic performance that influences parents' choices. However, if this is a crude or incorrect proxy for the information available to parents, then our estimates may understate the extent to which parents care about academic performance in choosing a school. Table A.I presents exploded-logit results for our original measure plus three alternative measures of school academic performance. To preserve space, we only report the coefficients on the measure of school academic performance (and its interactions with income and student baseline score), along with the log likelihood for each model as an indicator of overall fit. The first panel is our baseline specification. This specification assumes that parents correctly forecasted which students would choose each school and used these students' scores from the prior year as an indicator of how good the school would be in terms of academics. The second panel uses the average score in 2003 of the students in the school in 2003. This measure implies that parents correctly foresaw student sorting and outcomes and made their choices based on that. The third panel uses the 2002 average scores for students in each school in 2002. This measure implies that parents used historical student assignment and outcomes as the best indicator of future school performance. The final panel uses a "value-added" measure of each school's impact on academic achievement in the prior year. We estimated each school's "value-added" by regressing a student's test score performance in 2003 on math and reading performance in the prior year, demographic characteristics, grade fixed effects, and fixed effects for each school. The fixed effect estimated for each school represents our estimate of a school's average impact on student performance.

The results in Table A.I show that the preference estimates across the three measures of average test scores are quite similar. The specification in panel 1, using our base specification, has the highest likelihood value, and in this sense it fits the choice data best. The value-added measure in panel 4 does not fit the observed choice data well at all. This may not be surprising since this statistic is not available to parents and is not easy to calculate given the available average score statistics. Cullen, Jacob, and Levitt (2006) report that high-demand schools in Chicago's high school choice program tend to be schools with high average test scores, not high value-added. Note that our specifications using test score levels do allow parents to make a crude race-adjusted "value-added" calculation when choosing a school (since these specifications separately control for the racial makeup of the school).

Overall, the results from Table A.I suggest that our preference estimates are not particularly sensitive to the use of reasonable alternative measures of school academic performance. Because our baseline specification provides the best fit of the data, we will rely on this specification for the final simulations.

Strategy

As noted above, parents may have had an incentive to misrepresent their true preferences. If they understood the allocation mechanism, a parent with an undesirable home-school might want to hedge against being assigned to the home-school. They would do so by picking less desirable schools than they actually prefer – trading off desirability for an increased chance of being admitted. This strategy could make it appear that low-income families (with lower-performing home-schools) under-value academics even though they do not. They pick lower performing schools in order to increase their chance of admission, not because they place a lower weight on academics. However, it is not at all clear that parents had the information or experience in the first year of choice to understand how to exploit the incentives of the allocation mechanism. Parents did not know their lottery numbers or the assignment mechanism. In addition, parents were instructed by the district to list the schools they wanted on their choice form.[1]

We test for the presence of strategic behavior in the first year of choice by exploiting the redrawing of school boundaries. Many of those who lived in the same contiguous school assignment zone in 2001-2002, were given different school assignments in 2002-2003. Hence, among those with the same school assignments who lived in the same neighborhood in 2001-2002, some students experienced positive or negative shocks to the quality of their guaranteed school. Table A.II shows the average difference for students who had positive versus negative shocks to their home-school quality given that they lived in a 2001-2002 contiguous assignment zone that was split into new assignment zones. The table shows that the difference in scores was large and significant for many of the students affected by redistricting. If strategy was a major component of parental choices, we would expect to see very different choices for those with negative versus positive shocks to the quality of their home-school, given a contiguous 2001-2002 school assignment zone. In particular, we should see significantly lower weight placed on average test scores for students who had a negative shock to the average quality of their home-schools.

Table A.III presents an exploded-logit specification using choice data for the subsample of students who lived in 2001-2002 assignment zones that were split by redistricting. We estimated the model using the same covariates as we do in our base specification (Table IV), however we add full interactions with an indicator if the parent experienced a negative shock to the test score of their guaranteed school as a results of redistricting. We report all coefficients related to test scores, and look for negative and significant coefficients on interactions with "Redistricting Loss" as evidence of strategic behavior. We also report joint significance of all coefficients on interactions with "Redistricting Loss" in the final table row. The results show that there no significant difference in the school-score preference estimates for redistricting losers versus winners consistent with strategy.[2] For all groups except the Not-White-Lunch

[1] This stands in contrast with other long-standing and limited choice programs, such as the Boston Public School choice program which told parents to consider carefully what schools they chose to list (Abdulkadiroglu et al. (2006)). Abdulkadiroglu et Al. (2006) also show that many parents appear to behave non-strategically in Boston Public Schools limited and long-standing school choice program and that these parents would benefit most from strategizing on their first choice by picking a less popular school first.

[2] We also tested for differences in preferences using a reduced form regression of the average test score of the chosen school on the average test score at the home-school, controlling for 2001-2002 school assignment fixed effects and student demographic information. This compares the average scores of first-choice schools within 2001-

group the interactions with "Redistricting Loss" are jointly insignificant. In the case of the Not-White-Lunch group (last column), the coefficient on the redistricting loss interaction with school test scores is significant for elementary school students, but *positive* instead if negative. In addition, the point estimates do not follow the expected pattern if strategy was a key component in choosing schools. Coefficients on interactions with Redistricting Loss are positive as often as they are negative, showing no clear pattern in support of strategic choice.

Overall, we do not see evidence that parents with poorer home-school assignments hedged their bets. It is possible that such strategic behavior may develop over time as parents became more familiar with the system.[3] Our findings suggest that there is little evidence of hedging behavior in the first year of choice when parents most likely did not understand the allocation mechanism.

Residential Sorting

The exogenous reassignment of nearly half of the students in the district is also useful for testing the extent to which residential sorting affects our preference estimates. Residential sorting may lead us to overstate preferences for proximity if parents had already sorted to live next to the schools they prefer. What we interpret as a strong preference for proximity influencing school choice may actually be the opposite – strong preference for a school influencing proximity. Both redistricting and the multiple choices in our data will help identify preferences for proximity from preferences for other school characteristics.

To test for the potential effects of residential sorting on our estimates, we re-estimate our model for the subsample of students who were reassigned (whose school assignments under the busing plan in 2001-2002 were different from their home-school in 2002-2003).[4] Table A.IV provides summary statistics comparing the reassigned sample to the sample of students who were not reassigned. Because of the nature of the prior system of bussing, students who were reassigned were much more likely to be non-white and eligible for lunch subsidies. (The busing

---

2002 assignment zones across students with positive and negative shocks to home-school quality. We find no significant difference in the average scores of schools chosen across redistricting winners and losers.

[3] Evidence from laboratory experiments using simple extensive-form games between small numbers of players indicates that it takes time for players to learn how to play the game. In games with incomplete information on others' payoffs, learning and convergence to the perfect equilibrium is slower and sometimes does not occur (Roth and Erev (1995)).

[4] These students include both students whose 2001-2002 school assignment zone was split into two new home-school zones as well as those whose entire 2001-2002 school assignment zone was reassigned to a new home-school.

plan often assigned students living in neighborhoods with large concentrations of minority students to attend school in neighborhoods with lower concentrations.) But within the four demographic groups, the reassigned students looked similar to those who were not reassigned in terms of baseline test scores and median income. More interestingly, parents of reassigned 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 parents of reassigned students have less preference for their home-school: Parents of children not reassigned 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, parents of reassigned children faced a less clear choice since their last year's school was no longer their home-school.

Table A.V.a reports results from the exploded-logit model estimated on the sample of students who were reassigned, and Table A.V.b provides the corresponding standard errors on those estimates. The most striking feature of these estimates is their similarity to estimates from the full sample. The only apparent exception is in the White-Lunch column for elementary school students, but this exception is most likely due to the relatively small sample size in this subgroup. Overall, these estimates suggest that endogenous residential location is not a major source of bias in this data. In addition, the similarity of the results is not too surprising if we believe our model is using the information in multiple choices to identify preferences. Recall from Table I that a substantial fraction of parents who listed their home-school as their first choice also listed subsequent choices. For these parents, multiple choices simulate reassignment whether or not they were actually reassigned.

Other Robustness Checks

A range of alternative specifications yielded similar quantitative and qualitative results. 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. In addition, estimations using splines in distance indicated that the linear functional form used in our model was appropriate. We experimented with a range of alternative proxies for academic quality of a school. Using closely related measures, such as 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. In addition, estimates using splines in school test scores indicated that the linear model fit the data well for most segments of the population. 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 by about 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 exploded-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 the covariance estimates could often be unstable, yielding corner solutions in some circumstances (near perfect correlation between some idiosyncratic preferences). 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 these terms are not of first order importance to simulations of demand. We thus focused on the key covariance terms that could impact demand response to school test score improvement – correlations between idiosyncratic preferences for school test scores and proximity (home-school and driving distance).

Table A.I: Comparing Alternative Measures of School Academic Achievement

		Not Receiving Lunch Subsidies				Receiving Lunch Subsidies			
		White		Not White		White		Not White	
Score Measure: Spring 2002 Scores of Students in School in Spring 2003									
-Log Likelihood		-42570		-33784		-9677		-82724	
School Score	Elem.	1.508	(0.067)	1.939	(0.066)	0.172	(0.127)	1.376	(0.054)
	Mid.	1.776	(0.065)	2.462	(0.063)	1.067	(0.128)	1.405	(0.050)
Test Score*School Score	Elem.	0.172	(0.044)	0.304	(0.049)	0.114	(0.089)	0.355	(0.030)
	Mid.	0.382	(0.035)	0.364	(0.039)	0.245	(0.073)	0.226	(0.026)
Income*School Score	Elem.	0.014	(0.002)	0.015	(0.002)	0.004	(0.004)	0.012	(0.002)
	Mid.	0.008	(0.001)	0.010	(0.002)	0.006	(0.003)	0.009	(0.001)
Score Measure: Spring 2003 Scores of Student in School in Spring 2003									
-Log Likelihood		-42729		-34028		-9678		-82895	
School Score	Elem.	1.406	(0.066)	1.779	(0.066)	0.041	(0.124)	1.249	(0.054)
	Mid.	1.420	(0.060)	1.995	(0.058)	1.016	(0.120)	1.141	(0.047)
Test Score*School Score	Elem.	0.193	(0.044)	0.357	(0.050)	0.121	(0.089)	0.350	(0.031)
	Mid.	0.381	(0.034)	0.378	(0.038)	0.239	(0.071)	0.218	(0.026)
Income*School Score	Elem.	0.015	(0.002)	0.015	(0.002)	0.003	(0.004)	0.011	(0.002)
	Mid.	0.007	(0.001)	0.008	(0.002)	0.005	(0.003)	0.006	(0.001)
Score Measure: Spring 2002 Scores of Student in School in Spring 2002									
-Log Likelihood		-43005		-34617		-9687		-83228	
School Score	Elem.	1.069	(0.051)	1.061	(0.050)	0.468	(0.087)	0.856	(0.045)
	Mid.	0.526	(0.051)	0.703	(0.046)	-0.048	(0.089)	0.191	(0.039)
Test Score*School Score	Elem.	0.226	(0.041)	0.365	(0.049)	0.105	(0.090)	0.383	(0.030)
	Mid.	0.481	(0.034)	0.386	(0.038)	0.231	(0.069)	0.191	(0.026)
Income*School Score	Elem.	0.004	(0.001)	0.010	(0.002)	0.002	(0.004)	0.013	(0.002)
	Mid.	0.009	(0.001)	0.005	(0.002)	0.005	(0.002)	0.000	(0.001)
Value Added: Average Regression Adjusted Gains in Test Scores from 2002-2003[a]									
-Log Likelihood		-43751		-34987		-9697		-83385	
School Score	Elem.	-0.363	(0.215)	-0.635	(0.193)	-1.955	(0.339)	-0.405	(0.186)
	Mid.	-1.265	(0.219)	-0.295	(0.190)	0.853	(0.374)	-0.964	(0.178)
Test Score*School Score	Elem.	0.331	(0.203)	0.697	(0.227)	0.206	(0.393)	0.192	(0.139)
	Mid.	1.936	(0.182)	1.777	(0.201)	0.616	(0.392)	0.335	(0.137)
Income*School Score	Elem.	0.050	(0.007)	0.033	(0.009)	-0.003	(0.015)	0.044	(0.007)
	Mid.	-0.008	(0.006)	-0.015	(0.008)	-0.027	(0.013)	-0.048	(0.007)

Notes: Estimates from rank-ordered logit (exploded-logit) using the same data and specification as Table IV, but changing the measure of school academic performance to each of the alternative measures specified in row headings above. For the third specification, Score Measure: Sping 2002 Scores of Students in School in Spring 2002, the ethnic composition of the school choice in 2002 was also used in place of the 2003 ethnic composition. In the case that the school-choice did not exist in 2001-2002 school year, the 2002 characteristics of students attending the school in Spring 2003 were substituted instead. [a] Value Added calculated as school fixed effects in a regression of 2003 standardized test scores on 2002 standardized scores, controlling for student characteristics, such as race, lunch recipient status, and grade level. Empirical Bayes measures of Value Added were calculated and were correlated with Value Added ($\rho = 0.95$).

Table A.II: Differences in Scores Across Redistricted Polygons

		Ave. Difference in 2002-2003 Home-school Scores in a 2001-2002 School Assignment Polygon	
		Mean	St. Dev.
Not Receiving Lunch Subsidies	White	0.2925	0.3165
	Not White	0.2570	0.2552
Receiving Lunch Subsidies	White	0.2891	0.2836
	Not White	0.2866	0.3111

Table A.III: Exploded-logit Estimates of School Choice, Interactions with Redistricting Loss to Home-School Test Scores

	White, Not-Lunch	Not-White, Not-Lunch	White, Lunch	Not-White, Lunch
Elementary Student Interacted with:				
School Score	1.714**	2.182**	-0.030	1.213**
	(0.299)	(0.192)	(0.521)	(0.159)
Interaction with Redistricting Loss	-0.445	-0.205	0.095	0.369*
	(0.331)	(0.226)	(0.606)	(0.164)
SchoolScore*StudentScore	0.090	0.312**	-0.187	0.484**
	(0.130)	(0.102)	(0.282)	(0.076)
Interaction with Redistricting Loss	0.098	-0.028	0.493	-0.203
	(0.157)	(0.154)	(0.365)	(0.115)
SchoolScore*Income	0.020	0.027**	-0.003	0.008
	(0.011)	(0.006)	(0.009)	(0.006)
Interaction with Redistricting Loss	-0.006	-0.006	0.002	0.017*
	(0.011)	(0.007)	(0.010)	(0.008)
Middle School Student Interacted with:				
School Score	1.804**	2.575**	1.106*	1.399**
	(0.369)	(0.356)	(0.440)	(0.395)
Interaction with Redistricting Loss	0.035	-0.306	-0.463	-0.275
	(0.424)	(0.331)	(0.498)	(0.367)
SchoolScore*StudentScore	0.302*	0.328**	0.119	0.124*
	(0.148)	(0.075)	(0.228)	(0.050)
Interaction with Redistricting Loss	-0.095	0.089	0.343	0.105
	(0.198)	(0.149)	(0.249)	(0.080)
SchoolScore*Income	0.007	0.012**	0.009*	0.011
	(0.004)	(0.004)	(0.004)	(0.006)
Interaction with Redistricting Loss	-0.006	0.003	-0.005	-0.002
	(0.008)	(0.006)	(0.008)	(0.007)
Observations	421,828	198,168	66,573	417,187
Number of groups	7,162	3,220	1,056	6,634
Joint P-Value	0.803	0.896	0.491	0.009

Notes: * significant at 5%; ** significant at 1%. Robust standard errors in parentheses

Table A.IV: Summary Statistics Comparing Reassigned and Non-Reassigned Students

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

Table A.V.a: Coefficients from Exploded-Logit Estimates of School Choice for Full Sample and Redistricted Subsample

		White, Not-Lunch		Not-White, Not-Lunch		White, Lunch		Not-White, Lunch	
		All	Redistricted	All	Redistricted	All	Redistricted	All	Redistricted
School Test Score	Elem	1.508	1.078	1.939	1.902	0.172	-0.430	1.376	1.095
	Mid	1.776	1.757	2.462	2.506	1.067	0.902	1.405	1.387
School Score*Student Score	Elem	0.172	0.221	0.304	0.393	0.114	-0.052	0.355	0.354
	Mid	0.382	0.478	0.364	0.348	0.245	0.050	0.226	0.197
School Score*Income	Elem	0.014	0.021	0.015	0.020	0.004	0.006	0.012	0.007
	Mid	0.008	0.008	0.010	0.013	0.006	0.008	0.009	0.010
Driving Distance	Elem	-0.254	-0.226	-0.229	-0.225	-0.303	-0.261	-0.230	-0.234
	Mid	-0.234	-0.255	-0.188	-0.200	-0.210	-0.216	-0.175	-0.177
Home-Schol	Elem	1.690	2.417	1.545	2.086	1.239	1.951	1.388	1.931
	Mid	1.656	1.864	1.387	1.598	1.388	1.731	1.354	1.630
Choice Zone School	Elem	1.225	1.213	0.969	0.960	1.426	1.581	1.241	1.253
	Mid	0.893	0.751	0.994	1.048	1.243	1.153	1.165	1.176
Last-year's School	Elem	3.239	3.562	3.116	3.475	2.346	2.956	2.627	3.175
	Mid	2.200	2.300	1.904	2.152	1.582	2.220	1.421	1.781
School % Black	Elem	1.574	1.138	4.660	5.105	-0.359	-1.933	2.835	2.502
	Mid	3.144	3.024	4.545	4.249	1.214	0.094	1.632	0.711
School % Black Squared	Elem	-1.916	-2.003	-2.107	-2.577	-0.829	0.020	-0.978	-1.041
	Mid	-4.003	-3.875	-2.650	-2.564	-1.923	-1.238	-1.079	-0.411
Observations		1,025,071	360,641	384,401	193,540	137,948	52,829	794,834	497,838
Number of groups		16,187	6,059	6,194	3,253	2,134	861	12,372	7,974

Notes: Results from rank-ordered logit model of school choice using the same sample as the one that produces results in the main text.

Table A.V.b: Standard Errors from Exploded-Logit Estimates of School Choice for Full Sample and Redistricted Subsample

		White, Not-Lunch		Not-White, Not-Lunch		White, Lunch		Not-White, Lunch	
		All	Redistricted	All	Redistricted	All	Redistricted	All	Redistricted
School Test Score	Elem	(0.067)	(0.106)	(0.066)	(0.093)	(0.127)	(0.194)	(0.054)	(0.073)
	Mid	(0.065)	(0.094)	(0.063)	(0.082)	(0.128)	(0.171)	(0.050)	(0.061)
School Score*Student Score	Elem	(0.044)	(0.071)	(0.049)	(0.069)	(0.089)	(0.141)	(0.030)	(0.037)
	Mid	(0.035)	(0.053)	(0.039)	(0.050)	(0.073)	(0.102)	(0.026)	(0.030)
School Score*Income	Elem	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)	(0.005)	(0.002)	(0.002)
	Mid	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
Driving Distance	Elem	(0.004)	(0.007)	(0.006)	(0.008)	(0.011)	(0.017)	(0.004)	(0.005)
	Mid	(0.003)	(0.005)	(0.004)	(0.005)	(0.007)	(0.011)	(0.003)	(0.003)
Home-Schol	Elem	(0.034)	(0.053)	(0.047)	(0.061)	(0.076)	(0.111)	(0.029)	(0.035)
	Mid	(0.025)	(0.039)	(0.032)	(0.043)	(0.059)	(0.086)	(0.022)	(0.027)
Choice Zone School	Elem	(0.039)	(0.065)	(0.044)	(0.062)	(0.088)	(0.140)	(0.027)	(0.033)
	Mid	(0.024)	(0.035)	(0.028)	(0.036)	(0.056)	(0.075)	(0.019)	(0.021)
Last-year's School	Elem	(0.035)	(0.055)	(0.044)	(0.060)	(0.076)	(0.118)	(0.029)	(0.036)
	Mid	(0.032)	(0.047)	(0.039)	(0.050)	(0.075)	(0.106)	(0.027)	(0.033)
School % Black	Elem	(0.269)	(0.434)	(0.344)	(0.496)	(0.598)	(0.937)	(0.237)	(0.303)
	Mid	(0.271)	(0.423)	(0.307)	(0.418)	(0.610)	(0.865)	(0.218)	(0.265)
School % Black Squared	Elem	(0.281)	(0.435)	(0.303)	(0.430)	(0.535)	(0.802)	(0.188)	(0.235)
	Mid	(0.256)	(0.387)	(0.254)	(0.338)	(0.503)	(0.695)	(0.169)	(0.200)

Appendix B: Standard Errors for Demand Parameter Estimates

Appendix Table B.1: Standard Errors for Parameter Estimates Presented in Table IV

	Not Receiving Lunch Subsidies				**Lunch Subsidy Recipient**			
	White		**Not-White**		**White**		**Not-White**	
	Coeff.	St.Error	Coeff.	St.Error	Coeff.	St.Error	Coeff.	St.Error
Elementary School Mean Preference								
Distance (normal)	-1.093	(0.0003)	-1.257	(0.0000)	-0.947	(0.0001)	-1.252	(0.0001)
Last Year's School	4.830	(0.0018)	4.505	(0.0003)	4.128	(0.0003)	4.022	(0.0019)
Home-school	2.110	(0.0005)	1.801	(0.0000)	1.804	(0.0003)	1.709	(0.0003)
In Choice Zone	1.666	(0.0010)	1.325	(0.0001)	2.170	(0.0002)	1.711	(0.0010)
School Score	1.461	(0.0008)	2.103	(0.0002)	0.102	(0.0007)	1.444	(0.0005)
Own Baseline Score*School Score	0.378	(0.0012)	0.379	(0.0000)	0.152	(0.0001)	0.379	(0.0000)
Income* School Score	0.017	(0.0001)	0.014	(0.0000)	0.004	(0.0000)	0.010	(0.0000)
Percent Black	3.162	(0.0597)	5.565	(0.0009)	1.812	(0.0185)	4.414	(0.0081)
Percent Black Squared	-4.582	(0.0642)	-2.666	(0.0005)	-3.424	(0.0172)	-2.032	(0.0066)
Middle School Mean Preference								
Distance (normal)	-1.173	(0.0004)	-1.572	(0.0000)	-1.423	(0.0001)	-1.672	(0.0002)
Last Year's School	3.343	(0.0011)	2.696	(0.0002)	3.062	(0.0003)	1.986	(0.0010)
Home-school	2.150	(0.0002)	1.659	(0.0001)	1.953	(0.0000)	1.714	(0.0005)
In Choice Zone	1.213	(0.0001)	1.306	(0.0000)	1.902	(0.0000)	1.562	(0.0004)
School Score	1.888	(0.0005)	2.714	(0.0003)	1.157	(0.0005)	1.398	(0.0000)
Own Baseline Score*School Score	0.587	(0.0005)	0.451	(0.0000)	0.309	(0.0002)	0.243	(0.0000)
Income* School Score	0.015	(0.0001)	0.013	(0.0000)	0.010	(0.0000)	0.004	(0.0000)
Percent Black	4.912	(0.0263)	6.064	(0.0004)	3.174	(0.0111)	3.559	(0.0084)
Percent Black Squared	-6.523	(0.0328)	-3.977	(0.0001)	-4.102	(0.0101)	-2.856	(0.0077)
Cholesky factors								
Distance (normal)	0.523	(0.0001)	0.574	(0.0000)	0.679	(0.0000)	0.668	(0.0002)
Last Year's School	2.606	(0.0010)	2.858	(0.0003)	3.330	(0.0001)	3.332	(0.0029)
Home-school	-0.785	(0.0028)	0.597	(0.0008)	-0.178	(0.0020)	0.012	(0.0009)
In Choice Zone	1.117	(0.0001)	1.223	(0.0000)	1.590	(0.0000)	1.279	(0.0008)
School Score	0.359	(0.0005)	0.095	(0.0000)	0.639	(0.0004)	0.266	(0.0002)
Own Baseline Score*School Score	3.114	(0.0100)	1.512	(0.0003)	2.306	(0.0027)	1.037	(0.0008)
Income* School Score	0.307	(0.0001)	0.059	(0.0000)	0.396	(0.0002)	-0.227	(0.0000)
Percent Black	-0.024	(0.0003)	-0.122	(0.0001)	-0.423	(0.0004)	-0.112	(0.0010)
Percent Black Squared	-0.734	(0.0041)	-0.650	(0.0008)	-1.487	(0.0004)	-1.599	(0.0010)

Notes: Robust standard errors in parentheses, calculated using the numerical hessian and score of the likelihood function.