School Choice, School Quality and Postsecondary Attainment

David J. Deming

Justine S. Hastings

Thomas J. Kane

Douglas O. Staiger

July 2011[1]

We study the impact of a public school choice lottery in Charlotte-Mecklenburg (CMS) on postsecondary attainment. Students were assigned to their neighborhood school by default but could apply to attend another public school in the district. We develop a model where the academic gain from attending a chosen school depends on the gain in school quality and on an idiosyncratic match between the student and the school. Students with low quality neighborhood schools are more likely to apply to another school, and they have higher expected achievement gains from choice. We test the predictions of the model using CMS administrative data matched to the National Student Clearinghouse (NSC), a nationwide database of college attendance. Among applicants with low quality neighborhood schools, lottery winners are more likely than lottery losers to graduate from high school, attend a four year college, and earn a bachelor's degree. They are about twice as likely to earn a degree from an elite university. The results suggest that school choice can have an important impact on students' longer-term life chances, but only when students gain access to schools that are better on observed dimensions of quality.

[1] We would like to thank Lawrence Katz, Susan Dynarski, Brian Jacob and Christopher Jencks for reading early drafts of this paper and for providing essential guidance and feedback. We benefited from the helpful comments of Josh Angrist, Amitabh Chandra, Bridget Long, Erzo Luttmer, Dick Murnane, Seth Richards-Shubik, Lowell Taylor and seminar participants at Harvard University, Columbia University, the University of Michigan, RAND and the Association for Public Policy Analysis and Management (APPAM) and Society for Research on Educational Effectiveness (SREE) meetings. Special thanks to Brian Kovak for helpful discussions of the model and to Andrew Baxter at CMS for help with matching the student files to the NSC.

School choice is an increasingly important feature of the U.S. education landscape. Policies such as intra-district open enrollment, charter schools and vouchers for private school tuition decouple neighborhood residence and school attendance, loosening the monopoly on education held by local school districts and potentially increasing competition between schools for students (Hoxby, 2003). Yet even among the set of studies with high-quality research designs, evidence on the benefits of school choice options is inconsistent across policies and settings.

Studies of lottery-based admission to public schools in Chicago and Charlotte-Mecklenburg have found mixed evidence of impacts on academic and non-academic outcomes (Cullen, Jacob and Levitt 2006; Cullen and Jacob 2007; Hastings, Kane and Staiger 2008; Deming 2011). Similarly, studies of lottery-based admission to charter schools in New York City and Boston find impacts on test scores that range from modest to large (Hoxby and Rockoff 2004; Hoxby and Murarka 2009; Abdulkadiroglu et al 2011; Dobbie and Fryer 2011; Angrist et al 2011; Hastings, Neilson and Zimmerman 2011). Voucher programs in Milwaukee, New York City and Washington DC have found mixed impacts on student achievement (Witte 1997; Rouse 1998; Howell and Peterson 2002; Krueger and Zhu 2004; Wolf et al 2009). In nearly all of these studies, the outcome of interest is achievement test scores. Booker et al (2011) follow the earlier literature on Catholic schools (e.g. Evans and Schwab 1995, Neal 1997, Altonji et al 2005) by using proximity as an instrument for attendance at a charter school, and Lavy (2010) uses differences-in-differences and regression discontinuity approaches to estimate the impact of public school choice in Israel on high school graduation and other outcomes. While both studies find substantial impacts of choice on educational attainment, we are aware of no study that examines the impact of choice on postsecondary attainment using a lottery-based design.

In this paper we study the impact of winning an admissions lottery to attend a public high school in Charlotte-Mecklenburg Schools (CMS), which implemented an open enrollment public school choice program in the Fall of 2002. Students were guaranteed admission to their neighborhood school but were allowed to choose and rank up to three other schools in the district, including magnet schools. When demand for school slots exceeded supply, allocation was determined by random lottery. We model the decision to apply to a non-neighborhood school as a function of students' expected achievement gains from attending the new school. In our model, student achievement is determined by individual ability, school quality, and the quality of the match between the student and the school. Families living in neighborhoods with high quality schools are less likely to apply to another school, and those families that apply do so primarily to escape a "bad match". In contrast, families with low-quality neighborhood schools make choices that are predictable based on observed measures of school quality.

The main prediction of the model is that the expected achievement gains from choice are greater for students with low quality neighborhood schools, and that the magnitude of the difference is proportional to the relative weight placed on school quality in the achievement function. Intuitively, if very little of a school's impact is common across students, then observed measures of school quality will do a poor job of predicting improvements in academic outcomes. On the other hand, when school quality is an important determinant of student achievement, the measured quality difference between two schools will be a good approximation of the likely benefits of attending one versus the other.

We test the implications of the model using a long and detailed panel of administrative data from CMS, which we match to the National Student Clearinghouse (NSC), a national database of postsecondary enrollment. The match is done directly using personal identifying information, so dropouts and transfers can be followed and attrition from the sample is limited only by the coverage of the NSC.[2] Rising 9th graders in the Fall of 2002 could potentially have completed up to five years of post-secondary enrollment, enabling us to look at degree completion over a reasonable horizon. We also estimate the impact of winning a school choice lottery on high school test scores, graduation, and other school outcomes.

The central finding of this paper is that students from low quality neighborhood schools benefit greatly from choice. Students from these neighborhoods who win the lottery to attend another school are more likely to graduate from high school, attend a four-year college and earn a bachelor's degree. They are about twice as likely to earn a degree from an elite institution such as UNC-Chapel Hill or Duke. We find no impact on 9th grade test scores, but relatively large gains on other school outcomes such as grade point average, math course-taking, and absences. In contrast, we find no evidence of benefits from choice for students in other neighborhoods.

The results show that access to a higher quality school can improve students' longer-term life outcomes. While our tests are far from definitive, we find no evidence that improved match quality increases achievement. This is important since finding better matches for students is one way that school choice could increase allocative efficiency and enhance welfare (Hoxby, 2000). In this setting, school choice only increases achievement when students gain access to better schools, although we are unable to measure competitive effects of choice that enhance productivity in all schools (Hoxby, 2003; Chan and McMillan 2009; Figlio and Hart 2010).

---

[2] The NSC covers over 90 percent of 4-year college enrollment nationwide and in North Carolina. CMS sent every student who had ever been enrolled in any grade to the NSC for matching. See the Data Appendix for details.

Our model also provides a framework for understanding the conditions under which lottery-based impacts can provide a good approximation to the average causal effect of a school on all students. Lotteries allow us to estimate the impact of randomly assigned admission, but within the selected sample of students who chose to apply. We derive a simple expression for the expected achievement gain among lottery applicants that is equal to the measured quality difference between schools plus a selection term that is a function of unobserved match quality. Furthermore, we show that the magnitude of the selection term decreases as the probability of choice approaches one. Intuitively, for a neighborhood school of low enough quality, even families with very good matches will find it beneficial to choose another school, and the lottery sample will be selected from the full distribution of match quality. In that case the impact of winning the lottery will be equal to the impact from randomly assigning all students to schools, a finding that is similar in spirit to the "identification at infinity" result in Chamberlain (1986) and Heckman (1990).

The average causal impact of a school on student outcomes is a parameter of direct policy interest, and underlies the logic behind value-added measures of school effectiveness (e.g. Raudenbush and Bryk, 2002). We estimate a variety of school value-added measures and test their performance against the impact of winning the lottery. In our preferred specification, we cannot reject the hypothesis that school quality estimates equal the lottery-based treatment effects for three of the five main outcomes in the paper. The ideal way to validate school value-added measures would involve large-scale random assignment of students to schools (with forced compliance). Since this is both infeasible and unethical, lottery-based estimates may be the most promising alternative. This paper provides a framework for understanding when such estimates can be used to judge the effectiveness of schools for all students.

2. Background and Data Description

Charlotte-Mecklenburg is the 20th largest school district in the nation. The school district encompasses all of Mecklenburg County, which includes both the inner city areas of Charlotte and the more affluent suburbs surrounding it. Thus neighborhoods in CMS and vary widely by race and income. In 1971, the Supreme Court (in *Swann v. Charlotte-Mecklenburg Board of Education)* ruled that this variation resulted in neighborhood schools that were *de facto* segregated, and for over 30 years CMS schools were desegregated under a court order that bused students all over the district to preserve racial balance in the schools. Particularly at the high school level, this meant in practice that inner-city

and largely African-American neighborhoods were divided up and bused out to more affluent and white suburbs or to “midpoint schools” in different parts of the county.

After several years of legal challenges, the historic court order was overturned and the busing plan was terminated. In December of 2001 the CMS School Board voted to move forward with district-wide open enrollment for the 2002-2003 school year. Because CMS was no longer allowed to use race explicitly in student assignments, the school boundaries were redrawn as contiguous neighborhood school zones. Children who lived within each zone received guaranteed access to their neighborhood school, which was usually (but not always) the closest to their home address. This resulted in a change in assigned neighborhood high school for about 35 percent of households.

The school choice lottery took place in the spring of 2002. To maximize the number of parents that exercised choice, CMS conducted an extensive information campaign. They held a well-advertised fair at the Charlotte convention center, set up “choice booths” in local shopping malls, and sent volunteers door-to-door in low-income and non-English speaking neighborhoods to talk to families about the plan (Hastings, Kane and Staiger 2008). CMS also developed a comprehensive booklet with information about each school, as well as smaller brochures for individual schools. As a result, over 95 percent of parents submitted a choice application in the spring of 2002.

Parents were allowed to submit up to three choices, which included schools as well as special programs within schools.[3] Students were guaranteed admission to their neighborhood school, and admission for all other students was subject to grade-specific capacity limits that were set by the district beforehand but were not known to families at the time of the lottery (Hastings, Kane and Staiger 2008). Children with siblings already in enrolled in a school also received guaranteed access. CMS was also divided into four “choice zones” and free transportation was provided by the district only within each zone, although families could provide their own transportation to any school.[4] The district expanded capacity at schools where they anticipated high demand in an attempt to give everyone their first choice. Still, many high schools were oversubscribed.

When demand for slots among non-guaranteed applicants exceeded supply, admission was allocated by random lotteries that occurred within the following lexicographic priority groups:

1) Students that attended the school in the previous year and their siblings.

[3] Parents who listed 3 non-guaranteed choices were automatically assigned their “home school” as a 4th choice.
[4] The choice zones were constructed so that there was at least one predominately white suburban and at least one predominantly black inner-city school in each zone.

2) Free or reduced price lunch eligible (i.e. low-income, "FRPL") students applying to schools where less than half of the previous year's school population was FRPL.
3) Students applying to a school within their own choice zone.

Applicants were sorted by priority group according to these rules, and then assigned a random lottery number. Slots at each school were first filled by students with guaranteed access, and then remaining slots were allocated within each priority group according to students' lottery numbers. If all members of a priority group could be offered admission, slots were allocated to the next priority group in the order of lottery numbers. CMS administered the lottery centrally and applied an algorithm known as a "first choice maximizer" (Abdulkadiroglu and Somnez, 2003). This meant that CMS first allocated slots to all those who listed a school as their first choice and only then moved to second choices. As the name indicates, this maximized the percentage of students who received their first choice, but it also meant that students who lost the lottery to attend their first choice school often found that their second choice had already been filled up in the previous round. While there is the potential for strategic choice with this type of lottery mechanism, Hastings, Kane and Staiger (2006) show that this is not likely to have been a large problem in CMS, at least in the first year of the choice plan.

We match these files to information on college attendance from the National Student Clearinghouse (NSC), a non-profit organization that maintains enrollment information for over 90 percent of colleges nationwide. In collaboration with CMS, we provided each student's full name, date of birth and (when applicable) high school graduation date, which the NSC used to match to its database. Rather than restricting to CMS graduates, we matched all students who were old enough to have been enrolled in college, regardless of the last grade they been enrolled in CMS. The NSC data contain information on enrollment spells and degrees for all covered colleges that a student attended. Information is available on full or part-time status and degree receipt in some cases. Unfortunately, we have no information about college experiences such as grades or choice of major. Although not all colleges provide information to the NSC, the coverage is very good in North Carolina and the surrounding states. The Data Appendix contains a list of colleges by coverage and a detailed analysis of the match process using data from the Department of Education's Integrated Postsecondary Data Source (IPEDS) as a reference.[5] Students who leave CMS are followed in these data. Attrition is subject

[5] The major two-year college in Charlotte, Central Piedmont Community College (CPCC), did not provide information to the NSC until 2006. To fill in this gap, we obtained enrollment data directly from CPCC for all years. This data was more detailed than what colleges typically provide to the NSC. The data from CPCC contain

only to the NSC's coverage and the quality of the match. Unless coverage is differential for lottery winners and losers, the results may be attenuated but will not be biased.

We match the lottery applicant file, with individual lottery numbers and priority groupings, to a panel of administrative data from CMS. The data span seven years before and after the choice lottery (from 1995 to 2009) and contain detailed information on students' enrollment histories, test scores, course-taking and other outcomes of interest. The North Carolina Department of Public Instruction requires all school districts to assemble and send them a standardized set of files under the state's accountability regime. In addition to enrollment records, this includes students' scores on standardized End-of-Grade (EOG) exams in math and reading for grades 3-8 and End-of-Course (EOC) exam scores in high school subjects such as Algebra I and II, Geometry, English I, Biology and Chemistry. These tests are administered to all public school enrollees and schools are required to use the scores as some component of students' grades. Importantly, these reporting requirements help to ensure that CMS's records are of high quality, with longitudinally linked and consistent student records.

3. Sample Characteristics and School Quality Measures

CMS received high school lottery applications from 29,584 high school students. We first limit the sample to students who were enrolled in any CMS school in the previous year. About six percent of applicants come from outside the district, and these students are much less likely to be enrolled in CMS the following fall. Since previous enrollment status is fixed at the time of the lottery, this sample restriction does not affect the validity of the randomization. We also exclude from the sample the small number of students who apply to special education programs. Finally, about five percent of this remaining sample does not show up in any CMS school in the fall of 2002. Although these students can still be matched to the NSC data and are included in those analyses, we have no other outcome information for them. This restriction results in a sample of 25,564 rising 9th-12th graders, or 86 percent of the original sample. Finally, we exclude rising 12th graders from the analysis sample because of concerns about correct randomization.[6] This leaves an analysis sample of 20,021 students.

---

information on type of enrollment (i.e. degree-seeking or correspondence course), credit accumulation and GPA. We also used the CPCC data to verify the NSC's match process. See the Data Appendix for details.

[6] We have analyzed the individual choice lotteries to confirm that random numbers determine offers of admission, and have found that they hold perfectly except for in the 12th grade. In reviewing the historical documentation and in conversation with CMS, we have some concern that additional slots may have been made available at schools for rising 12th grade applicants. Thus we exclude from the analysis the 85 rising 12th grade applicants who were in marginal priority groups (about 4 percent of the lottery sample).

Table 1 presents descriptive statistics for the 14 neighborhood school zones in CMS. School zones vary widely in income, demographic composition, average student test scores and postsecondary attainment. Median household income ranges from $94,799 (in 2000 dollars) in Providence to $27,278 in West Charlotte. Similarly, the share of nonwhite and free lunch (an indicator of poverty) students ranges from less than 10 percent to over 90 percent. Average 9th grade test scores in math and English have a range of around 1.5 standard deviations. Most strikingly, the share of students attending a college that is judged as "very competitive" or better by the Barron's Rankings ranges from nearly 50 percent in the highest income neighborhood to 2 percent in the poorest neighborhood.[7]

The last three rows show the same measures for students who attend one of the three magnet high schools in CMS. Since these schools have no neighborhood zones, they are attended by students from one of the other 14 school zones in CMS. Magnet high schools serve predominately nonwhite students in the lower end of the income distribution. This is due in part to their location in the central city, whereas many of the higher-income schools are located in the surrounding suburbs. Magnet schools rank around the district average on measures such as average test scores, high school graduation and college attendance.

Despite having higher average test scores and higher rates of college attendance, schools in higher income neighborhoods are not necessarily better. Families that choose to live in these neighborhoods are likely to make other investments in their children that improve academic outcomes. Thus a simple comparison of mean outcomes across neighborhood school zones will overstate differences in school quality. We estimate each school's contribution to student outcomes using a general value-added framework:

$$A_{ij} = \beta X_{ij} + \nu_{ij}\ , \quad \text{where} \quad \nu_{ij} = \mu_j + \varepsilon_{ij} \tag{1}$$

[7] School in North Carolina with a rating of "very competitive" or higher include Appalachian State University, Duke University, Elon University, North Carolina State University, UNC-Asheville, UNC-Chapel Hill, UNC-Wilmington, and Wake Forest University. Four year colleges in North Carolina that are less than "very competitive" include East Carolina University, Fayetteville State University, North Carolina A&T University, North Carolina Central University, UNC-Charlotte, UNC-Greensboro, UNC-Pembroke, Western Carolina University and Winston-Salem State University. "Most competitive" schools in North Carolina are Davidson, Duke, UNC-Chapel Hill and Wake Forest.

The X_{ij} vector includes basic demographic controls and third order polynomials in state-standardized 8th grade math and reading end-of-grade (EOG) exams.[8] The residual υ_{ij} is a linear combination of a school effect μ_j (which is constant for all students in the school by assumption) and an idiosyncratic student-level error term (ε_{ij}). Under the strong assumption that there is no sorting on the unobserved determinants of A_{ij} after conditioning on X_{ij}, the parameter μ_j is a causal estimate of a school's mean impact on the academic achievement of its students (Raudenbush and Willms 1995, Kane and Staiger 2008, Rothstein 2010).

We estimate equation (1) by ordinary least squares (OLS) and take the mean residual across students as our estimate of μ_j.[9] We estimate equation (1) separately for 2 cohorts of first-time 9th graders in Fall 2002 and 2003, and we take the average of the two estimates as our measure of school quality. [10] We attribute each student only to the first high school that they attended, and we exclude students from grades 10-12 from the estimates to prevent bias from differential rates of grade attainment and dropout. Finally, we exclude the approximately 1,800 students who are in our lottery sample to avoid any mechanical correlation between μ_j and the lottery treatment effects.

In principle, we could use data from earlier cohorts of 9th grade students to make our estimates more precise. However, because of the sudden change in school composition caused by neighborhood rezoning and choice, school effects that differ by year seem particularly plausible in this setting. Thus our preferred specification is based only on the post-rezoning cohorts of 2003 and 2004.

Our estimates of school quality for the five main outcomes of the paper are in columns 2 through 6 of Table 2. The two test score outcomes are in state standard deviation units, and the educational attainment outcomes are measured in percentage points. If the estimates are unbiased, they can be interpreted as the change in each outcome for a randomly chosen student attending that school, relative to the average school in CMS. Column 7 reports the average of the five school quality measures, scaled to have a mean of zero and standard deviation of one. The neighborhood school zones are sorted by the average quality measure in Column 7.

---

[8] The demographic covariates are median household income in Census tract, and indicators for race (black, Asian, Latino, white), gender, free or reduced price lunch, special education and limited English proficiency. All covariates are collected in the 8th grade year.

[9] In principle, a more efficient alternative to OLS would be to use a hierarchical linear model (HLM) that accounts for the nested random effects implicit in our error structure (Raudenbush and Bryk, 2002). Because of the large number of students in each school, HLM and OLS yield virtually identical estimates.

[10] For the remainder of the paper, we will refer to our value-added estimates as measures of "school quality". We recognize that there are many different dimensions of quality that we cannot measure.

The correlation between family income and the average quality measure (0.39, in Column 7) is of moderate size but much weaker than the unadjusted comparison in Table 1. Magnet schools appear particularly strong after adjusting for student characteristics. All 3 magnets are above average on value-added measures of high school graduation and four-year college attendance, and two of the three rank in the top five schools in CMS on the average quality measure in Column 7. However, since families must actively apply to magnet schools and students often must travel a greater distance from home, quality estimates for magnet schools are particularly likely to be biased by unobserved sorting of motivated families into the school.

Column 8 shows the share of students assigned to each neighborhood school that list that school as their first choice. We interpret this as a revealed preference measure of school quality. Families who are already guaranteed access to a high-quality school will be less likely than others to want to leave it.[11] Reassuringly, the probability that a student will choose their neighborhood school is negatively correlated (-0.27) with the school's average quality estimate. However, the negative correlation between choice and income is much stronger (-0.9). This raises the question of how much parents actually observe about quality and their motivations for choice. Recent evidence (including some from CMS) indicates that providing information on basic measures of quality can have a significant impact on parental decision-making (Hastings and Weinstein 2009; Andrabi et al 2009). [12]

4. Empirical Strategy

Of the 20,021 students in the analysis sample, about 48 percent (9,719) chose a school other than the neighborhood school to which they were assigned. About 49 percent (4,736) of these students applied to schools that were not oversubscribed, and thus were automatically admitted. Another 32 percent (3,118) were in priority groups where no one was admitted. That leaves the remaining 19 percent (1,865) of students who applied to schools where admission was determined by random lottery numbers. Our main results are based on this sample of students.

---

[11] It is not a perfect measure of quality, for 3 reasons: 1) If a family lives much closer to the neighborhood school than to any other, they may find it too costly to switch to a better school; 2) Families may care about factors other than quality when selecting a school (friends, sports teams, etc.) 3) Families may not observe school quality – in fact, they may sort directly on observed but imperfect proxies for quality such as income.

[12] In fact, in modeling school choice, Hastings, Kane and Staiger (2008) test several measures of elementary and middle school achievement within their choice model and find no impact of traditional value added calculations on choice, though parents do place significant implicit weights on average test scores and racial composition of schools.

We begin by following the standard approach in lottery-based studies of school choice, which estimate the average impact of winning the lottery across multiple schools and grades (Cullen, Jacob and Levitt 2006; Hoxby and Rockoff 2004; Hoxby and Murarka 2009; Abdulkadiroglu et al 2011; Deming 2011). We estimate:

$$A_{ij} = \delta W_{ij} + \beta X_{ij} + \Gamma_j + \varepsilon_{ij} \quad (2)$$

Where W_{ij} is an indicator variable that is equal to one if student i has a winning lottery number for admission to school j, X_{ij} is a vector of pre-lottery covariates that is included only for balance, Γ_j is a set of lottery fixed effects, and ε_{ij} is a stochastic error term.[13] We use only first choices in the model, so the number of observations in the regression is simply equal to the number of students in the sample. In principle we could estimate a nested model that incorporates multiple choices and accounts for students' "risk sets" (Abdulkadiroglu et al 2011). However, since students who lost the lottery to attend their first choice school were generally shut out of other oversubscribed schools, there is almost no randomization on 2nd and 3rd choices.

Lottery fixed effects Γ_j are necessary to ensure that the *ex ante* probability of admission to a first-choice school does not differ between lottery winners and losers (Rouse 1998). If, for example, savvy families had some prior knowledge about the chance of admission, they might apply to schools with a higher probability of acceptance. Thus comparing winners and losers across lotteries might induce bias. In equation (2), δ gives the weighted average of outcome differences summed over each individual lottery, with weights equal to $N * [p(1-p)]$ where N is the number of applicants and p is the probability of admission (Cullen, Jacob and Levitt 2006). We can test the validity of the randomization by replacing the outcomes A_{ij} in equation (2) with pre-determined covariates such as race, gender and prior test scores. If the randomization was conducted correctly, winners and losers should be balanced on all characteristics that are fixed at the time of the lottery. We test this in Appendix Table A1 and find that only 2 of 20 coefficients are significant at the 10 percent level, which is consistent with sampling variability.

[13] The lotteries were actually conducted at the school-grade-priority group level, so the number of lotteries is greater than the number of schools. We suppress subscripts for grade and priority group for notational convenience. The X_{ij} vector includes controls for race, gender, free or reduced price lunch, 8th grade math and reading test scores, and indicators for the level of math taken in 8th grade (since some students are already enrolled in advanced math).

In Table 3 we show estimates from equation (2) for our key outcomes of interest – 9th grade Math and English test scores, high school graduation, and college attendance and persistence.[14] Columns 1 and 3 show the mean value of each outcome for students who lost the lottery (the "control mean"), and Columns 2 through 4 show the mean effect of winning the lottery on each outcome, with standard errors that are clustered at the individual lottery level. We find no impact of winning the lottery on math or English test scores, although lottery winners are significantly more likely than lottery losers to take an end-of-course (EOC) math exam at all.[15] Imputing the (standardized) 8th grade math score for students who did not take a 9th grade exam results in a point estimate of 0.000 with a standard error of 0.039. There is no difference in English test-taking by lottery status. Thus we can rule out even relatively small (0.05σ) impacts on test scores of winning the lottery.

Overall, we find small positive impacts of winning the lottery on educational attainment. Lottery winners are 3.6 percentage points more likely to graduate from high school, 3.2 percentage points more likely to attend a 2-year college, and 2.4 percentage points more likely to ever attend a 4-year college (only high school graduation is even marginally significant, at the 10 percent level). However, we see a relatively large proportional impact on college quality. Lottery winners are 2.4 percentage points more likely to attend a "very competitive" college. That estimate is a 25% increase from the control mean baseline of 9.6 percentage points, and is statistically significant at the 5 percent level. Although the 0.8 percentage point increase in "most competitive" college attendance is small and statistically insignificant, it is a 33% increase from the 2.4 percentage point baseline among lottery losers.

We have college attendance data from the NSC through the Spring of 2011. This means that rising 9th grade students who progress normally through high school would be able to attend a maximum of 10 semesters (Fall 2006 to Spring 2011) of college. Thus for rising 9th grade students in 2002 our outcome is completion of a degree within 5 years of high school graduation, with additional years available for 10th and 11th graders. We construct two measures of college persistence – 1) enrollment in at least 4 semesters; 2) completion of a degree by the Spring of 2011. In the righthand panel of Table 3 we examine impacts of winning the lottery on these two measures of persistence. The results are similar to the attendance outcomes. The impacts on persistence are proportionally very

---

[14] We only estimate test score impacts for rising 9th graders (N=1,146), but we show results for educational attainment outcomes for the full sample (N=1,865) of rising 9th through 11th grade students.

[15] Lottery winners are only about 2.5 percentage points more likely to remain in a CMS school, so selective attrition can only explain about 25% of the difference in test-taking. The remainder is explained by students who are enrolled but not taking math, or taking math in a non-standard (usually remedial) course such as pre-Algebra.

similar to the impacts on attendance for very competitive colleges, while the impact of winning the lottery on persistence in any 4 year college is near zero and statistically insignificant.

Next we examine the impact of winning the lottery for different groups of students. We estimate equation (2) for the five main outcomes in the paper, and we add interactions between winning the lottery and indicator variables for race (nonwhite/white), free or reduced price lunch, and gender. Table 4 presents those results in a similar format to Table 3, with the addition in Column 5 of an F-test for equality of coefficients across groups. The point estimates for all five outcomes are greater for white students than for nonwhite students, although only the difference for 9^{th} grade English score is significant at even the 10 percent level. The two measures of college attendance are positive and significantly different from zero for free lunch-eligible students. Interestingly, free lunch-eligible students gain on all three measures of educational attainment despite negative impacts on both Math and English test scores. Finally, we note the large and statistically significant difference in impacts on four-year college attendance for girls versus boys. Girls who win the lottery are about 7 percentage points more likely to attend a four year college. In contrast, boys are actually slightly less likely to attend a four year college. This matches the growing body of evidence that girls benefit academically more than boys from educational interventions (Hastings et al 2006; Anderson 2008; Angrist, Lang and Oreopoulos 2009; Deming 2009; Jackson 2010; Lavy, Silva and Weinhardt 2010).

Overall, while there are some interesting differences in outcomes across groups, most do not pass the threshold of statistical significance. Moreover, interpreting the pattern of results across outcomes and groups of students is difficult. The treatment is an offer of admission to a student's first choice school, but it is not necessarily true that the offer (even if taken up) will improve student outcomes. First, families must choose schools that are actually higher quality in the sense that they lead to improvements in the outcomes that we can measure. If preferences for school attributes are heterogeneous, then only families who place a high weight on school quality can expect to benefit from choice (Hastings, Kane and Staiger 2008). Second, the impact of winning the lottery depends on the difference in quality between a student's chosen school and the quality of their outside option. If students who lose the lottery to attend a high quality school are able to find another comparable alternative, the estimated impacts will be small even if the school itself is quite good. One can try to assess the difference in quality for lottery winners by estimating equation (2) with measures of quality and other school characteristics as outcomes. However, the interpretation of δ will also vary with the group of students who apply to each lottery.

In the next section we lay out a simple framework that highlights both the nature of selection into the lottery sample and the expected impact of choice across students and neighborhoods. The model motivates our focus on students who are assigned to low quality neighborhood schools.

5. Model of School Choice

We assume a simple education production function $A = A(b_i, S_j, \gamma_{ij})$ where academic achievement of student i in school j depends on the ability of the student b_i, the quality of the school the student attends S_j, and γ_{ij} is the quality of the match between the student and the school. We interpret S_j broadly as any observable feature of a school that affects all students and would be considered similarly by parents who are selecting neighborhoods in which to live (quality of instruction, school culture, management, etc.) Notably, this includes peers, and measures of peer quality such as average test scores as well as the "value added" measures in Table 2.

Student achievement is also a function of match quality γ_{ij}. We think of γ broadly as anything that is idiosyncratic to a student's relationship to the school, such as a particularly positive or negative peer group influence, a favorite teacher or set of teachers, etc. The key difference between S and γ is that families do not observe match quality until the child actually attends school.[16] Instead, we model match quality as a mean zero i.i.d. shock that is uncorrelated with S. Thus families would prefer to reside in neighborhoods with higher quality schools, yet the potential achievement of a student in each school is not perfectly predicted by S and depends also on whether they receive a good or bad match. Although match quality is unknown initially, once realized it does not vary by school year for the same student, nor does S.[17] We also assume that match quality is independent across schools for an individual student, i.e. $cov(\gamma_{ij}, \gamma_{ik}) = 0$.[18]

[16] If families perfectly observed match quality in all schools, the model would be a standard parametric selection model in the spirit of Borjas (1987). The effect of relaxing the assumption of unobserved match quality depends on what we assume about the correlation between matches across schools and students. In most cases, it would complicate the model but not change our interpretation of the main results. See the discussion in the Results section for details.

[17] We assume constant values of S and γ even for students who are changing school levels (i.e. middle school to high school) but remaining in the same feeder pattern. This is reasonable since peers are a potentially important component of both S and γ. We can relax the assumption of invariance of γ across years by allowing for correlated shocks within the same school/feeder pattern (and we could assume a higher variance shock in the rising grades). In general, allowing for new draws at match quality in different years reduces the relative importance of match quality in the choice decision.

[18] This is less restrictive than it seems, since we can always expand the definition of student ability b_i to include unobserved student characteristics that covary across schools (such as being a "bad seed", or a good study partner).

We do not explicitly model the initial distribution of students across neighborhoods with different levels of S. Epple and Romano (2003) show that, under reasonable assumptions, the equilibrium outcome of families' residential location decisions is hierarchical sorting of neighborhoods schools by income and school quality, and the descriptive statistics in Table 1 bear this out.[19]

Students experience a realized level of academic achievement $A(b_i, S_j, \gamma_{ij})$ in the year prior to choice.[20] We assume a linear and additive specification of A, and $\gamma \sim N(0, \sigma^2)$.

$$A_{ij} = b_i + S_j + \gamma_{ij} \tag{3}$$

Once the school choice policy is announced, students are no longer bound to attend their neighborhood school and can apply to attend any other school in the district. Families are risk neutral in academic achievement and select a school to maximize utility $U_{ij} = E(A_{ij}) - C_{ij}$.[21] They will select a non-neighborhood school if:

$$E(A_{ij}) - C_{ij} > A_{in} \quad \text{for any j} \neq \text{n} \tag{4}$$

C_{ij} is the cost of traveling to a non-neighborhood school for student i. We normalize $C_{in} = 0$ and assume that C is independent of A - that is, travel cost can affect the probability of choice but does not directly affect achievement. Since match quality γ is mean zero by construction, $E(A_{ij}) = b_i + S_j$ and we can rewrite (4) as:

$$b_i + S_j - C_{ij} > b_i + S_n + \gamma_{in}$$
$$\gamma_{in} < (S_j - S_n) - C_{ij} \tag{5}$$

[19] While we could add this sorting process to the model, it adds no additional insight beyond Epple and Romano (2003). The intuition is that when school expenditures are equalized (as they are within a district), the main determinant of school quality is average peer ability. If the demand for school quality is normal, and income and ability are positively correlated, then richer families with more able students will pay more to live in neighborhoods with better peers (and thus better schools). Even if income and ability are uncorrelated, stratified equilibria can still exist if demand for school quality is increasing in student ability.

[20] Epple and Romano (2003) also show that, in the absence of transportation costs, the equilibrium result from open enrollment is to equalize quality across neighborhoods. If the policy were anticipated when families select neighborhoods, they would have no reason to pay higher prices for access to neighborhood schools, absent a capacity constraint. In Charlotte-Mecklenburg, the school choice policy was relatively unanticipated, and it did enforce capacity constraints so that non-neighborhood families were not always guaranteed admission. Here we are interested in the short-run implications of such a policy, before any possible resorting of families in response to the policy.

[21] Implicitly we assume that families have homogeneous preferences but vary in their access to schools. This is a departure from Hastings, Kane and Staiger (2008), who use variation in parental choices to estimate preferences for school quality and other factors. Here, every parent cares about school quality equally.

Define $E(A_{ij}{}^{*}) - C_{ij}{}^{*} = \ b_i + S_j{}^{*} - C_{ij}{}^{*}$ as expected achievement in the choice of school that maximizes the expected utility for each student. Thus for any family, the probability that they will choose a non-neighborhood school is:

$$= \Pr(\gamma_{in} < S_j{}^{*} - S_n - C_{ij}{}^{*})$$

$$= \Pr[\frac{\gamma_{in}}{\sigma_n} < \frac{S_j{}^{*} - S_n - C_{ij}{}^{*}}{\sigma_n}]$$

$$= \Phi\left[\frac{\left(S_j{}^{*} - S_n\right) - C_{ij}{}^{*}}{\sigma_n}\right]$$

where Φ is the CDF of the standard normal distribution. The probability of choice is thus increasing in the quality of the chosen school $S_j{}^{*}$ and decreasing in neighborhood school quality S_n and travel cost C_{ij}. The probability of choice is also decreasing in σ_n, the variance of neighborhood school match quality. Intuitively, this is because "bad matches" can drive families to choose non-neighborhood schools of similar or even lower quality. As this variance shrinks, observable features of choice such as relative quality and travel cost play a larger role. Our measure of travel cost is the crow's flight distance from each student's home address to each school. We normalize distance to the neighborhood school to zero and calculate the additional distance that a student would have to travel to attend each school.

In Table 5 we test whether selection into the lottery sample matches the predictions of the model. We estimate the probability that a student will choose their neighborhood school, controlling for demographic characteristics, prior test scores, and measures of relative distance to higher and lower quality schools. Column 1 looks at the probability that students will apply to a non-neighborhood school across all students in CMS, while Column 2 looks at within-neighborhood selection by controlling for neighborhood school fixed effects. Overall, applicants to non-neighborhood schools are more likely to be minority and low-income, even after controlling for neighborhood school fixed effects. Proximity to higher quality alternatives strongly predicts choice among students within the same neighborhood. Consistent with the predictions of the model, each additional mile of distance to a higher quality alternative leads to a 3 percentage point decrease in the probability of choice. Interestingly, proximity of lower quality alternatives does not affect the probability of choice.

We are also interested in the expected difference in achievement for families who choose a non-neighborhood school.

$$E(A_j |\ choice) - E(A_n | choice)$$

$$= b_i + {S_j}^* - E(b_i + S_n + \gamma_{in} | choice)$$
$$= b_i + {S_j}^* - b_i - S_n - E(\gamma_{in} | \, choice)$$
$$= {S_j}^* - S_n - E\left(\gamma_{in} | \, \frac{\gamma_{in}}{\sigma_n} < \frac{{S_j}^* - S_n - {C_{ij}}^*}{\sigma_n}\right)$$
$$= {S_j}^* - S_n - \sigma_n E\left(\frac{\gamma_{in}}{\sigma_n} | \, \frac{\gamma_{in}}{\sigma_n} < \frac{{S_j}^* - S_n - {C_{ij}}^*}{\sigma_n}\right)$$
$$= {S_j}^* - S_n + \sigma_n \frac{\phi\left(\frac{{S_j}^* - S_n - {C_{ij}}^*}{\sigma_n}\right)}{\Phi\left(\frac{{S_j}^* - S_n - {C_{ij}}^*}{\sigma_n}\right)}$$
$$= {S_j}^* - S_n + \sigma_n \lambda(z) \qquad (6)$$

Where $z = \left(\frac{{S_j}^* - S_n - {C_{ij}}^*}{\sigma_n}\right)$ and $\lambda(z) = \frac{\phi(z)}{\Phi(z)}$ is the inverse Mills ratio, which in this case is the negative of the hazard rate for the expression z (Heckman, 1979). Thus the expected gain in academic achievement for students who choose a non-neighborhood school is increasing in the direct effect of ${S_j}^* - S_n$ but decreasing in the indirect effect of ${S_j}^* - S_n$ that operates through sample selection. This is because as ${S_j}^* - S_n$ increases, choice applicants are drawn from a larger share of the neighborhood school match quality distribution. Conversely, students with zero or even negative values of ${S_j}^* - S_n$ can still benefit academically from receiving a better match.

Because the direct and indirect effects of ${S_j}^* - S_n$ have opposite signs, the relationship between neighborhood school quality and the expected achievement gains from choice is not obvious. We might think that for large enough values of $\lambda(z)$, achievement gains would be greater for students who come from high quality neighborhood schools but have very bad matches.

Yet, all else equal, achievement gains will always be greater for students with lower neighborhood school quality. Furthermore, the additional achievement gain from choice for students with lower values of S_n will be greater when school quality is an important determinant of student achievement. To see this, let $\alpha \in [0,1]$ denote the weight on school quality relative to match quality in the achievement function.[22] Because α is a constant, we get very similar expressions for z and for expected achievement:

$$\tilde{z} = \frac{\alpha({S_j}^* - S_n) - {C_{ij}}^*}{(1-\alpha)\sigma_n}$$

[22] For simplicity we do not consider the weight on student ability. Since it is differenced out of equation (4) we can give it an arbitrary weight and leave the main implications of the model unchanged.

and $E(A_j|\ choice) - E(A_n|choice)$ is simply:

$$= \alpha({S_j}^* - S_n) - (1-\alpha)E(\gamma_{in}|\ choice)$$
$$= \alpha({S_j}^* - S_n) + (1-\alpha)\sigma_n\lambda(\tilde{z}) \qquad (7)$$

The derivative of equation (7) with respect to S_n is:

$$-\alpha[1 + \lambda'_{S_n}(\tilde{z})] \qquad (8)$$

Since $\alpha \in [0,1]$ and $\lambda'(z) < 0$ with $\lim_{z\to\infty} \lambda'(z) = 0$ and $\lim_{z\to-\infty} \lambda'(z) = -1$, equation (8) will always be negative. As neighborhood school quality increases, the expected achievement gain from choice decreases (and vice versa), with a gradient that is proportional to α.

As $\alpha \to 1$, school quality becomes the only important predictor of academic achievement, and $S_j - S_n$ will be the predicted achievement difference for any student, not just the sample of students that applied. As $\alpha \to 0$, the weight on match quality gets very large, until (in the limit) families are selecting simply for a new draw and without regard to quality. Thus the difference in achievement gains from choice for students with low versus high quality neighborhood schools is a partial test of the relative importance of school quality for academic achievement.

It is useful to compare the expected achievement gain from choice in equation (7) with the impact of randomly assigning all students to schools. With random assignment, there would be no self-selection of students with "bad matches" and the expected achievement gain relative to attending the neighborhood school is simply:

$$= E\big(b_i + \alpha S_j + (1-\alpha)\gamma_{ij}\big) - E(b_i + \alpha S_n + (1-\alpha)\gamma_{in})$$
$$= \alpha(S_j - S_n)$$

Notice that since $\lim_{z\to\infty} \lambda(z) = 0$, the expected achievement impact for lottery applicants will approach the observed school quality differential as the expression z increases. While the $\lambda(z)$ term is commonly used as an expression for selection bias in regressions based on observational data (Heckman 1974; Heckman, 1979), here we use it show how the impact of attending a different school that is estimated on a self-selected sample of lottery applicants differs from the impact of attendance on a randomly chosen student. The impact of winning a school choice lottery approaches the average treatment effect from random assignment as $S_j - S_n$ increases, as $\to 1$, and as $\Phi(z)$ approaches 1. This

is similar to an “identification at infinity” result where the model is identified when the support of z is large enough to allow the probability of selection to approach one as z approaches infinity (Chamberlain 1986, Heckman 1990). Intuitively, the lottery treatment effect will approach $\alpha(S_j - S_n)$ when the weight on school quality relative to match quality is high, when students are choosing a much better school, when travel to that school is not much more costly than travel to their neighborhood school, and when the variance of match quality σ_n is low.[23]

The coefficient on δ will more closely approximate $S_j - S_n$ as the selection term $\sigma_n \lambda(z)$ goes to zero, which happens as z approaches infinity and $\phi(z)$approaches one. Thus we can test the implications of the model by estimating equation (9) for sets of students with different values of S_n and z. We estimate:

$$A_{ij} = \delta\left(W_{ij} * S_{in}^{LQ}\right) + \theta\left(W_{ij} * S_{in}^{HQ}\right) + \eta S_{in}^{LQ} + \beta X_{ij} + \Gamma_{\mathrm{j}} + \varepsilon_{\mathrm{ij}} \tag{9}$$

S_{in}^{LQ} and $S_{in}^{HQ}$are indicator variables equal to one if a student is zoned to neighborhood schools of “low” and “high” quality respectively, and the rest of the notation is identical to equation (2). Since equation (9) controls for lottery fixed effects, allowing the impact of winning the lottery to vary with S_n is equivalent to allowing it to vary with $\left(S_j - S_n\right)$. We stratify every school into one of the two categories based on several different metrics, including our value-added estimates of school quality and revealed preference. We can interpret $(\delta - \theta)$ as a differences-in-differences (DID) estimate of the relative impact of coming from a low quality neighborhood school on winning the lottery. Since equation (9) includes lottery fixed effects, we are comparing applicants *to the same school and in the same grade* who came from different neighborhood schools. Furthermore, since the derivative of the expected achievement gains from choice with respect to S_n (equation (8)) is proportional to α, we interpret the magnitude of $(\delta - \theta)$ as a measure of the relative weight on school quality in the achievement production function for each outcome. If there is no variation by neighborhood school quality in the impact of winning the lottery on college attendance, for example, that is evidence that schools (at least variation in the schools within CMS) are not an important contributor to college attendance relative to individual ability b_i or match quality γ_{ij}.

[23] The expected achievement gain conditional on choice is increasing in σ_n. However, when σ_n is large, the relative influence of school quality and travel cost on expected achievement decreases. This is because the selection bias term $\sigma_n \lambda(z)$ increases, but also because the marginal influence of changes in S or C on $\lambda(z)$ becomes very small when σ_n, which is in the denominator of both $\phi(z)$ and $\Phi(z)$, is large.

In addition to stratifying the analysis by neighborhood school quality, we can allow the lottery treatment effect to vary continuously with our estimate of the school quality differential $(\widehat{S_j} - \widehat{S_n})$:

$$A_{ij} = \delta W_{ij} + \eta(\widehat{S_j} - \widehat{S_n}) + \theta[W_{ij} * (\widehat{S_j} - \widehat{S_n})] + \beta X_{ij} + \Gamma_{\mathrm{j}} + \varepsilon_{\mathrm{ij}} \tag{10}$$

Intuitively, if the estimated school quality difference is a good approximation of the impact that each school has on its students, then the impact of winning the lottery should be larger for applicants with a larger estimated quality difference. If all students were randomly assigned to schools, then the impact of switching a student from school j to school n would be $S_j - S_n$. In that case the specification in equation (10) would yield a coefficient of zero on the main effect and one on the interaction term. On the other hand, if the quality differential does not do a good job of explaining the lottery treatment effect, the slope coefficient θ will be close to zero and δ in equations (2) and (10) will be similar. We also estimate similar specifications where winning the lottery is interacted with $\Phi(z)$, the probability that students choose their neighborhood school. We could think of this as an alternative “revealed preference” measure of school quality that is robust to misspecification of S.

There are two obstacles in using equation (10) as a test of the model. The first is that our estimates of S might be biased by unobserved sorting of students into schools. The second obstacle is that while we can make predictions about the relative magnitude of the selection bias term for different types of students, we do not know α, the relative importance of S and γ in the achievement function. Empirically, we cannot separate these two issues without further assumptions.

By assumption, travel cost C_{ij} affects the probability of choice $\Phi(z)$ but does not enter the achievement function directly. This suggests a partial test of the importance of match quality. To see this, note that the derivative of equation (7) with respect to C_{ij} is simply $-(1-\alpha)\lambda'_{C_{ij}}(\tilde{z})$, which is positive everywhere. This implies that, all else equal, students who are farther away from their chosen school will benefit more from attending it. Moreover, this increased benefit operates entirely through self-selection. As a result, the marginal impact of increased travel cost on academic achievement is larger when α is closer to zero. Thus we estimate a version of equation (10) where winning the lottery is interacted with the individual student’s difference in distance between the chosen school and their neighborhood school, which is our measure of C_{ij}. The coefficient on that interaction will be larger when match quality is more important.

6. Results

We split the analysis sample into two groups according to the attributes of their neighborhood school. The four lowest-ranked schools on the measure of average quality in Column 7 of Table 2 are Waddell, West Mecklenburg, Olympic and Vance. The four lowest-ranked schools on the revealed preference measure of quality in Column 8 are Garinger, West Charlotte, West Mecklenburg and Olympic. We call the six schools that rank in the bottom four on either measure "low quality" neighborhood schools, and we call the other eight neighborhood schools "high quality". In Appendix Tables 2 and 3 we report results for the four schools that are in each category separately. Overall, the pattern of results is not sensitive to small changes in the choice of schools.

Table 6 presents the results from estimation of equation (9) for the main outcomes of the paper. The format is similar to Table 4, with control means in Columns 1 and 3, point estimates and standard errors in Columns 2 and 4, and an F-test for equality of coefficients across categories in Column 5. The pattern of results is striking. Nearly all of the impacts are larger in the low quality neighborhood school subsample, and the difference between the samples is statistically significant for about two-thirds of all outcomes and 4 of the 5 main outcomes in the paper. The impact of winning the lottery on 9th grade test scores is near zero for students from low quality neighborhood schools, but the impact on math is negative and statistically significant in the high quality sample. However, if we account for the fact that students who win the lottery are more likely to take the test by imputing 8th grade math scores for 9th grade applicants with missing exams, we find a small and marginally significant positive impact on test scores in the low quality sample.

The main finding of this paper is that lottery winners from low quality neighborhood school zones are 8.7 percentage points more likely to graduate from high school, 6 percentage points more likely to attend a four year college, and 4.6 percentage points more likely to earn a four year college degree. This is consistent with our model, which predicts that both the probability of choice and the achievement gains from choice will be increasing in the school quality differential that we estimated in Table 4. Moreover, the large and statistically significant difference in outcomes between the two samples suggests that the benefits of school choice can be large, but only for students who are admitted to a school of higher quality than their neighborhood school. Unlike the results for high school graduation and college attendance, lottery winners from high quality neighborhood schools are actually more likely to attend a "very competitive" college, although we cannot reject that the two coefficients are equal. Furthermore, the gap between the two samples for very competitive colleges narrows for the

persistence measure and actually reverses for degree completion.[24] The impact on four-year college attendance and degree completion for lottery winner s from low-quality neighborhood school zones is proportionally fairly large (about 20 percent and 35 percent of the respective control means). However, for "most selective" colleges, the impacts are an order of magnitude larger. While the point estimates for attendance, persistence and completion are small (between 1.1 and 1.4 percentage points), they represent a doubling of the probability that a student will attend and obtain a degree from one of these elite institutions.

While we have limited evidence on the things that schools did to increase the educational attainment of their students, we find that lottery winners from low quality neighborhood schools had better performance in high school on a range of mediating outcomes. We investigate these outcomes in Table 7. When looking at course-taking, we restrict our attention to math since the courses were more likely to have a standard curriculum, and because of recent evidence that mathematics courses may increase earnings later in life (Goodman, 2009). Lottery winners from low quality neighborhood schools had significantly higher grade point averages overall and in math and science courses, and they took more math courses overall. They also had significantly fewer absences from school in the first school year after the lottery. We find some evidence of reductions in absences in later years, but the estimates are imprecise. In contrast, we found no evidence of any impacts on mediating outcomes in the high quality sample, and point estimates were often negative. Notably, the differences across the two samples on the mediating outcomes are all very large and statistically significant. Such large differences are evidence of the relative importance of school quality, which makes sense because schools have much more control over outcomes such as course-taking and academic preparation than over the college selection and attendance process. [25]

To further understand the pattern of results, we estimate the impact of winning the lottery on enrollment and school characteristics in Table 8. The first row examines selective attrition from the lottery sample. We find that lottery winners in both samples are about 2.5 percentage points more likely to still be enrolled in a CMS school in the Fall of 2002. Neither coefficient is statistically significant. While

[24] The pattern of estimates implies that, among students who live in high quality neighborhood school zones, degree completion conditional on enrollment at a selective college is much lower for lottery winners than for lottery losers. One possibility is that the first choice schools do a better job of placing students in colleges but an equal or worse job of preparing them for college-level work (relative to the neighborhood school). This is consistent with the zero or small negative impacts on grades and course-taking in the bottom panel of Table 6.

[25] It is important to note that the differences in impacts by neighborhood school quality are not simply a proxy for heterogeneous effects by student characteristics. We estimated versions of Table 6 for the student subgroups in Table 4 and found that, even within groups, neighborhood school quality strongly predicted the benefits from choice.

the impact on selective attrition is small, it could bias our results. However, we find that the main results are relatively robust to a variety of assumptions about attrition.[26] In the second row we see that the first stage impact of winning the lottery on fall enrollment is strong in both samples, suggesting that the offer of admission was relatively attractive. As indicated by the control means, however, a significant share of lottery losers still managed to enroll in their chosen school. This could happen because the student moved into the neighborhood over the summer, or because they were admitted subsequently off of a waitlist that we do not observe. Among lottery losers, students who subsequently enroll in their chosen school ("always takers" in the terminology of Angrist, Imbens and Rubin (1996)) have higher family incomes, are more likely to be white, and have substantially higher average test scores and rates of high school graduation and college attendance. This holds even with the low quality neighborhood school sample. Lottery winners with low quality neighborhood schools are about 36 percentage points more likely to attend a magnet school, compared to about 19 percentage points in the "high quality" sample. This helps to explain why there is little change in the demographic composition of lottery winners' schools, in either group.

Finally, we examine the impact of winning the lottery on school quality. In the last five rows of Table 8, the outcome is the "value-added" school quality estimate from Table 2 for each of the five main outcomes in the paper. If these measures were unbiased estimates of school quality, they would represent the average impact of the school on student outcomes. In that case, we would expect the impact for a randomly chosen (not self-selected) student to be simply equal to the coefficient on each outcome. The quality estimates predict that lottery winners in the low quality sample will be more likely to graduate from high school and attend a four year college, but will score no higher on math and English tests. They also predict that in the "high quality" sample, lottery winners will have modestly lower test scores and show no gains on measures of educational attainment. With the exception of the very competitive college attendance outcome, the school quality estimates do a good job of predicting the lottery treatment effects in Table 6.

We formalize the comparison between school value-added and the lottery treatment effects by allowing for a continuous relationship between the estimated school quality difference and the impact of a randomly assigned offer of admission. We estimate equation (10), where the impact of winning the

---

[26] We reestimate the models in Table 5 assuming values of zero for students who were missing in the Fall of 2002. The impact on high school graduation drops by about 35 percent but is still statistically significant in the "low quality" sample. This assumption is probably too restrictive, especially since we possess information on the college-going of students who leave prior to the Fall of 2002 and we still find impacts on four-year college attendance.

lottery is interacted with $\widehat{S_j} - \widehat{S_n}$, the difference between the value-added of the school to which a student applies and their neighborhood school. If all students were randomly assigned to schools, then the impact of switching a student from school j to school n would be $\alpha(S_j - S_n)$. In that case the specification in equation (10) would yield a coefficient of zero on the main effect δ, and one on the interaction term θ. A rejection of this "random assignment" joint hypothesis can be interpreted in two ways: 1) that the match quality term $\sigma_n\lambda(z)$ plays an important role in lottery winners' achievement gains, and 2) that $\widehat{S_j} - \widehat{S_n}$ is misspecified.

One potential complication is that students who lose the lottery to attend their first choice often attended another school that was not oversubscribed (as explained in Section 2, virtually no students were admitted to oversubscribed schools unless they listed that school as their first choice). Since we are interested in the school quality lottery winners would have experienced had they lost, we construct an estimate of counterfactual school quality by using applicants' second and third choices. If their second choice was not oversubscribed (or if it was their neighborhood school), then we use the quality difference between their first and second choice in equation (9). Similarly, if their second choice was oversubscribed, and no one was admitted, we move to their third choice, and then to their neighborhood school (the implicit fourth choice for all students). In the small number of cases where there was a lottery in the 2nd or 3rd rounds, we use a weighted average quality difference where the weights are the probability of admission.[27]

In Table 9 we estimate versions of equation (10) that differ only in the covariates used to adjust the estimates of $\widehat{S_n}$ that we produced in equation (1) and Table 2. Column 1 contains the baseline model without any interaction term and is identical to the results in Table 3. Column 2 interacts an indicator for winning the lottery with unadjusted school-level means for each outcome, which are just demeaned versions of the descriptive statistics in Table 1. The coefficients on the interactions are significantly different from zero only for the 9th grade math and high school graduation outcomes, and the "random assignment" hypothesis is strongly rejected for all five outcomes.

Column 3 shows results where winning the lottery is interacted with the school quality estimates in Table 2, which include demographic covariates and polynomials in 8th grade math and reading test scores. The performance of the model improves dramatically, and we fail to reject the

---

[27] Formally, we estimate the school quality differential as VA(W)-VA(L) where VA(L) = p2*(VA2) + (1-p2)*[(p3*VA2) + (1-p3)*VA(n)]. VA(W) and VA(L) are value-added if the students wins and loses the lottery respectively, p2 and p3 are the probability of admission to the students' 2nd and 3rd choices, and VA2, VA3 and VA(n) are value-added for the 2nd choice, 3rd choice and neighborhood schools respectively. Estimates using just the neighborhood school as the counterfactual are very similar.

"random assignment" hypothesis for 9th grade math, high school graduation and 4 year college attendance. The school quality estimates do a particularly good job of predicting gains in high school graduation and four year college attendance. The coefficients on the main effects are 0.015 and 0.005 respectively, and the interaction terms are 0.937 and 0.866. This says that students who win the lottery to attend a school with identical estimated quality to the school they would have attended if they lost are 1.5 percentage points more likely to graduate and 0.5 percentage points more likely to attend a 4 year college (so some self-selection on match quality or misspecification or remains). However, for every 1 percentage point of value-added gained, lottery winners are 0.94 and 0.87 percentage points more likely to graduate from high school and attend a four-year college. The fact that both are significantly different from zero is a rejection of the null hypothesis that value-added estimates have no power to predict the treatment effect of winning the lottery.

School value-added does a very poor job of predicting 9th grade English scores and very competitive college attendance. This could be because match quality is particularly important for these outcomes, or because we cannot adequately control for their determinants, and so $\widehat{S_n}$ is misspecified. Column 4 makes use of a richer model of value-added that controls for 3rd through 8th grade test scores plus absences, out-of-school suspensions and grade repetition in middle school. Perhaps surprisingly, these additional covariates have very little impact, suggesting that a basic set of controls adequately controls for selection for three of the five main outcomes.

Column 5 presents estimates from a model where winning the lottery is interacted with $\Phi(z)$, the share of students assigned to a neighborhood school that apply to attend another school. Here we measure school quality by relying on applicants' revealed preferences for schools other than their own. The pattern of results closely mirrors Columns 3 and 4, suggesting that our estimates of school quality conform closely to parental preferences for schools.

Finally, in Column 6 we interact an indicator for winning the lottery with the additional distance in miles that an applicant must travel (relative to their neighborhood school) to attend their preferred school. In the model, students can gain academically through self-selection out of a "bad match". This generates the prediction that students who face a higher travel cost will benefit more from choice. Since the marginal academic achievement gain from increased travel cost is larger when match quality is relatively more important, we interpret the results in Column 6 as a test of the relative importance of match quality. The interaction term on relative distance is very small and statistically indistinguishable from zero for all five outcomes. Yet we do not wish to push this result too strongly. Our assumption that distance does not enter the achievement function directly may not hold. Also, we are limited in the

outcomes we can observe. A family that chooses a very distant school may benefit in ways that we do not observe. More generally, our model will not do a good job of capturing any determinants of choice that increases family utility but not student achievement.

Our empirical implementation of the model relies on comparisons between students who applied to the same school but who vary in the quality of their assigned neighborhood school. We interpret $(\delta - \theta)$, the differences-in-differences (DID) estimate from equation (9), as the relative impact of school quality on each outcome, rather than as an allocative efficiency gain that comes from an improved match between student and school. We assume in the model that students only observe their match quality in the neighborhood school, but not in their first choice school. Thus the main threat to our interpretation of the results is if match quality in the choice school γ_{ij} were 1) known in advance by applicants; AND 2) negatively correlated with S_n. This would be true if students from low quality neighborhood schools had much better idiosyncratic matches than students from high quality neighborhood schools, *within each lottery.* This would cause us to overstate the relative importance of school quality for these students. Since about 43 percent of first choices were one of the three magnet schools located in the inner city, we might be concerned that applicants from low quality neighborhood schools had particularly good potential matches in these schools and programs. However, nearly all the main results in the paper hold up when we exclude magnet applications from the analysis.[28] A related concern is that 10th and 11th grade applicants might be more likely to sort on idiosyncratic matches, yet we find that the pattern of results holds when we restrict everything to rising 9th grade applicants.

Throughout the paper we use the intent-to-treat (ITT) estimate of winning the lottery rather than instrumenting for enrollment and estimating a local average treatment effect (LATE) for applicants who comply with their lottery status. If we used LATE rather than ITT in Table 9, all of the estimates would be scaled upward by the inverse probability of enrollment, which is between 0.5 and 0.6 in Table 7. This would push many of the interaction coefficients in Table 8 above 1, implying that the school quality estimates under-predict the impact of winning the lottery. This makes sense, given that students who do not comply with their lottery status may have match-specific reasons for doing so. Concretely, the assumption that students do not observe match quality in the choice school is less likely to hold by the 20th day of the Fall 2002 semester (when yearly enrollment is officially recorded). We would expect

[28] Results are available upon request. The only difference is in selective college enrollment and degree completion, where excluding the magnets makes the estimates in the low quality sample insignificant but still significantly larger than for the HQ sample in many cases.

some students who learn of a bad match to attend another school, which would drive the lottery-based estimates for compliers above the average effect for all students.

Finally, it is useful to think also about the welfare gains from a school choice policy. Leaving aside compositional changes in S from increased population or peer effects, moving students to higher quality schools clearly increases overall achievement. Consider instead a policy that allows for school choice but keeps each school's capacity fixed, so that spots only open in schools when students from the neighborhood leave voluntarily.[29] In that case, absent non-linear peer effects, all of the net achievement gain from choice comes through the match quality term. Since $\alpha\sigma_n\lambda(\tilde{z}) > 0$ for all values of $\tilde{z}$, the net achievement gain in the population will be increasing in α and σ_n.[30] Thus our inability to find evidence of gains from match quality may suggest that there are no allocative efficiency gains from school choice. However, our research design does not allow us to say anything about the possible productive efficiency gains from competition between schools for students (Hoxby, 2000; Chan and McMillan 2009; Figlio and Hart 2010).

7. Discussion and Conclusion

In this paper we study the impact of winning an admissions lottery to attend a public high school in Charlotte-Mecklenburg. We find that lottery winners who are assigned to low-quality neighborhood schools are more likely to graduate from high school, attend a four-year college, and earn a bachelor's degree. They are about twice as likely to earn a degree from an elite institution such as UNC-Chapel Hill or Duke. In contrast, we find no evidence of benefits for lottery winners from higher-quality neighborhood schools. Our model predicts that achievement gains from school choice will be greater for students with lower-quality neighborhood schools, and that the difference will be particularly large when school quality is a more important component of student achievement. This framework can be applied to understand the different impacts of school choice policies by design and setting. The quality of a student's default option matters just as much as the quality of the school they choose.

The results show that school choice can have an important impact on educational attainment, but only when students can gain access to schools that are better on observed dimensions of school quality. Furthermore, we find that value-added measures of school quality which control for a prior test

---

[29] CMS increased capacity at schools where they anticipated high demand. The number of extra slots was decided prior to the lottery but was unknown to students at the time of application (Hastings, Kane and Staiger 2006).

[30] As $\alpha \to 0$ and $\sigma_n \to 0$, student achievement is dominated by observed quality differences and very few students would exit the highest quality schools, leaving fewer spots available to applicants.

score and a basic set of demographic covariates do a reasonable job of predicting the treatment effects from lottery-based random assignment.

Unfortunately, we cannot say much about the underlying explanation for the gains experienced by lottery winners from low quality neighborhood schools. Because the choice schools were often magnet schools, with specialized programs such as career academies, arts education, and intensive college prep, the benefits could come primarily from improved student engagement in high school. It is possible that having demographically similar but more able peers led to increased student learning and engagement inside the classroom. Better peers could also have an impact on behavior inside and outside of the classroom. The impacts on mediating variables as early as 9th grade are inconsistent with a scenario where "better" schools are simply more organized at getting students to surmount administrative hurdles such as accumulating enough credits in the right high school courses, or filling out college applications and financial aid forms. That could generate increases in educational attainment, but would be unlikely to raise grade point averages, increase math course-taking and reduce absences.

Interestingly, the performance of value-added measures of school quality is best for high school graduation and four year college attendance, middling for math, and very poor for competitive college attendance and English test scores. Provided that these results could be replicated in other settings, they have important implications for the design of school accountability policies. It makes little sense to hold schools accountable for outcomes that they cannot control. School quality estimates that approximate impacts from random assignment are less likely to be biased by unobserved determinants of achievement, and thus may serve as better candidates for judging performance.

References

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

Abdulkadiroğlu, Atila , Susan M. Dynarski, Thomas J. Kane and Parag A. Pathak (2010). "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots". *Quarterly Journal of Economics,* forthcoming.

Altonji, Joeseph G., Todd E. Elder and Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113 (2005):151 184.

Anderson, Michael L. (2008). "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool and Early Training Projects". *Journal of the American Statistical Association,* 103(484): 1481-1495.

Andrabi , Tahir, Jishnu Das and Asim Ijaz Khwaja (2009). “Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets”. Unpublished Working Paper.

Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Jon B. Fullerton, Thomas J. Kane, Parag A. Pathak and Christopher R. Walters (2011). “Student Achievement in Massachusetts Charter Schools”. Working paper.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin (1996). “Identification of Causal Effects Using Instrumental Variables”. *Journal of the American Statistical Association*, Vol. 91, No. 434 (Jun., 1996), pp. 444-455.

Angrist, Joshua D., Daniel Lang and Philip Oreopoulos (2009). “Incentives and Services for College Achievement: Evidence from a Randomized Trial”, *American Economic Journal: Applied Economics,* 1(1): 136-163.

Booker, Kevin, Tim R. Sass, Brian Gill and Ron Zimmer (2011). “The Effect of Charter High Schools on Educational Attainment”, *Journal of Labor Economics,* 29(2): 377-415.

Chamberlain, Gary (1986). “Asymptotic Efficiency in Semi-Parametric Models with Censoring”, *Journal of Econometrics,* 32(2): 189-218.

Chan, P.C. Winnie and Robert McMillan (2009). “School Choice and Public School Performance: Evidence from Ontario’s Tuition Tax Credit”. Unpublished Working Paper.

Cullen, Julie Berry, Brian A. Jacob, and Steven Levitt. (2006). “The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries.” *Econometrica*, 74(5):1191-1230.

Cullen, Julie Berry, Brian Jacob and Steven Levitt. (2005). “The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools,” *Journal of Public Economics*, 89(5-6): 729-760.

Cullen, Julie Berry and Brian Jacob (2007). “Is Gaining Access to Selective Elementary Schools Gaining Ground? Evidence from Randomized Lotteries”, National Bureau of Economic Research Working Paper No. 13443.

Deming, David (2009). “Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start”. *American Economic Journal – Applied Economics,* 1(3): 111-134.

Deming, David (2011). “Better Schools, Less Crime?” *Quarterly Journal of Economics,* forthcoming.

Dobbie, Will and Roland G. Fryer Jr. (2011). “Living in the Zone: An Analysis of a Bold Social Experiment in Harlem”. *American Economic Journal: Applied Economics,* forthcoming.

Epple, Dennis N. and Richard Romano (2003). “Neighborhood Schools, Choice, and the Distribution of Educational Benefits”. In *The Economics of School Choice,* Caroline M. Hoxby (ed.), p. 227-286, University of Chicago Press.

Evans, William N. and Robert M. Schwab, “Finishing High School and Starting College: Do Catholic Schools Make a Difference?,” *Quarterly Journal of Economics*, 110 (1995): 941-974.

Figlio, David N. and Cassandra M.D. Hart (2010). “Competitive Effects of Means-Tested School Vouchers”. NBER Working Paper #16056.

Goodman, Joshua (2009). “The Labor of Division: Returns to Compulsory Math Coursework”. Unpublished Working Paper.

Hastings, Justine S., and Jeffrey M. Weinstein. (2008). “Information, School Choice, and Academic Achievement: Evidence from Two Experiments.” *Quarterly Journal of Economics*.

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(2):232-236.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. (2008). “Heterogeneous Preferences and the Efficacy of Public School Choice”. Working Paper.

Hastings, Justine S. , Christopher Neilson and Seth Zimmerman. (2011). “The Effect of School Choice on Intrinsic Motivation and Academic Outcomes.” NBER Working Paper 2011.

Heckman, James J. (1974). “Shadow Prices, Market Wages and Labor Supply”. *Econometrica,* 42(4): pp. 679-694.

Heckman, James J. (1979). “Sample Selection Bias as a Specification Error”. *Econometrica,* 47(1): pp. 153 161.

Heckman, James J. (1990). “Varieties of Selection Bias”. *American Economic Review,* 80(2): pp. 313-318.

Howell, William G. and Paul Peterson (2002). *The Education Gap: Vouchers and Urban Schools,* Washington, DC: Brookings Institution Press.

Hoxby, Caroline M. (2000). “Does Competition among Public Schools Benefit Students and Taxpayers?” *American Economic Review,* 90(5): pp. 1209-1238.

Hoxby, Caroline M. (2003). “School choice and school productivity: could school choice be a tide that lifts all boats?” *The Economics of School Choice*, ed. CM Hoxby, 287-341. Chicago: The University of Chicago Press.

Hoxby Caroline M. and Sonali Murarka (2009). “Charter Schools in New York City: Who Enrolls and How They Affect Their Students’ Achievement”. NBER Working Paper No. 14852.

Hoxby Caroline M., and Jonah Rockoff (2004). “The impact of charter schools on student achievement.” Working Paper, Columbia University.

Jackson, C. Kirabo (2009). “Student Demographics, Teacher Sorting and Teacher Quality: Evidence from the End of School Desegregation”. *Journal Labor Economics,* forthcoming.

Jackson, C. Kirabo (2010). “Peer Quality or Input Quality? Evidence from Trinidad and Tobago”. NBER Working Paper No. w16598.

Kane, Thomas J. and Douglas O. Staiger (2008). “Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation.” NBER Working Paper #14607.

Krueger, Alan B. and Pei Zhu (2004). “Another Look at the New York City Voucher Experiment.” *American Behavioral Scientist.* 47(5): pp. 658-698.

Lavy, Victor (2010). “Effects of Free Choice Among Public Schools”. *Review of Economic Studies,* 77(3): 1164-1991.

Lavy, Victor, Olmo Silva and Felix Weinhardt (2010). “The Good, the Bad and the Average: Evidence on the Scale and Nature of Ability Peer Effects in Schools”. NBER Working Paper No. w15600.

Neal, Derek, “The Effects of Catholic Secondary Schooling on Educational Achievement”, *Journal of Labor Economics*, 15 (1997):98-123.

Pathak, Parag A., and Tayfun Sönmez. (forthcoming). “Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism.” American Economic Review.

Rothstein, Jesse (2006). “Good principals or good peers? Parental valuations of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions.” *American Economic Review,* 96(4):1333-50.

Raudenbush, Stephen W. and Anthony S. Bryk (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Russell Sage: New York, NY.

Raudenbush, Stephen W. and J. Douglas Willms (1995). “The Estimation of School Effects”. *Journal of Educational and Behavioral Statistics,* 20(4): 307-335.

Rothstein, Jesse (2009). “Student sorting and bias in value added estimation: Selection on observables and unobservables”. National Bureau of Economic Research Working Paper No. 14666.

Rothstein, Jesse (2010). “Teacher Quality in Educational Production: Tracking, Decay and Student Achievement.” *Quarterly Journal of Economics,* 125(1): 175-214.

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

Witte, John F. (1997). “Achievement effects of the Milwaukee voucher program.” Working Paper, University of Wisconsin at Madison.

Wolf, Patrick et al (2009). “Evaluation of the DC Opportunity Scholarship Program: impacts after two years.” U.S. Department of Education, Institute of Education Science (NCEE 2000-4023).

Table 1 - Descriptive Statistics by Neighborhood School

							College Attendance			
	Median HH Income	Nonwhite	Free Lunch	9th Grade Math	9th Grade English	HS Graduate	Any 4 Year	Very Competitive	Most Competitive	N
Neighborhood School	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Providence	94,799	0.06	0.03	0.41	0.79	0.82	0.75	0.49	0.15	1,364
South Mecklenburg	72,830	0.15	0.16	0.24	0.58	0.79	0.66	0.39	0.10	1,590
North Mecklenburg	68,275	0.31	0.24	-0.26	0.25	0.76	0.58	0.27	0.09	1,303
Hopewell	67,473	0.31	0.27	-0.31	0.19	0.73	0.54	0.24	0.07	1,410
Butler	66,843	0.13	0.16	0.07	0.58	0.76	0.57	0.27	0.08	1,272
Myers Park	61,782	0.32	0.33	0.02	0.36	0.74	0.60	0.38	0.12	1,422
Olympic	50,914	0.52	0.50	-0.40	-0.15	0.69	0.41	0.12	0.03	1,148
East Mecklenburg	50,806	0.45	0.47	-0.20	0.15	0.64	0.44	0.19	0.06	1,376
Vance	50,394	0.66	0.54	-0.51	-0.08	0.69	0.44	0.12	0.05	1,979
Independence	48,664	0.58	0.53	-0.51	0.04	0.69	0.41	0.10	0.03	1,804
Waddell	45,536	0.58	0.61	-0.53	-0.22	0.63	0.31	0.08	0.02	842
Garinger	38,359	0.73	0.71	-0.60	-0.25	0.64	0.35	0.07	0.03	1,331
West Mecklenburg	38,305	0.70	0.71	-0.53	-0.43	0.60	0.29	0.06	0.02	1,923
West Charlotte	27,278	0.91	0.92	-0.83	-0.67	0.52	0.22	0.02	0.01	1,257
Magnet Schools										
Harding University	43,643	0.68	0.55	-0.28	0.15	0.75	0.53	0.14	0.07	1,016
Berry Academy	41,568	0.79	0.76	-0.61	-0.42	0.68	0.41	0.07	0.02	691
Northwest Arts	52,654	0.39	0.38	-0.69	0.15	0.72	0.48	0.17	0.06	454
Correlation with Income		-0.94	-0.96	0.87	0.91	0.86	0.92	0.93	0.90	

Notes: All data are based on the students that are assigned to each neighborhood school in the Fall of 2002, not the students that actually attend. Column 1 is the average value of median household income for families in each neighborhood school zone, based on the 2000 Census and calculated at the tract level. Columns 4 and 5 show students' average 9th grade math and English end-of-course (EOC) exams. College competitiveness measures are defined by the Barron's Rankings. See the text of the paper for details.

Table 2: Neighborhood School Quality Estimates

		School Quality Measures (Value-Added)						
	Median HH Income	9th Grade Math	9th Grade English	HS Graduate	4 Year College	Very Competitive	Avg. Quality (standardized)	Chose Another School - Φ(z)
Neighborhood School	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
South Mecklenburg	72,799	0.112	0.057	0.053	0.042	0.050	1.388	0.30
North Mecklenburg	68,319	-0.008	0.032	0.047	0.056	0.037	0.967	0.49
Myers Park	61,891	0.027	-0.020	0.011	0.013	0.055	0.532	0.41
Providence	94,801	0.004	-0.030	0.002	0.032	0.049	0.465	0.08
Garinger	38,359	0.059	-0.011	0.020	0.006	-0.004	0.232	0.82
Independence	48,669	-0.039	0.082	0.005	-0.005	-0.028	0.110	0.41
West Charlotte	27,267	0.075	-0.097	-0.024	0.015	0.023	0.061	0.78
East Mecklenburg	50,838	0.078	0.039	-0.045	-0.017	-0.011	0.019	0.55
Hopewell	67,493	-0.085	-0.039	0.004	-0.013	-0.031	-0.466	0.31
Butler	66,871	-0.003	0.029	0.003	-0.043	-0.071	-0.480	0.43
Vance	50,487	-0.111	-0.034	-0.041	-0.026	-0.014	-0.589	0.45
Olympic	50,942	-0.044	-0.096	-0.002	-0.046	-0.046	-0.845	0.58
West Mecklenburg	38,320	0.046	-0.117	-0.046	-0.052	-0.034	-0.898	0.72
Waddell	45,537	-0.146	-0.092	-0.066	-0.049	-0.025	-1.386	0.47
Magnet Schools								
Harding University	43,643	0.054	0.039	0.016	0.053	-0.005	0.639	
Berry Academy	41,568	-0.091	-0.114	0.006	0.007	-0.013	-0.567	
Northwest Arts	52,654	-0.055	0.159	0.004	0.029	0.004	0.523	
Correlation with Income		-0.01	0.30	0.43	0.28	0.36	0.39	-0.90

Notes: Column 1 is the average value of median household income for families in each neighborhood school zone, based on the 2000 Census and calculated at the tract level. Columns 2 through 6 contain value-added measures of school quality for the five main outcomes in the paper. See equation (1) in the paper and the accompanying discussion for details. The table is sorted by Column 7, which is the normalized average of the five quality measures in columns 2 through 6. Column 8 is a "revealed preference" measure of school quality - the probability that students assigned to each neighborhood school will choose to attend it.

Table 3 - Impact of Winning the Lottery on Test Scores and Educational Attainment

	Ctrl Mean			Ctrl Mean	
High School Outcomes	(1)	(2)	**Persistence (4+ sems)**	(3)	(4)
9th Grade Math Score	-0.441	-0.025	Any 2 year college	0.197	0.018
		[0.051]			[0.029]
9th Grade English Score	-0.152	-0.017	Any 4 year college	0.283	-0.005
		[0.044]			[0.026]
Taken 9th Grade Math Exam	0.815	0.089***	Very Competitive	0.080	0.022*
		[0.022]			[0.012]
Grad from CMS High School	0.601	0.036*	Most Competitive	0.021	0.005
		[0.020]			[0.006]
College Attendance			**Earned a Degree**		
Any 2 year college	0.530	0.032	Any 2 year college	0.026	0.011
		[0.031]			[0.013]
Any 4 year college	0.367	0.024	Any 4 year college	0.177	0.005
		[0.029]			[0.015]
Very Competitive	0.096	0.024**	Very Competitive	0.065	0.014*
		[0.011]			[0.017]
Most Competitive	0.024	0.008	Most Competitive	0.020	0.002
		[0.007]			[0.007]
Sample Size	1865				

Notes : Columns 1 and 3 are mean values of the outcome in each row for students who lost the lottery. Columns 2 and 4 are estimates of the effect of winning the lottery from equation (2) in the paper, with lottery fixed effects and basic covariates for balance. Standard errors are below each estimate in brackets and are clustered at the lottery level. "Very Competitive" and "Most Competitive" colleges are defined by the Barron's rankings - see the text of the paper for details

Table 4 - Impact of Winning the Lottery by Subgroups

	Nonwhite		White		F(equal)
	Ctrl Mean		Ctrl Mean		
	(1)	(2)	(3)	(4)	(5)
9th Grade English Score	-0.389	-0.063*	0.224	0.090	0.078
		[0.035]		[0.082]	
9th Grade Math Score	-0.657	-0.043	-0.133	0.014	0.708
		[0.055]		[0.126]	
Grad from CMS High School	0.597	0.022	0.607	0.063*	0.389
		[0.028]		[0.033]	
Attended a 4 Year College	0.333	0.013	0.411	0.044	0.311
		[0.032]		[0.037]	
Very Competitive	0.035	0.019	0.175	0.034	0.714
		[0.017]		[0.029]	
	Free Lunch		Not Free Lunch		
9th Grade English Score	-0.378	-0.084	0.218	0.121*	0.007
		[0.050]		[0.064]	
9th Grade Math Score	-0.655	-0.055	-0.124	0.032	0.255
		[0.038]		[0.087]	
Grad from CMS High School	0.537	0.055	0.683	0.003	0.407
		[0.032]		[0.043]	
Attended a 4 Year College	0.258	0.042**	0.506	-0.009	0.193
		[0.020]		[0.053]	
Very Competitive	0.028	0.028**	0.182	0.017	0.637
		[0.010]		[0.021]	
	Male		Female		
9th Grade English Score	-0.227	-0.036	-0.073	0.009	0.586
		[0.052]		[0.065]	
9th Grade Math Score	-0.465	-0.036	-0.415	-0.010	0.595
		[0.053]		[0.059]	
Grad from CMS High School	0.541	0.030	0.669	0.044	0.763
		[0.024]		[0.036]	
Attended a 4 Year College	0.342	-0.014	0.395	0.071***	0.008
		[0.035]		[0.025]	
Very Competitive	0.078	0.021	0.116	0.028*	0.804
		[0.018]		[0.016]	
Sample Size	1865				

Notes: Columns 1 and 3 are mean values of the outcome in each row for students who lost the lottery. Columns 2 and 4 are estimates from equation (2) in the paper, with a full set of indicator variables for lottery status interacted with the groups indicated in the Table. Standard errors are below each estimate in brackets and are clustered at the lottery level. Column 5 contains the p-value for an F-test that columns (2) and (4) are equal. "Very Competitive" College is defined by the Barron's rankings - see the text of the paper for a list of NC colleges that apply.

Table 5 - Selection into the Lottery Sample

	Pr(Choice)	Pr(Choice)
	(1)	(2)
Relative Distance - Higher Quality School	-0.004	-0.031***
	[0.001]	[0.001]
Relative Distance - Lower Quality School	-0.007***	-0.001
	[0.001]	[0.001]
Distance to Neighborhood School	0.017***	0.015***
	[0.002]	[0.002]
Male	-0.010	-0.003
	[0.007]	[0.006]
African-American	0.088***	0.072***
	[0.010]	[0.009]
Latino	0.047**	0.029
	[0.021]	[0.019]
Free/Reduced Lunch	0.084***	0.072***
	[0.010]	[0.009]
Ln(Tract Median Income)	-0.171***	-0.084***
	[0.013]	[0.013]
8th Grade Math	-0.010	-0.012**
	[0.006]	[0.005]
8th Grade Reading	-0.015**	-0.015**
	[0.006]	[0.006]
Neighborhood School Fixed Effects		X
Includes Students Who Choose Neighborhood School	X	X
Sample Size	16,201	16,201
Adjusted R-squared	0.122	0.172

Notes: Column 1 is a regression of the probability that a student in the analysis sample will choose a non-neighborhood school on the variables listed in each row. Column 2 is the same model but with neighborhood school fixed effects. Our measures of relative distance exclude the highest and lowest quality schools in CMS by construction, so those schools are not included in the regressions.

Table 6 - Impact of Winning the Lottery by Neighborhood School Quality

	HQ Neighborhood School		LQ Neighborhood School		F(equal)
	(1)	(2)	(3)	(4)	(5)
9th Grade English Score	-0.124	-0.058	-0.164	-0.003	0.286
		[0.033]		[0.052]	
9th Grade Math Score	-0.375	-0.219***	-0.469	0.038	0.002
		[0.052]		[0.049]	
9th Grade Math (impute 8th)	-0.449	-0.173***	-0.576	0.059*	0.001
		[0.052]		[0.032]	
Graduated from CMS High School	0.650	-0.055*	0.568	0.087***	0.006
		[0.032]		[0.030]	
Attended a 2 Year College	0.511	0.057*	0.542	0.018	0.437
		[0.031]		[0.043]	
Attended a 4 Year College	0.444	-0.042	0.316	0.060*	0.006
		[0.034]		[0.031]	
Very Competitive	0.119	0.035**	0.080	0.015	0.219
		[0.015]		[0.012]	
Most Competitive	0.038	-0.004	0.015	0.014**	0.133
		[0.011]		[0.006]	
Persistence (4+ sems)					
Any 2 year college	0.201	0.042	0.195	0.009	0.451
		[0.042]		[0.029]	
Any 4 year college	0.360	-0.038	0.232	0.031	0.042
		[0.040]		[0.020]	
Very Competitive	0.104	0.032	0.064	0.027**	0.797
		[0.019]		[0.010]	
Most Competitive	0.033	-0.004	0.013	0.012**	0.080
		[0.009]		[0.005]	
Earned a Degree					
Any 2 year college	0.025	0.032*	0.026	-0.001	0.081
		[0.017]		[0.013]	
Any 4 year college	0.246	-0.069**	0.131	0.046***	0.001
		[0.027]		[0.016]	
Very Competitive	0.094	-0.017	0.046	0.025***	0.025
		[0.022]		[0.008]	
Most Competitive	0.032	-0.016	0.011	0.011**	0.005
		[0.010]		[0.005]	
Sample Size	654		1211		

Notes : Columns 1 and 3 are mean values of the outcome in each row for students who lost the lottery. Columns 2 and 4 are estimates from equation (9) in the paper, with a full set of indicator variables for lottery status and low or high quality neighborhood school. Standard errors are below each estimate in brackets and are clustered at the lottery level. Column 5 contains the p-value for an F-test that columns (2) and (4) are equal. Competitive colleges are defined by the Barron's rankings - see the text of the paper for details.

Table 7 - Impact of Winning the Lottery on Mediating High School Outcomes

	HQ Neighborhood School		LQ Neighborhood School		F(equal)
	Ctrl Mean		Ctrl Mean		
	(1)	(2)	(3)	(4)	(5)
Grade Point Average	2.166	0.009	2.003	0.149***	0.030
		[0.046]		[0.041]	
GPA -Math and Science	1.809	-0.065	1.604	0.125**	0.003
		[0.048]		[0.047]	
Number of Math Courses	2.566	-0.095	2.727	0.220***	0.022
		[0.106]		[0.068]	
Took Math Every Year	0.653	-0.045	0.597	0.080***	0.061
		[0.053]		[0.028]	
Absences in 2003	10.61	1.19	12.51	-1.95**	0.009
		[0.89]		[0.82]	
Sample Size	654		1211		

Notes : Columns 1 and 3 are mean values of the outcome in each row for students who lost the lottery. Columns 2 and 4 are estimates from equation (9) in the paper, with a full set of indicator variables for lottery status and low or high quality neighborhood school. Standard errors are below each estimate in brackets and are clustered at the lottery level. Column 5 contains the p-value for an F-test that columns (2) and (4) are equal. Competitive colleges are defined by the Barron's rankings - see the text of the paper for details.

Table 8 - Impact of Winning the Lottery on Enrollment and School Characteristics

	HQ Neighborhood School		LQ Neighborhood School		F(equal)
	Ctrl Mean		Ctrl Mean		
Enrollment in Fall 2002	(1)	(2)	(3)	(4)	(5)
In Any CMS School	0.903	0.023	0.913	0.028	0.853
		[0.021]		[0.023]	
In 1st Choice	0.442	0.496***	0.311	0.586***	0.201
		[0.079]		[0.063]	
Neighborhood School	0.407	-0.405***	0.421	-0.364***	0.501
		[0.060]		[0.040]	
Magnet School	0.136	0.195	0.114	0.364***	0.092
		[0.116]		[0.123]	
Distance to Fall 2002 school	6.65	2.76***	5.99	0.72	0.003
		[0.60]		[0.58]	
Percent African-American	0.334	0.059	0.542	-0.001	0.164
		[0.058]		[0.040]	
Percent Free Lunch	0.373	0.040	0.592	-0.034	0.061
		[0.053]		[0.038]	
Avg. 8th Grade Math Score	0.175	-0.041	-0.220	0.105	0.028
		[0.087]		[0.065]	
School Quality Measures					
9th Grade Math	0.034	-0.069**	-0.011	-0.037	0.246
		[0.022]		[0.038]	
9th Grade English	0.036	-0.063*	-0.045	0.001	0.008
		[0.032]		[0.036]	
High School Graduation	0.001	0.010	-0.022	0.034***	0.042
		[0.009]		[0.008]	
4 Year College Attendance	-0.000	0.001	-0.024	0.036***	0.008
		[0.013]		[0.007]	
Very Competitive College	-0.004	-0.011	-0.020	0.012	0.039
		[0.010]		[0.008]	
Sample Size	654		1211		

Notes : Columns 1 and 3 are mean values of the outcome in each row for students who lost the lottery. Columns 2 and 4 are estimates from equation (9) in the paper, with a full set of indicator variables for lottery status and low or high quality neighborhood school. Standard errors are below each estimate in brackets and are clustered at the lottery level. Column 5 contains the p-value for an F-test that columns (2) and (4) are equal. The quality estimates for each school come from Table 2.

Table 9 - Comparison of School Value-Added Estimates to the Impact of Winning the Lottery

		Measured School Quality Difference (S_j - S_n)				
		Unadjusted	8th Grade Xs	All Covariates	Pr(choice)	Relative Distance
Main Outcomes	(1)	(2)	(3)	(4)	(5)	(6)
9th Grade English Score	-0.017	0.035	0.052	0.056	-0.104	0.054
	[0.044]	[0.030]	[0.033]	[0.039]	[0.154]	[0.051]
Interaction Term		0.083	0.482	0.549	0.240	-0.005
		[0.077]	[0.354]	[0.439]	[0.214]	[0.009]
F(main effect = 0, interaction = 1)		*0.000*	*0.002*	*0.001*		
9th Grade Math Score	-0.025	0.003	0.023	0.028	-0.470**	0.012
	[0.051]	[0.052]	[0.065]	[0.068]	[0.188]	[0.086]
Interaction Term		0.296**	0.756	0.803	0.773**	-0.008
		[0.132]	[0.520]	[0.538]	[0.237]	[0.013]
F(main effect = 0, interaction = 1)		*0.000*	*0.464*	*0.447*		
Graduated from CMS	0.036*	-0.002	0.015	0.023	-0.127	0.052
	[0.020]	[0.026]	[0.026]	[0.024]	[0.103]	[0.034]
Interaction Term		0.522**	0.937*	0.979**	0.289	-0.002
		[0.209]	[0.463]	[0.428]	[0.187]	[0.006]
F(main effect = 0, interaction = 1)		*0.023*	*0.808*	*0.431*		
Attended a 4 Year College	0.024	0.018	0.005	0.007	-0.072	0.049*
	[0.029]	[0.037]	[0.028]	[0.028]	[0.082]	[0.024]
Interaction Term		0.095	0.866*	0.974	0.172	-0.005
		[0.196]	[0.459]	[0.611]	[0.120]	[0.005]
F(main effect = 0, interaction = 1)		*0.000*	*0.954*	*0.967*		
Very Competitive College	0.024**	0.020**	0.020**	0.020**	0.032	0.006
	[0.011]	[0.009]	[0.009]	[0.009]	[0.029]	[0.010]
Interaction Term		0.068	0.301	0.320	-0.019	0.001
		[0.077]	[0.344]	[0.409]	[0.046]	[0.002]
F(main effect = 0, interaction = 1)		*0.000*	*0.001*	*0.003*		
Sample Size	1865					

Notes: Column 1 is the δ coefficient from equation (2) for each of the five main outcomes in the paper. Columns (2) through (5) estimate equation (10), where winning the lottery is interacted with the estimated difference in quality between the school that applicants attend if they win the lottery and the school they attend or would have attended had they lost. Column 3 is our preferred estimate, and measures quality for each school by the estimates in Table 2. Column 5 interacts winning the lottery with the probability that students in each neighborhood school zone apply to another school. Column 6 interacts winning the lottery with the additional distance beyond the neighborhood school that an applicant has to travel to attend their chosen school.

Table A1 - Impact of Winning the Lottery on Pre-Lottery Covariates

	High Quality Neighborhood School	Low Quality Neighborhood School
	(1)	(2)
Median HH Income	-1319	-149
	[2,136]	[902]
Male	0.013	0.012
	[0.046]	[0.025]
Black	0.064*	0.033
	[0.037]	[0.026]
Latino	-0.014	-0.015
	[0.015]	[0.016]
FRPL	0.022	0.014
	[0.029]	[0.036]
Special Education	0.018	-0.024
	[0.030]	[0.025]
8th Grade Math	-0.067	0.069
	[0.088]	[0.052]
8th Grade Reading	-0.113	0.030
	[0.076]	[0.045]
Distance to Home	0.23	0.10
	[0.26]	[0.20]
Distance to Choice	1.07**	-0.51
	[0.41]	[0.41]
Sample Size	1865	

Notes : Columns 1 and 2 are estimates from a form of equation (8) in the paper, with a full set of indicator variables for lottery status and low or high quality neighborhood school. Standard errors are below each estimate in brackets and are clustered at the lottery level.