NBER WORKING PAPER SERIES

THE DEMAND FOR EFFECTIVE CHARTER SCHOOLS

Christopher R. Walters

Working Paper 20640
http://www.nber.org/papers/w20640

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2014

I am grateful to Joshua Angrist, Parag Pathak, and David Autor for their guidance and support. Seminar participants at APPAM, Cornell University, Duke University, Mathematica Policy Research, MIT, Northeastern University, the Spencer Foundation, the Stanford Graduate School of Business, the St. Louis Federal Reserve, UC Berkeley, UC Davis, UCLA, the University of Chicago Harris School of Public Policy, the University of Michigan, the University of Virginia, and Warwick University provided useful comments. I also thank Aviva Aron-Dine, David Card, David Chan, Sarah Cohodes, William Darrity Jr., Susan Dynarski, Maria Ferreyra, Patrick Kline, Michal Kurlaender, Bridget Terry Long, Enrico Moretti, Christopher Palmer, Jesse Rothstein, Stephen Ryan, Steven Stern, Xiao Yu Wang, Tyler Williams, and numerous others for helpful discussions. Special thanks go to Carrie Conaway and the Massachusetts Department of Elementary and Secondary Education for suggestions, assistance and data. This work was supported by a National Science Foundation Graduate Research Fellowship, a National Academy of Education/Spencer Dissertation Fellowship, and Institute for Education Sciences award number R305A120269. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Christopher R. Walters. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Demand for Effective Charter Schools
Christopher R. Walters
NBER Working Paper No. 20640
October 2014
JEL No. C25,I21,J2,J24

ABSTRACT

This paper uses a structural model of school choice and academic achievement to study the demand for charter schools in Boston, Massachusetts, with an emphasis on comparative advantage in school choice. I combine an optimal portfolio choice model of charter school application and attendance decisions with a selection correction approach that links students' school choices to the achievement gains generated by charter attendance. To estimate the model, I use instrumental variables derived from randomized entrance lotteries, together with a second set of instruments based on distance to charter schools. The estimates show that charter schools reduce achievement gaps between high- and low-achieving groups, so disadvantaged students and low-achievers have a comparative advantage in the charter sector. Higher-income students and students with high prior achievement have the strongest demand for charter schools, however, which implies that preferences for charters are inversely related to potential achievement gains. The structural estimates show a similar pattern of selection on unobservables. These findings imply that students do not sort into charter schools on the basis of comparative advantage in academic achievement; instead, disadvantaged students are less likely to apply to charter schools despite larger potential achievement gains. I use simulations of an equilibrium school choice model to quantify the consequences of this demand-side pattern for the effects of charter school expansion. The results suggest that in the absence of significant behavioral or institutional changes, the effects of charter expansion may be limited as much by demand as by supply.

Christopher R. Walters
Department of Economics
University of California at Berkeley
530 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
crwalters@econ.berkeley.edu

1 Introduction

Reforms that broaden the scope for school choice are an increasingly common phenomenon in US public school districts. One such reform is the creation of charter schools, which are publicly funded, non-selective schools that typically have more freedom than traditional public schools to set curricula and make staffing decisions. While evidence on the effects of non-urban charter schools is mixed,[1] studies based on entrance lotteries show that attendance at charter schools in Boston and New York's Harlem Children's Zone boosts academic achievement sharply for poor, minority applicants (Abdulkadiroglu et al. 2011; Dobbie and Fryer 2011). Angrist et al. (2012, 2013a, 2013b), Dobbie and Fryer (2013), Gleason et al. (2010), Hoxby and Murarka (2009), and Hoxby and Rockoff (2004) also report positive effects for urban charters. These findings suggest that urban charter schools may have the potential to reduce achievement gaps between racial and socioeconomic groups. Reflecting this hope, the Massachusetts legislature recently relaxed the state's charter school cap with the explicit goal of reducing racial and socioeconomic disparities in academic performance (Commonwealth of Massachusetts 2010).

Despite the large literature documenting the causal effects of charter schools and other school-choice programs, little attention has been paid to the demand for these programs. Existing studies typically restrict attention to samples of lottery applicants, among whom lottery offers are randomly assigned (see, e.g., Abdulkadiroglu et al. 2011 and Deming et al. 2014). Understanding the application decisions that generate these samples is important both for interpreting existing evidence and for evaluating the potential effects of charter school expansion. Of particular interest is whether students sort into the charter sector on the basis of comparative advantage. That is, are students with more to gain from charter attendance more likely to apply to charter schools? If the students with the most to gain are most likely to sign up for charter lotteries, then local average treatment effects (LATE) derived from lottery-based instruments will overstate potential gains for non-applicants and provide a misleading picture of the effects of charter expansion (Imbens and Angrist 1994; Rothstein 2004).[2] On the other hand, the parents of low-achieving students may be unlikely to investigate alternatives to traditional public school, despite evidence that urban charter schools are especially effective for such students (Angrist et al. 2012). The case for school choice is also stronger if increasing choice allows students to sort according to comparative advantage, boosting aggregate achievement.

This paper studies the demand for charter middle schools in Boston, with a focus on absolute and comparative advantage in school choice. I answer two related questions: (1) Do students sort into the charter sector on the basis of potential achievement gains? and (2) What are the likely effects of charter attendance for students who would be induced to attend by a marginal expansion of the charter sector? To answer these questions, I develop and estimate a structural model that links charter application and enrollment decisions to potential achievement gains in a parametric selection framework.

Students in Boston can apply to any combination of charter schools, and face uncertainty in the form of an admissions lottery at each charter. I study this process using a dynamic, unordered discrete choice model

[1] Gleason et al. (2010) find that non-urban charters are no more effective than traditional public schools. Angrist et al. (2013b) find negative effects for non-urban charter middle schools in Massachusetts. In an observational study of 27 states, CREDO (2013) finds that charter schools are slightly more effective than traditional public schools on average.

[2] Rothstein (2004, p.82) writes of the Knowledge is Power Program (KIPP), a high-performing urban charter operator: "[T]hese exemplary schools...select from the top of the ability distribution those lower-class children with innate intelligence, well-motivated parents, or their own personal drives, and give these children educations they can use to succeed in life."

of application portfolio choices, lottery outcomes, and school attendance decisions. The model is similar to the stochastic portfolio choice problems considered by Chade and Smith (2006) and Chade et al. (2009): students submit charter application portfolios to maximize expected utility, taking account of admission probabilities and non-monetary application costs. I estimate the model using admissions records for seven of the nine Boston charter middle schools operating during my sample period, including in years when these schools were undersubscribed and did not hold lotteries. This allows me to develop a complete model of the charter application portfolio choice set for these seven schools. As far as I know, this is the first paper to estimate a dynamic discrete choice model of application portfolio and school attendance choices. The model estimated here may be useful in other settings in which students submit many applications and face uncertainty in admission decisions, including the college admissions process.

I combine the portfolio-choice model of charter demand with a flexible model of academic achievement that describes potential achievement in charter schools and traditional public schools. The link between choices and achievement takes the form of a selection correction, which captures the relationship between potential outcomes and the latent preferences that drive charter application and attendance decisions (Heckman 1979). This approach is related to the models analyzed by Lee (1983), Dubin and McFadden (1984), and Dahl (2002), which extend the canonical Heckman (1979) sample selection model to multinomial choice settings. The model estimated here is dynamic, and the set of available portfolio choices is high-dimensional. To reduce the dimensionality of the problem, I impose a "single-index" restriction, which requires that endogenous selection into the charter sector is driven by a one-dimensional unobserved preference. I show that charters in Boston are highly homogenous, and that students treat them as close substitutes conditional on applying; this suggests that a one-factor selection model is reasonable. To identify the parameters of the model, I combine instruments from randomized entrance lotteries with a second set of instruments based on proximity to charter schools. The use of a selection correction approach in a dynamic discrete choice model with endogeneity is a second methodological contribution of the paper.

Estimates of the model reveal that students do not sort into charter schools on the basis of comparative advantage. Richer students and those with higher previous achievement have absolute advantages in both charter schools and traditional public schools, but charter schools boost scores more for poor students and low-achievers; disadvantaged students therefore have a comparative advantage in the charter sector. Richer students and high-achievers have stronger tastes for charter schools, however, which implies that charter applicants are *negatively* selected on potential achievement gains. The structural estimates show a similar pattern of selection on unobservables. This pattern is consistent with the possibility that effective charter schools partially compensate for differences in human capital investments across families, but motivated parents who invest more at home are also more likely to enroll their children in charter schools. As a result, existing estimates *understate* the potential achievement effects of charter schools for non-applicants: effects of treatment on the treated (TOT) for current applicants are roughly 25 percent smaller than the model's implied population average treatment effects (ATE), which are smaller still than the potential effects of treatment on the non-treated (TNT). This suggests that efforts to recruit students who are currently unlikely to apply to charters could boost productivity in the charter sector.

I quantify these findings by simulating the effects of charter expansion in an equilibrium school choice model assuming constant returns to scale (CRTS) on the supply side. While this assumption is unlikely to hold in reality, most departures from CRTS, such as inelastic supply of teachers and other inputs, are likely to reduce the efficacy of new charter schools. The results therefore provide a plausible upper bound on the benefits of charter expansion due to demand-side behavior. The simulations show that while effects for marginal applicants are substantial, the effects of charter expansion may nevertheless be limited by weak demand: students act as if charter application costs are high, and many of the highest-benefit students prefer to attend traditional public schools even when charters offering guaranteed admission are located in close proximity. These results suggest that in the absence of significant behavioral or institutional changes, the effects of charter expansion may be limited as much by demand as by supply.

In additional to the methodological literatures on dynamic discrete choice and multinomial sample selection models, this paper contributes to a nascent literature assessing the relationships between educational choices, treatment effects, and socioeconomic status. Arcidiacono et al. (2013) study differences in selection into college majors between minority and non-minority students. Hastings et al. (2009) show that higher-socioeconomic-status parents are more likely to choose schools with high average test scores in a public school choice plan in Charlotte, North Carolina. Kirkeboen et al. (2014) find evidence of comparative advantage in field choices in the Norwegian higher-education system. Using a nationally representative sample from the Early Childhood Longitudinal Study, Butler et al. (2013) show that richer children are more likely to enroll in charter schools. In the higher-education sphere, Hoxby and Avery (2012) show that seemingly-qualifed poor students are unlikely to apply to selective colleges. Similarly, Brand and Xie (2010) argue that students with larger potential returns are less likely to enroll in college, while Dillon and Smith (2013) document the prevalence of "mismatch" between student ability and college quality in the National Longitudinal Survey of Youth. Farther afield, Ajayi (2013) shows that students from lower-performing elementary schools in Ghana are less likely to apply to selective secondary schools. Consistent with the findings from this literature, I find that disadvantaged students are unlikely to apply to effective charter schools, despite large potential achievement gains.

This paper also adds to a series of studies using structural models to analyze educational decisions and outcomes. Epple and Romano (1998) study a theoretical model of competition between public and private schools. Epple et al. (2003, 2006, 2013) analyze theoretical and empirical models of college quality, student sorting, and financial aid. In more closely related work, Ferreyra and Kosenok (2011) and Mehta (2011) study models of charter school entry and student sorting. Rather than modeling charter entry decisions, I focus on estimation of the joint distribution of student preferences and achievement gains. The detailed student-level application, lottery offer, and achievement data used here allow for a more flexible specification of student preferences than those used in previous studies. The use of entrance lotteries also allows me to identify achievement effects using random assignment rather than the selection-on-observables assumptions implicit in these studies. Finally, this is the first paper to model the charter application decision as an optimal portfolio choice problem.[3]

The rest of the paper is organized as follows: The next section gives background on charter schools in Boston and describes the data. Section 3 describes the reduced-form patterns in charter application decisions and lottery-based estimates that motivate the structural analysis to follow. Section 4 outlines the structural model of charter demand and academic achievement, and Section 5 discusses identification and estimation of the model. Section 6 reports

[3]Mehta (2011) and Ferreyra and Kosenok (2011) do not use admissions lottery data, so they cannot observe application portfolio choices.

the structural estimates. Section 7 discusses implications of the estimates and uses the model to simulate the effects of counterfactual policies. Section 8 concludes.

2 Setting and Data

2.1 Context: Charter Schools in Boston

Non-profit organizations, teachers, or other groups wishing to operate charter schools in Massachusetts submit applications to the state's Board of Education. If authorized, charter schools are granted freedom to organize instruction around a philosophy or curricular theme, as well as budgetary autonomy. Charter employees are also typically exempt from local collective bargaining agreements, giving charters more discretion over staffing than traditional public schools.[4] Charters are funded primarily through per-pupil tuition payments from local districts. Charter tuition is roughly equal to a district's per-pupil expenditure, though the state Department of Elementary and Secondary Education partially reimburses these payments (Massachusetts Department of Elementary and Secondary Education 2011). The Board of Education reviews each charter school's academic and organizational performance at five year intervals, and decides whether charters should be renewed or revoked.

Enrollment at Massachusetts charter schools is open to all students who live in the local school district. If a charter school receives more applications than it has seats, it must accept students by random lottery. Students interested in multiple charter schools must submit separate applications to each charter, and may receive multiple offers through independent school-specific lotteries. This system of independent enrollment processes is in contrast to the centralized enrollment mechanism used for Boston's traditional public schools, which collects lists of students' preferences over schools and generates a single offer for each student (Pathak and Sonmez 2008).

The Boston Public Schools (BPS) district is the largest school district in Massachusetts, and it also enrolls an unusually large share of charter students. In the 2010-2011 school year, 14 charter schools operated in Boston, accounting for 9 percent of BPS enrollment. The analysis here focuses on middle schools, defined as schools that accept students in fifth or sixth grade; 12 percent of Boston middle schoolers attended charter schools in 2010-2011. Panel A of Appendix Table A1 lists names, grade structures and years of operation for the nine Boston charter middle schools that operated through the 2010-2011 school year. I use admissions records from seven of these schools to produce the estimates reported below.

Many of Boston's charter schools adhere to a model known as "No Excuses," a set of practices that includes extended instruction time, strict behavior standards, a focus on traditional reading and math skills, selective teacher hiring, and teacher monitoring (Wilson 2008). A growing body of evidence suggests that these practices boost student achievement (Angrist et al., 2013b; Curto and Fryer, 2011; Dobbie and Fryer, 2013; Fryer, 2011). Consistent with this evidence, Abdulkadiroglu et al. (2011) use entrance lotteries to show that Boston's charter schools substantially increase achievement among their applicants. Their estimates imply that a year of charter middle school attendance raises test scores by 0.4 standard deviations (σ) in math and 0.2σ in reading. Similarly,

[4] Massachusetts has two types of charter schools: Commonwealth charters, and Horace Mann charters. Commonwealth charters are usually new schools authorized directly by the Board of Education, while Horace Mann charters are often conversion schools and must be approved by the local school board and teachers' union prior to state authorization. Horace Mann employees typically remain part of the collective bargaining unit. I focus on Commonwealth charter schools. No Horace Mann charter middle schools operated in Boston during my data window.

Angrist et al. (2013a) show that Boston's charter high schools have substantial effects on longer-term outcomes like SAT scores and four-year college enrollment.

These encouraging findings make Boston an appealing setting for studying the demand for effective charter schools. The demand for charters in Boston is also relevant to an ongoing policy debate. In recent years, the growth of charters in Massachusetts has been slowed by the state's charter cap, a law that limits expenditures on charter schools to 9 percent of the host district total.[5] The Board of Education stopped accepting proposals for new Boston charters in 2008 when charter expenditure hit the cap (Boston Municipal Research Bureau 2008). In 2010, the Massachusetts legislature relaxed the charter cap for school districts in the state's lowest test score decile. For these districts, the limit on charter expenditures is to rise incrementally from 9 percent in 2010 to 18 percent in 2017 (Commonwealth of Massachusetts 2010). Through 2011, the Board of Education received 51 charter applications under the new law and granted 20 charters, eleven to schools in Boston (Massachusetts Department of Elementary and Secondary Education 2012b). Panel B of Appendix Table A1 lists the six charter middle schools opened through the 2012-2013 school year. Column (6) indicates existing charters operated by the same organizations. Boston's charter sector may continue to expand in the near future; recently-proposed legislation would eliminate the charter cap in Boston and other low-performing districts (Levitz 2013).

2.2 Data Sources and Sample Construction

The data used in my analysis comes from three sources. First, I obtain demographics, school attendance, and test scores from an administrative database provided by the Massachusetts Department of Elementary and Secondary Education (DESE). Second, I draw spatial location data from student addresses provided by the BPS district. Finally, I obtain information on charter school applications and offers from lottery records gathered from individual charter schools.

The DESE database covers all Massachusetts public school students from the 2001-2002 school year through the 2012-2013 school year. Key variables include sex, race, subsidized lunch status, limited English proficiency (LEP), special education status (SPED), town of residence, schools attended, and scores on Massachusetts Comprehensive Assessment System (MCAS) math and English Language Arts (ELA) achievement tests. I begin by selecting from the database the four cohorts of students who attended a traditional BPS school in 4th grade between 2005-2006 and 2008-2009. I also require students to have non-missing 4th grade demographics and test scores, as well as school attendance information and test scores in 6th, 7th, or 8th grade. I use only the earliest test taken by a given student in a particular subject and grade. Test scores are standardized to have mean zero and standard deviation one within each subject-year-grade in Massachusetts.

Next, I merge the student address database to the DESE administrative file using a crosswalk between BPS and state student identifiers. The address database includes a record for every year that a student attended a traditional BPS school between 1998 and 2011. I drop students in the state database without 4th-grade BPS address data. This restriction eliminates less than 1 percent of Boston 4th graders. The address information is used to measure proximity to each Boston charter school. I measure proximity using great-circle distance in miles.[6]

[5] Legislation also limits the total number of Commonwealth charter schools to 72 and the number of Horace Mann charters to 48, though these caps are not currently binding.

[6] I also performed the analysis using travel times measured by Google Maps, obtained using the STATA *traveltime* command. I chose to use great-circle distances instead because *traveltime* produced different results when queried at different times, making the

I then match the student data to admissions records from seven charter middle schools in Boston. These seven schools provided complete records for applicant cohorts attending 4th grade between 2006 and 2009, including in years when they were undersubscribed. Importantly, only nine charter middle schools operated in Boston during my sample period, so the admissions data provides a nearly-complete picture of charter application decisions during this period. Column 4 of Appendix Table A1 summarizes the availability of admissions records for the nine charter middle schools that operated between the 1997-1998 school year and the 2010-2011 school year.[7] Of the two schools without available records, one closed prior to the 2010-2011 school year; the other declined to provide records. In the analysis below, I treat these schools as equivalent to traditional public schools. I matched the available records to the administrative data by name, grade, year, and (where available) date of birth. This process produced unique matches for 92 percent of applicants. Though admission records for all seven schools were available for cohorts attending 4th grade between 2006 and 2009, not every school was oversubscribed in every year, so schools did not always hold lotteries. Column (5) of Table A1 shows that each of the seven schools held lotteries in at least two years. The analysis below uses applicant records for all four years, setting admission probabilities to one for undersubscribed years.

After matching the admission files to the student data, I constructed two subsamples for statistical analysis. The first is used to estimate causal effects for lottery applicants, and thus excludes students who did not apply to charter schools. The lottery sample includes 1,794 applicants to charter middle schools. A second sample, which includes both applicants and non-applicants, is used to investigate charter application behavior and estimate the structural model. The full sample includes 10,122 students who attended BPS schools in 4th grade between 2006 and 2009.

3 Patterns in Application Decisions and Test Score Effects

3.1 Charter Application Decisions

To motivate my investigation of comparative advantage in charter school choice, I begin with a reduced-form analysis of patterns in charter application decisions and achievement effects. Table 1 compares mean characteristics for the full Boston population (column (1)) to characteristics of charter applicants (column (2)). Eighteen percent of Boston students applied to at least one charter lottery, thirteen percent were offered a charter seat, and eleven percent attended a charter school. Five percent of students applied to more than one charter. Charter applicants tend to have higher socioeconomic status and fewer academic problems than non-applicants. Specifically, applicants are less likely to be eligible for subsidized lunch (a proxy for poverty), to have special education status, or to be classified as limited English proficient. Charter applicants are less likely to be Hispanic and slightly more likely to be white than non-applicants. Applicants also live slightly closer to charter schools on average (1.9 miles from the closest charter, compared to 2.1 miles for the average student).

Table 1 also investigates the relationship between charter application decisions and the quality of nearby traditional public school options. Students can choose between many schools in Boston's centralized traditional public

results difficult to replicate. Key estimates were very similar for this alternative distance measure.

[7] I classify charter schools as middle schools if they accept applicants in 5th or 6th grade. Two Boston charter schools accept students prior to 5th grade but serve grades 6 through 8. Since I restrict the analysis to students who attended traditional BPS schools in 4th grade, no students in the sample attend these schools.

school assignment mechanism, so the characteristics of the relevant traditional public school are unknown for students who attend charter schools. Most students attend a school close to home, however, so the average quality of schools in a student's zip code is a reasonable proxy for the quality of available traditional public options. I estimate the quality of nearby schools using a value-added regression of test scores for traditional public school students in grades 6 through 8 on 4th-grade test scores, the other demographic variables from Table 1, and school fixed effects. The value-added for a given school is the coefficient on the relevant fixed effect in this regression. I then compute the average value-added of the schools attended by students in each zip code, and assign each student the average of the math and ELA value-added measures for her zip code. I jackknife the value-added procedure so that a student's own score does not contribute to her own measure of value-added. Columns (1) and (2) of Table 1 show that the average value-added of nearby traditional public schools is similar for charter applicants and non-applicants.

The last two rows of Table 1 display information about 4th grade MCAS scores. Boston 4th graders lag behind the state average by 0.55σ and 0.66σ in math and ELA. Students who apply to charter schools have substantially higher scores than the general Boston population: applicants' 4th grade scores exceed the Boston average by more than 0.2σ in both subjects.

Column (3) summarizes these demographic patterns with a probit model of the form

$$Pr[A_i = 1|X_i, D_i] = \Phi\left(\gamma^0 + X_i'\gamma^x + \gamma^d D_i\right), \tag{1}$$

where A_i is an indicator for applying to at least one charter school, X_i is a vector of baseline characteristics, D_i is the distance to the closest charter, and $\Phi(\cdot)$ is the normal CDF. The reported coefficients are marginal effects on the charter application probability evaluated at the sample means of X_i and D_i. The results show that charter application decisions are strongly correlated with poverty status: students eligible for subsidized lunch are 9 percentage points less likely to apply to a charter school conditional on the other included covariates, a large effect relative to the baseline application rate of 18 percent. The probit estimates also reveal that special education students, students with limited English proficiency, and students with low baseline test scores are less likely to apply, and these relationships are statistically significant. The effect for whites is negative and significant, which implies that the higher unconditional application rate for whites is explained by the other covariates; the negative coefficients for whites and Hispanics imply that black students are more likely to apply to charter schools conditional on observed characteristics. The marginal effect for traditional public school value-added is also negative and significant. This suggests that after accounting for other observed characteristics, students with weaker traditional public school options are more likely to apply to charter schools. Taken together, the statistics in Table 1 suggest that the charter applicant pool is skewed towards economically-advantaged, higher-performing students.

3.2 Lottery Estimates of Test Score Effects

I next use instruments based on entrance lotteries to summarize the causal effects of charter schools in the selected sample of lottery applicants.[8] The estimating equations for the lottery analysis take the form

$$Y_i = \psi_\ell + \beta S_i + \epsilon_i, \tag{2}$$

where Y_i is a test score for charter applicant i, S_i is a dummy variable indicating charter school attendance, and ψ_ℓ is a set of lottery application portfolio fixed effects (dummies for all observed combinations of schools and applicant cohorts). I code a student as attending a charter school if she attends a charter at any time after the lottery and prior to the test. The first stage equation is

$$S_i = \kappa_\ell + \pi Z_i + \eta_i. \tag{3}$$

The instrument Z_i is one if student i received any charter offer before the start of the school year following the lottery. The 2SLS estimate of β can be interpreted as a weighted average of within-lottery local average treatment effects (LATE), defined as effects of charter schools on applicants induced to attend by lottery offers (Imbens and Angrist 1994).[9] To use all available test score information, the sample stacks scores in grades six through eight. Standard errors are clustered at the student level.[10]

Consistent with the results reported by Abdulkadiroglu et al. (2011), the 2SLS estimates show that Boston's charter schools have dramatic effects on student achievement for lottery applicants. As shown in column (1) of Table 2, receipt of a lottery offer increases the probability of charter attendance by 0.68. The second-stage estimates, reported in columns (2) and (3), imply that attending a charter school increases math scores by 0.59σ and boosts ELA scores by 0.41σ. These effects are precisely estimated ($p < 0.01$).

These pooled results mask substantial heterogeneity in the benefits of charter school attendance across demographic groups. Columns (3) though (8) of Table 2 show estimates from three 2SLS models that interact charter attendance with baseline characteristics, instrumenting with interactions of lottery offers and these characteristics. Columns (3) and (4) show that the estimated effects of charter attendance are larger for students with subsidized lunch status. The difference in effects between subsidized lunch and non-subsidized lunch students is significant at the 10-percent level in math, through insignificant in ELA (p=0.060 and 0.161, respectively). The remaining columns show that the effects of charter attendance are larger for non-white students and for students with below-median previous test scores than for students that lack these characteristics. The estimated effects for white students are small and statistically insignificant in both subjects. These results show that Boston's charter schools generate larger gains for disadvantaged, low-achieving, and non-white applicants.

[8]Appendix Table A2 shows that baseline covariates are balanced across lottery winners and losers, suggesting that randomization was successful in the lotteries used here. Even with random assignment, lottery-based instruments can be compromised by non-random attrition. Appendix Table A3 shows that follow-up rates are high: 89 percent of potential scores in grades 6 through 8 are observed in the full sample, while 87 percent are observed in the lottery sample. Moreover, there is no difference in follow-up rates between lottery winners and losers, so attrition is unlikely to bias my estimates.

[9]With a first stage saturated in offer-times-lottery interactions, 2SLS produces a weighted average of lottery-specific IV estimates, with weights proportional to the variance of the first-stage fitted values (Angrist and Imbens 1995). Estimates from this fully saturated model were similar to the more parsimonious model used here, which includes a single instrument and lottery fixed effects.

[10]To keep sample sizes consistent, I include all charter applicants in these models, including undersubscribed cohorts. Since all students in undersubscribed cohorts are offered seats, the lottery fixed effects absorb the offer dummy for these cohorts, and their inclusion has no effect on the estimates.

Taken together, the results reported in tables 1 and 2 suggest that disadvantaged students and low-achievers benefit more from charter school attendance, but are relatively unlikely to apply to charter schools. Figure 1 summarizes this pattern using the full set of demographic characteristics. To construct the figure, I use the probit model in equation (1) to compute a predicted charter application probability for each student. I then divide the sample of lottery applicants into quartiles of this predicted probability, and estimate a version of equation (2) that includes interactions of charter attendance with quartile dummies, instrumenting with interactions of the lottery offer and these dummies and controlling for quartile main effects. The results show that charter gains are *smaller* for applicants with higher predicted application probabilities. Equality of effects across quartiles is rejected at the 10-percent level in both subjects.

The standard Roy (1951) selection model predicts that the probability of applying to a charter school will be positively related to the achievement benefit generated by charter attendance. Figure 1 suggests the opposite pattern: charter schools produce larger gains for applicants with *smaller* predicted application probabilities. This suggests that students do not sort into the charter sector on the basis of comparative advantage. However, this figure may provide a misleading picture of the pattern of selection into charter schools for at least two reasons. First, the simple probit model in equation (1) is an inaccurate description of the charter school application decision. Students can apply to any combination of charter schools in Boston, and face different admission probabilities and distance costs at each charter. With heterogeneous schools and admission probabilities, predicted probabilities from equation (1) may not reflect the relationship between observed characteristics and tastes for charter schools. More importantly, there may be selection on unobserved dimensions of academic achievement. If students sort into the charter sector on the basis of private information about expected gains, applicants may experience atypically large benefits despite smaller predicted gains given their observed characteristics. The next section describes a structural model that incorporates the key features of the Boston charter landscape and allows selection on both observed and unobserved dimensions.

4 Modeling Charter School Attendance

4.1 Setup

I model charter application choices as a random utility optimal portfolio choice problem. Figure 2 explains the sequence of events described by the model. First, students decide whether to apply to each of K charter schools, indexed by $k \in \{1 \ldots K\}$. The dummy variable $A_{ik} \in \{0, 1\}$ indicates that student i applies to school k. Second, charter schools randomize offers to applicants. The dummy variable $Z_{ik} \in \{0, 1\}$ indicates an offer for student i at school k, and π_k denotes the admission probability for applicants to school k. In the third stage, students choose schools denoted S_i, where $S_i = 0$ indicates public school attendance. Any student can attend public school, but student i can attend charter school k only if she receives an offer at this school. Finally, students take achievement tests, with scores denoted Y_i.

4.2 Student Choice Problem

4.2.1 Preferences

Students' preferences for schools depend on demographic characteristics, spatial proximity, expected academic achievement, application costs, and unobserved heterogeneity. The utility of attending charter school k is

$$u_{ik} = \delta_k^0 + X_i'\delta^x + \gamma^d D_{ik} + \delta^y \tilde{Y}_i(k) + \nu_i + \upsilon_{ik} - c_i(A_i), \tag{4}$$

where X_i is a vector of observed characteristics for student i, D_{ik} measures distance to school k, and $\tilde{Y}_i(k)$ denotes student i's expected academic achievement if she attends school k. The utility of public school attendance is

$$u_{i0} = \upsilon_{i0} + \delta^y \tilde{Y}_i(0) + c_i(A_i). \tag{5}$$

Expected academic achievement in charter schools and public schools is assumed to take the form

$$\tilde{Y}_i(k) = \alpha_k^0 + X_i'\alpha_c^x + \eta_{ic} \quad \text{for } k > 0, \tag{6}$$

$$\tilde{Y}_i(0) = \alpha_0^0 + X_i'\alpha_0^x + \eta_{i0}. \tag{7}$$

The parameter α_k^0 is the academic quality of school k. This parameter captures factors that affect achievement of all students at the school, such as the quality of teachers, facilities, or the peer environment.[11] The subscript c indicates parameters and variables that are constant across charter schools. The variables η_{ic} and η_{i0} represent private information about potential scores in charter schools and public schools. As described in Section 4.3, realized achievement is modeled as $Y_i(k) = \tilde{Y}_i(k) + \mu_{ik}$, where μ_{ik} is unknown to the student at the time of the application and attendance decisions. This implies that students may have private information about their own idiosyncratic achievement benefits from charter schools compared to traditional public schools, but they do not make choices *between* charter schools on the basis of such information. Section 5 provides further motivation and evidence in support of this assumption.

Substituting equations (6) and (7) into equations (4) and (5) yields the following equivalent specification for student preferences:

$$U_{ik} = \gamma_k^0 + X_i'\gamma^x + \gamma^d D_{ik} + \theta_i + \upsilon_{ik} - c_i(A_i) \quad \text{for } k > 0, \tag{8}$$

$$U_{i0} = \upsilon_{i0} - c_i(A_i). \tag{9}$$

Though they do not explicitly include academic achievement, equations (8) and (9) implicitly allow for the possibility that tastes for schools are partially or entirely determined by expected achievement.[12]

The quantity $c_i(A_i)$ represents the utility cost of A_i, the application portfolio chosen by student i (Here and elsewhere, variables without k subscripts refer to vectors, so that $A_i \equiv (A_{i1}, \ldots, A_{iK})'$ and so on). Application costs

[11] To the extent that peer quality affect academic achievement, the parameter α_k^0 should change in counterfactuals that shift the composition of enrollees. Section 7 provides further discussion of the role of peer effects.

[12] Note that we can write $\gamma_k^0 = \delta_k + \delta^y(\alpha_k^0 - \alpha_0^0)$, $\gamma^x = \delta^x + \delta^y(\alpha_c^x - \alpha_0^x)$, and $\theta_i = \nu_i + \delta^y(\eta_{ic} - \eta_{i0})$.

include the disutility of filling out application forms and the opportunity cost of time spend attending lotteries. These costs may also capture frictions associated with learning about charter schools.[13] The application cost function is parameterized as

$$c_i(a) = \gamma^a |a| - \psi_{ia}.$$

The parameter γ^a is the marginal cost associated with an additional charter school application. The error term ψ_{ia} is a shock to the utility associated with a particular application portfolio. Applicants pay these costs whether or not they attend a charter.

The variables θ_i and υ_{ik} represent unobserved heterogeneity in tastes. θ_i, which characterizes student i's preference for charter schools relative to traditional public school, is the key unobservable governing selection into the charter sector. This variable includes any latent factors that influence students to opt out of traditional public school in favor of charter schools, such as the perceived achievement gain from attending charter schools, proximity or quality of the relevant traditional public school, or parental motivation. In the language of the random-coefficients logit model (see, e.g., Hausman and Wise 1978, Berry et. al. 1995, and Nevo 2000), θ_i is the random coefficient on a charter school indicator. The presence of θ_i implies that charter schools are closer substitutes for each other than for traditional public schools. I assume that θ_i follows a normal distribution with mean zero and variance σ_θ^2.

The υ_{ik} capture idiosyncratic preferences for particular schools, which are further decomposed as

$$\upsilon_{ik} = \tau_{ik} + \xi_{ik}.$$

Students know ψ_{ia}, τ_{ik}, and θ_i before applying to charter schools, and learn ξ_{ik} after applying. The post-application preference shock ξ_{ik} explains why some applicants decline charter school offers. To generate multinomial logit choice probabilities, ψ_{ia}, τ_{ik}, and ξ_{ik} are assumed to follow independent extreme value type I distributions, with scale parameters λ_ψ, λ_τ, and 1.[14]

4.2.2 School Lotteries

In the second stage of the model, schools hold independent lotteries. School k admits applicants with probability π_k. The probability mass function for the offer vector Z_i conditional on A_i is

$$f(Z_i|A_i;\pi) = \prod_k [A_{ik} \cdot (\pi_k Z_{ik} + (1-\pi_k)(1-Z_{ik})) + (1-A_{ik}) \cdot (1-Z_{ik})]. \tag{10}$$

I allow the admission probabilities π_k to vary by application cohort. If school k is undersubscribed and hence does not hold a lottery for a particular cohort, I set $\pi_k = 1$ for that cohort.

4.2.3 Application and Attendance Decisions

I derive students' optimal application and attendance rules by backward induction. A student is faced with a unique attendance decision after each possible combination of charter school offers, because the set of offers in hand

[13] Charter schools are not listed in informational resources provided to parents by the BPS district. For example, the "What Are My Schools?" tool located at www.bostonpublicschools.org provides a list of the BPS schools to which children are eligible to apply, but does not list charter schools (accessed September 13th, 2013).

[14] That is, ξ_{ik} follows a standard Gumbel distribution, which provides the scale normalization for the model.

determines the available school choices. Consider the decision facing a student at stage 3 in Figure 2. At this point, the student knows her charter offers, application costs are sunk, and there is no uncertainty about preferences. Student i can attend public school or any charter school that offers a seat. Her choice set is

$$C(Z_i) = \{0\} \cup \{k : Z_{ik} = 1\}.$$

Define

$$\tilde{U}_{ik}(\theta_i, \tau_{ik}) \equiv \gamma_k^0 + X_i'\gamma^x + \gamma^d D_{ik} + \theta_i + \tau_{ik} \quad \text{for } k>0,$$

with $\tilde{U}_{i0}(\theta_i, \tau_{i0}) \equiv \tau_{i0}$. Student i's optimal school choice is

$$S_i = \arg \max_{k \in C(Z_i)} \tilde{U}_{ik}(\theta_i, \tau_{ik}) + \xi_{ik},$$

and the probability that student i chooses school k at this stage is given by

$$Pr[S_i = k | X_i, D_i, Z_i, \theta_i, \tau_i] = \frac{\exp\left(\tilde{U}_{ik}(\theta_i, \tau_{ik})\right)}{\sum_{j \in C(Z_i)} \exp\left(\tilde{U}_{ij}(\theta_i, \tau_{ij})\right)}$$

$$\equiv P_{ik}(Z_i, \theta_i, \tau_i).$$

The expected utility associated with this decision (before the realization of ξ_i) is

$$W_i(Z_i, \theta_i, \tau_i) \equiv E\left[\max{}_{k \in C(Z_i)} \tilde{U}_{ik}(\theta_i, \tau_{i0}) + \xi_{ik} | X_i, D_i, Z_i, \theta_i, \tau_i\right]$$

$$= \nu + \log\left(\sum_{k \in C(Z_i)} \exp\left(\tilde{U}_{ik}(\theta_i, \tau_{ik})\right)\right),$$

where ν is Euler's constant.

Students choose charter applications to maximize expected utility, anticipating offer probabilities and their own attendance choices. Consider the application decision facing a student at stage 1 in Figure 2. The student knows θ_i, τ_i, and ψ_i, but does not know ξ_i, and her choice of A_i induces a lottery over Z_i at a cost of $c_i(A_i)$. Define

$$V_i(a, \theta_i, \tau_i) \equiv \sum_{z \in \{0,1\}^K} [f(z|a; \pi) \cdot W_i(z, \theta_i, \tau_i)] - \gamma^a \cdot |a|.$$

The expected utility associated with the choice $A_i = a$ is $V_i(a, \theta_i, \tau_i) + \psi_{ia}$, and the probability of choosing this portfolio is

$$Pr[A_i = a | X_i, D_i, Z_i, \theta_i, \tau_i] = \frac{\exp\left(\frac{V_i(a, \theta_i, \tau_i)}{\lambda_\psi}\right)}{\sum_{a' \in \{0,1\}^K} \exp\left(\frac{V_i(a', \theta_i, \tau_i)}{\lambda_\psi}\right)}$$

$$\equiv Q_{ia}(\theta_i, \tau_i).$$

Previous lottery-based studies of charter school effectiveness condition on students' application portfolios, which are typically referred to as "risk sets" because they determine the probability of a charter school offer (Abdulkadiroglu et al. 2011, Angrist et al. 2013, Dobbie and Fryer 2013). The probabilities $Q_{ia}(\theta_i, \tau_i)$ provide a model-based description of how students choose lottery risk sets.

4.3 Academic Achievement

Students are tested after application and attendance decision have been made. Potential academic achievement at school k is equal to expected achievement plus an idiosyncratic error: $Y_i(k) = \tilde{Y}_i(k) + \mu_{ik}$. To link the utility and achievement equations, I assume that the vector $(\theta_i, \eta_{ic}, \eta_{i0})'$ follows a multivariate normal distribution. This implies that potential academic achievement can be written

$$Y_i(k) = \alpha_k^0 + X_i'\alpha_c^x + \alpha_c^\theta \theta_i + \epsilon_{ik} \quad \text{for } k > 0, \tag{11}$$

$$Y_i(0) = \alpha_0^0 + X_i'\alpha_0^x + \alpha_0^\theta \theta_i + \epsilon_{i0}, \tag{12}$$

where $E[\varepsilon_{ik}|X_i, \theta_i] = 0$. The causal effect of attending charter k relative to traditional public school for student i is $Y_i(k) - Y_i(0)$. The observed score for student i is the potential score associated with her optimal school choice: $Y_i = Y_i(S_i)$. I assume that ϵ_{ik} follows a normal distribution with mean zero and variance σ_k^2, with σ_k^2 the same across charter schools but possibly different between charter and traditional public schools.

The coefficients α_c^θ and α_0^θ govern comparative and absolute advantage in charter school choice. If $\alpha_c^\theta > 0$, students with stronger tastes for charters have an absolute advantage in the charter sector, while these students have an absolute advantage in the traditional public sector if $\alpha_0^\theta > 0$. The difference $\alpha_c^\theta - \alpha_0^\theta$ determines whether sorting on unobservables is consistent with comparative advantage. If this difference is positive, students with larger potential gains prefer charter schools, and students sort based on comparative advantage. Choices are inconsistent with such sorting if the difference is negative.

4.4 Comments on Modeling Choices

Equations (8) through (12) provide a complete description of charter demand and potential academic achievement. This section provides intuition for some of the key modeling choices implicit in these equations.

First, the model emphasizes differences between charter and traditional public schools, while limiting differences between charter schools. Heterogeneity in preferences and achievement across students with different observed characteristics is governed by the vectors γ^x, α_c^x, and α_0^x. This specification allows observed characteristics to affect the choice of charter schools relative to traditional public schools, and to interact differently with achievement in charter and public schools, but requires that these characteristics affect preferences and achievement the same way at every charter. Similarly, equation (11) implies that the relationship between the unobserved taste θ_i and student achievement is the same at every charter school. Heterogeneity in preferences and achievement across charter schools is captured by the school specific intercepts γ_k^0 and α_k^0. These restrictions limit the number of parameters to be estimated while also parsimoniously summarizing heterogeneity across both students and schools. Moreover, this emphasis on differences between charters and traditional public schools mirrors the approach to identification described in Section 5.2, which emphasizes selection into the charter sector rather than across charter schools.

A second notable feature of the model is that potential achievement does not enter directly in students' utility functions. Instead, achievement and preferences are linked through the charter taste θ_i, which appears in the utility function (8) and the outcome equations (11) and (12). The inclusion of θ_i in the equations for academic achievement is a version of a selection correction approach to dealing with endogeneity (Heckman 1979; Dahl 2002). Selection into the charter sector is driven by θ_i, so comparisons of observed outcomes across schools yield causal effects after conditioning on θ_i. The coefficients α_c^θ and α_0^θ capture the influence of unobserved factors that are related to both tastes for charter schools and achievement.[15] The standard approach to selection correction involves computing the conditional expectation of the selection term and including it as a regressor in the outcome equation (Heckman 1979; Dubin and McFadden 1984). In the model estimated here, the expectation of θ_i conditional on a student's application and attendance choices does not have a closed form. Section 5.4 describes a simulation-based approach that allows me to estimate the parameters of the model despite the absence of a closed form for the selection correction.

5 Identification and Estimation

5.1 Exclusion Restriction

Identification of the parameters of equations (11) and (12) is based on the following exclusion restriction:

$$E[\epsilon_{ik}|X_i, Z_i, D_i, \theta_i, \upsilon_i, \psi_i] = 0. \tag{13}$$

Equation (13) embeds three identifying assumptions. First, the lottery offer vector Z_i is excluded from equations (11) and (12). This requires that offers have no direct affect on student achievement, a standard assumption in the charter lottery literature. Second, the school- and application-specific taste shocks υ_{ik} and ψ_{ia} are also excluded from these equations. Finally, distance to charter schools is excluded. I next discuss the latter two assumptions in detail and provide suggestive evidence in support of them.

5.2 Exclusion of School-Specific Preferences

In multinomial sample selection models, it is generally not feasible to allow potential outcomes to depend on preferences in an unrestricted way. Lee (1983), Dubin and McFadden (1984), and Dahl (2002) derive selection corrections under various restrictions on the selection process in static multinomial choice models. The model estimated here is dynamic, and the set of available application portfolios in the first stage is very large. To limit the dimensionality of the problem, I assume that the average charter taste θ_i is sufficient to capture the relationship between unobserved preferences and potential outcomes.

[15] Appendix A uses a simplified example to show that this model nests a standard Roy (1951) model in which preferences depend only on expected achievement.

The exclusion of v_i and ψ_i from equations (11) and (12) implies that selection on unobservables has a "single-index" form: endogenous selection into the charter sector depends only on students' overall tastes for charter schools relative to traditional public schools. This assumption is closely related to the assumptions underlying the marginal treatment effects (MTE) framework analyzed by Heckman et al. (2001), which characterizes average treatment effects conditional on the latent cost of seeking treatment. The analogue of this cost in the present context is the uniform random variable $\nu_i \equiv \Phi^{-1}(-\theta_i)$. The additional error terms v_{ik} and ψ_{ia} are necessary to capture the dynamic, multinomial nature of the charter choice process, but I assume that these errors are not systematically related to treatment effects. The difference $\alpha_c^{\theta} - \alpha_0^{\theta}$ governs the structure of marginal treatment effects.

The single-index restriction allows students to know about cross-site heterogeneity in the average effects of charter schools (captured by γ_k and α_k^0) and about their own suitability for the charter treatment in general (captured by θ_i), but it rules out the possibility that choices *between* charters are correlated with idiosyncratic treatment gains across schools. I present two pieces of evidence in favor of this assumption. First, I show that inputs and practices are highly homogeneous across Boston's charter middle schools. With homogeneous school practices, it seems less likely for a student to know that the environment at a specific school will lead to a particularly large benefit for him or her. Second, I show that application portfolio choices among charter applicants are determined mostly by distance. This suggests that there is not much heterogeneity in tastes for individual schools, which leaves little scope for matching on idiosyncratic achievement benefits.

Appendix Table A4 shows that practices are similar across Boston's charter middle schools. Columns (1) through (7) report responses to a survey on school practices for the seven charter middle schools included in the sample. For comparison, column (8) reports average responses for other charter middle schools in Massachusetts. Boston middle schools have more instructional time than other charter schools; five of seven have longer school years than the non-Boston average, and six of seven have longer school days. The seven Boston middle schools all strongly identify with the No Excuses educational approach, and emphasize the typical components of No Excuses, including traditional reading and math skills, discipline and comportment, and measurable results. With a few exceptions, Boston middle schools ask parents and students to sign commitment contracts, require students to wear uniforms, utilize formal merit/demerit systems to reward and punish student behavior, and use cold-calling and math and reading drills in the classroom. All of these practices are less common elsewhere in the state. Since educational practices are very similar across Boston's charter schools, it seems unlikely for a student to know that a specific school's practices will lead to an especially large idiosyncratic benefit.

To further motivate the exclusion of school-specific preferences, Appendix Table A5 shows that the choice of school conditional on applying is determined mostly by distance. In the model outlined above, the decision to choose one charter school over another is determined by the combination of distance and school-specific tastes. If application portfolio choices are dominated by distance, then there is no scope for selection on school-specific tastes, and the exclusion restriction requires only the exogeneity of distance. Fourty-one percent of applicants applied to the closest school, and these students traveled an average of 1.91 miles to their chosen schools. An additional twenty-two percent applied to the second closest charter, traveling an average of 1.12 miles beyond the closest school, and 16 percent choose the third closest, on average traveling 2.39 miles further than necessary. Less than ten percent of applicants chose the fourth-closest school, and the fractions who chose more distant schools are even smaller. A negligible fraction of applicants chose the most distant school. These facts show that although students are free to apply to distant schools, few do so; conditional on choosing to apply to a charter, most students apply to one close by, leaving little potential for matching on school-specific achievement gains.

5.3 Exclusion of Distance

In sample selection models, excluded instruments are necessary to identify potential outcome distributions without relying on functional form restrictions (Heckman 1990). Several sets of moments help to identify outcome distributions in the model used here. First, students who apply to more charter schools have stronger preferences for charters, so comparisons of outcomes between students who apply to different numbers of charters help to identify the selection correction parameters α_c^θ and α_0^θ. Second, I use distance and lottery instruments as exogenous shifters of charter tastes in the model's two choice stages (applications and attendance). Appendix B uses a simplified example to show how the combination of these instruments identifies the selection correction parameters. Intuitively, students who apply to charter schools from farther away are likely to have stronger unobserved tastes for charters, so a comparison of lottery-based estimates across students from different distances identifies the relationship between unobserved tastes and achievement gains. The maximum likelihood procedure implemented below efficiently combines information from all sources of identification. This section discusses potential threats to the validity of the distance instrument.

Distance is a valid instrument if it affects charter attendance and is uncorrelated with unobserved determinants of achievement. The use of this instrument parallels the use of proximity-based instruments in previous research on college and school choice (see, e.g., Card 1993 and Booker et al. 2011). The exclusion restriction requires that distance to charter schools is as good as random conditional on X_i. A sufficient condition for exclusion is that charter school leaders make choices between neighborhoods on the basis of averages of the characteristics in X_i. This seems plausible since X_i includes a rich set of student characteristics, including race, poverty, previous academic achievement, and a proxy for the academic quality of the surrounding public schools. These variables seem likely to capture many of the factors that might lead charter schools to locate in a particular area.

Columns (1) and (2) of Table 3 explore the validity of the distance instrument by examining the relationship between distance and baseline achievement. These columns report coefficients from ordinary least squares (OLS) regressions of 4th grade test scores on distance to the closest charter middle school, measured in miles. The estimates in the first row show that students who live farther from charter middle schools have significantly higher 4th grade test scores, suggesting that charter schools tend to systematically locate in lower-achieving areas of Boston. The second row shows that adding controls for observed characteristics shrinks these imbalances considerably and renders the math coefficient statistically insignificant. This suggests that observed demographic characteristics capture much of the relationship between charter locations and academic achievement. This lends plausibility to the use of distance as an instrument in models that control for these characteristics. The models estimated below also control directly for baseline test scores, which further limits the types of spatial correlation that might violate the distance exclusion restriction.

To directly compare the two sets of instruments used to estimate the model, columns (3) through (5) of Table 3 report 2SLS estimates using lottery offers and distance as instruments for charter attendance. The lottery estimates are replicated from Table 2. The distance models control for student characteristics and include the full sample. Column (3) shows that the distance instrument has a strong, statistically significant first stage effect on charter attendance: a one-mile increase in distance decreases the probability of charter attendance by 2.5 percentage points. Columns (4) and (5) show that the two instruments produce similar estimates of the effects of charter attendance, though the distance estimates are less precise. The distance instrument generates estimates of 0.60σ and 0.20σ in math and ELA, compared to the lottery estimates of 0.59σ and 0.41σ. While it is encouraging that these estimates are broadly similar, note that they needn't be; on average, the students induced to attend charter schools by the lottery and distance instruments may differ with respect to their observable characteristics X_i or their unobserved tastes θ_i.[16]

5.4 Estimation

I estimate the parameters of the model by maximum simulated likelihood (MSL). Let Ω denote the parameters of equations (8) through (12). The likelihood contribution of a student with outcome variables (A_i, Z_i, S_i, Y_i) can be written

$$\mathcal{L}_i(\Omega) = \int Q_{ia(i)}(\theta, \tau) \cdot f(Z_i|A_i; \pi) \cdot P_{is(i)}(Z_i, \theta, \tau)$$

$$\times \frac{1}{\sigma_{s(i)}} \phi \left(\frac{Y_i - \alpha^0_{s(i)} - X_i' \alpha^x_{s(i)} - \alpha^\theta_{s(i)} \theta}{\sigma_{s(i)}} \right) dF(\theta, \tau | X_i, D_i, \Omega). \quad (14)$$

Here the subscript $a(i)$ denotes the application bundle chosen by student i, while $s(i)$ denotes her school choice.[17]

I evaluate the integral in equation (14) by simulation. Let θ_i^r and τ_i^r be draws of θ and τ for student i in simulation r. Define

$$\hat{\ell}_i^r(\Omega) \equiv Q_{ia(i)}(\theta_i^r, \tau_i^r) \cdot f(Z_i|A_i; \pi) \cdot P_{is(i)}(Z_i, \theta_i^r, \tau_i^r) \cdot \frac{1}{\sigma_{s(i)}} \phi \left(\frac{Y_i - \alpha^0_{s(i)} - X_i' \alpha^x_{s(i)} - \alpha^\theta_{s(i)} \theta_i^r}{\sigma_{s(i)}} \right).$$

The simulated likelihood for observation i is

$$\hat{\mathcal{L}}_i(\Omega) = \frac{1}{R} \sum_{r=1}^{R} \hat{\ell}_i^r(\Omega),$$

where R is the number of draws. The MSL estimator is defined by

$$\hat{\Omega}_{MSL} = \arg\max_{\Omega} \sum_{i=1}^{N} \log \hat{\mathcal{L}}_i(\Omega).$$

If R rises faster than $\sqrt{N}$, the MSL estimator is $\sqrt{N}$-consistent and has the same asymptotic distribution as the conventional maximum likelihood estimator (Train 2003). I use 300 draws of θ_i and τ_i for each student, and

[16] It is also not necessary for lottery and distance compliers to *differ* with respect to average X_i or θ_i. Appendix B makes clear that identification of selection on unobservables is based on the *interaction* of the two instruments, not differences in average tastes across the two sets of compliers when the instruments are used separately.

[17] $s(i)$ is used to refer both to the specific school chosen by student i, as in the school-specific intercept $\alpha^0_{s(i)}$, and to the type of school chosen by student i (charter or public), as in the demographic coefficient vector $\alpha^x_{s(i)}$.

maximize the simulated likelihood using the Newton-Rhapson method, with the gradient of the objective function calculated analytically. The results were not sensitive to increasing the number of draws beyond around 100. I calculated standard errors using the average outer product of the score of the simulated likelihood.

6 Structural Estimates

I next discuss estimates of the key parameters of the structural model, which are reported in table 4, 5 and 6. These results were generated using 8th grade test scores as outcomes and assuming a bivariate normal distribution for the ϵ_{ik} across subjects. Estimates for 6th and 7th grade were very similar and are available upon request.[18]

6.1 Preference Parameters

Table 4 shows MSL estimates of the parameters governing preferences for charter schools. Column (1) reports estimates of the utility parameters, while column (2) reports standard errors. Column (3) shows average marginal effects of observed characteristics on the probability of applying to at least one charter school.[19] The covariate vector X_i is de-meaned in the estimation sample, so the intercept (computed as the average of γ_k^0 across schools) is the average utility of charter attendance. The estimated intercept is negative and statistically significant, which implies that on average, students prefer traditional public schools to charter schools even in the absence of application and distance costs.

Estimates of the vector γ^x are consistent with the demographic patterns reported in Table 1. Subsidized lunch status, special education, and limited English proficiency are associated with weak demand for charter schools, while black students and students with higher baseline math and ELA scores have stronger preferences for charters than other students. Preferences for charters are weaker among students with higher-quality local public school options as measured by value-added, though this coefficient is imprecisely estimated. This lack of precision is likely driven by the relatively small variance of the value-added measure: local public school value-added ranges from -0.12σ to 0.18σ in the estimation sample, with a standard deviation of only 0.05σ. As in Table 1, the estimates show that poverty status has a substantial effect on application behavior. Holding other variables constant, subsidized lunch status reduces the probability of submitting a charter application by 7 percentage points.

The bottom half of Table 4 reports estimates of the parameters governing preferences for distance, application costs, and heterogeneity in unobserved tastes. Increased distance significantly reduces the utility of charter school attendance. The marginal effect in column (3) shows that a one-mile increase in distance to a particular charter reduces the probability of applying to that school by 0.6 percentage points, which is large relative to mean application rates at individual schools (2 to 3 percent). The estimate of the application cost γ^a is positive, large, and statistically significant. Its magnitude suggests that applying to a charter school involves a utility cost equivalent to a 5.6-mile increase in distance.

The estimates also reveal important unobserved heterogeneity in preferences for charter schools. In utility

[18]The estimates for other grades were also reported in an earlier draft of this article (Walters 2013).

[19]Marginal effects for discrete variables are computed by simulating the model first with the relevant characteristic set to zero for each student and then with it set to one, and computing the average difference in application probabilities across these simulations. Marginal effects for continuous variables are average simulated numerical derivatives of the application probability. The marginal effect for distance is the average effect of a one-mile increase in distance to a particular school on the probability of applying to that school.

terms, a one-standard-deviation increase in θ_i is equivalent to a 13-mile increase in distance to all charter schools. The equivalent estimates for ξ_{ik}, τ_{ik}, and ψ_{ia} are smaller (7.4 miles, 0.34 miles, and 1.5 miles).[20]The preference estimates therefore suggest that there is more unobserved heterogeneity in tastes for charter schools as a whole than for individual charters or application bundles. This is further evidence that a single-index selection model is reasonable.

6.2 Achievement Parameters

Table 5 reports estimates of the parameters of 8th-grade potential achievement distributions. In each panel, column (1) shows estimates for charter schools, column (3) shows estimates for public schools, and column (5) shows the difference, which is the causal effect of charter attendance. Columns (2), (4), and (6) report standard errors. The reported charter intercept is the mean of the school-specific intercepts α_k^0, weighted by the enrollment shares of each school. The intercept in column (5) can therefore be interpreted as the population average treatment effect (ATE) of charter attendance.

The estimates in Table 5 reveal that charters have larger effects on test scores for more disadvantaged students. The constant term implies that charter attendance raises 8th-grade math and ELA scores by 0.64σ and 0.56σ on average. Subsidized lunch students, non-white students, and students with lower baseline scores receive further benefits. A comparison of columns (1) and (3) reveals that black students, Hispanics, and poor students lag behind other students in public school, but these characteristics are not predictive of potential scores in charter schools conditional on the other covariates. In this sense, charter schools close achievement gaps between racial and socioeconomic groups.

Estimates of the selection correction parameters α_c^θ and α_0^θ reveal that stronger unobserved preferences for charters are associated with slightly *smaller* achievement benefits from charter attendance. Column (3) shows that students with stronger preferences for charters do better in traditional public schools. A one-standard-deviation increase in θ_i is associated with a 0.1σ increase in public school math scores and a 0.027σ increase in ELA scores. Similar to the pattern for observed characteristics, the relationship between θ_i and achievement is weaker in charter schools. Students with stronger unobserved preferences therefore experience smaller benefits from charter attendance: when θ_i increases by one-standard deviation, the achievement benefits generated by charter attendance fall by 0.09σ and 0.06σ in math and ELA.

6.3 Absolute and Comparative Advantage in Charter School Choice

Taken together, the structural preference and achievement estimates can be used to characterize selection into the charter sector on both observed and unobserved dimensions. To summarize patterns of absolute and comparative advantage, define the preference index

$$\mathcal{P}_{\mathrm{i}} \equiv X_i'\gamma^x + \theta_i.$$

[20]The standard deviation of τ_{ik} is $\lambda_\tau \cdot \left(\pi/\sqrt{6}\right)$, and similarly for the other extreme-value errors.

$\mathcal{P}_i$ indexes student i's preference for charter schools relative to public schools as a function of both observed characteristics and unobserved tastes. The relationship between charter preferences and potential charter achievement is summarized by the function

$$\alpha_c(p) \equiv \sum_{k=1}^{K} w_k \cdot E[Y_i(k)|\mathcal{P}_i = p].$$

$\alpha_c(p)$ is the average potential charter school outcome for students with preference p. I set the weights w_k proportional to charter enrollment shares. Similarly, $\alpha_0(p) \equiv E[Y_i(0)|\mathcal{P}_i = p]$ is the average potential public school outcome for students with preference p. The average achievement benefit generated by charter attendance is $\beta(p) \equiv \alpha_c(p) - \alpha_0(p)$.

The key finding in the paper is shown in Figure 3. This figure reveals that students with stronger demand for charter schools have absolute advantages in both the charter sector and the traditional public sector, but students with weaker preferences have a comparative advantage in the charter sector. Panels A and B plot the conditional expectation functions $\alpha_c(p)$ and $\alpha_0(p)$ for math and ELA. In both subjects, mean potential outcomes rise with charter preferences. This is driven both by observed characteristics (since disadvantaged students have weaker tastes for charters and lower scores) and unobserved characteristics (since high-θ_i students have stronger tastes for charters and higher scores). The slope of $\alpha_c(p)$ is less steep than the slope of $\alpha_0(p)$, however, so the benefit associated with charter attendance falls as charter preference rise. This can be seen clearly in panel C, which plots $\beta(p)$. Consistent with the reduced form evidence in Figure 1, the structural estimates reveal that achievement benefits are smaller for students with stronger demand for charter schools.

Sorting on the basis of comparative advantage requires that $\beta(p)$ rises with p; instead, the results reported here show that the students with the largest potential benefits are precisely those who are least likely to enter the charter sector. This inverse relationship between preferences and achievement gains is striking in view of the standard Roy model, which predicts a positive relationship between the propensity to seek treatment and the benefit from doing so. While it is not possible to conclusively determine the cause of the "reverse Roy" pattern documented in Figure 3, one plausible explanation is worth noting. Parents who invest more in human capital on other margins may also be more motivated to enroll their children in charter schools. Charter schools weaken the relationships between student characteristics and academic achievement, however, which suggests that they partially compensate for differences in human capital investments across families. In this scenario, children with more motivated parents will have absolute advantages in both sectors and will be more likely to enroll in charters, but will experience smaller gains from charter attendance. This description matches the patterns of absolute and comparative advantage documented in Figure 3.

6.4 School Effects

Table 6 reports estimates of the model's school-specific parameters, including the average utilities γ_k^0, the admission probabilities π_k (averaged across applicant cohorts), and the average test score effects $(\alpha_k^0 - \alpha_0^0)$. The utility estimates show that some charters are more popular than others, but all of the estimates are negative, indicating that on average students prefer traditional public schools to attending any charter. The admission probabilities range from 0.39 to 0.88. The achievement estimates in columns (3) and (4) show that the large effects of Boston's charters are not driven by any particular school: all seven schools boost achievement in both math and ELA. Interestingly, the most effective schools do not seem to be the most popular; schools 4, 6, and 7 have the largest test score effects, but also the three lowest average utilities. This is further evidence that school choices are not driven by achievement gains in the Boston charter setting.

7 Policy Implications

The structural estimates reported in Section 6 reveal that students with larger potential achievement benefits are less likely to apply to charter schools. The preference estimates in Table 4 also imply that demand for charters is relatively weak: students act as if charter applications are costly, and the average utility of charter attendance is below the utility associated with traditional public school. I next explore the policy implications of these findings. First, to quantify how much the pattern of selection matters for productivity in the charter secotr, I use the structural model to compare treatment effects for current charter students to potential treatment effects for other groups of students, holding the size of the charter sector fixed. I then combine the model of charter demand with assumptions about supply to quantify the potential effects of charter expansion at the margin.

7.1 Charter School Treatment Effects

Table 7 reports three key treatment effects computed using the structural estimates. Column (1) shows the effect of treatment on the treated (TOT), the effect of charter attendance for students who attend charter schools in the current system. This effect can be written $TOT \equiv E[\beta(\mathcal{P}_i)|S_i \neq 0]$. Column (2) shows the population average treatment effect (ATE), given by $ATE \equiv E[\beta(\mathcal{P}_i)]$. Column (3) shows the effect of treatment on the non-treated (TNT), defined as $TNT \equiv E[\beta(\mathcal{P}_i)|S_i = 0]$. Estimates of these effects are generated by simulating observations from the structural model and computing average treatment effects for students who make each treatment choice.

The results in Table 7 show that the pattern of selection into charter schools produces meaningful differences in treatment effects between charter and non-charter students. The ATE for 8th-grade math scores is 0.64σ, while the TOT is 0.48σ, roughly 25 percent smaller. The TNT is 0.68σ. This implies that replacing a randomly selected charter student with a randomly selected non-charter student would increase average math achievement in the pair by 0.1σ $((0.68\sigma - 0.48\sigma)/2)$. Similarly, the TOT for ELA is 25 percent smaller than the ATE (0.42σ compared to 0.56σ). These results suggest that outreach programs targeting students who are currently unlikely to apply to charter schools could substantially boost overall productivity in the charter sector.

One potential caveat to the interpretation of the results is worth noting. The treatment effects in Table 7 capture average causal effects of charter attendance for randomly selected students in the charter sector (TOT), non-charter sector (TNT), or in the population (ATE), holding all inputs in the charter and non-charter sectors fixed. If part of the effect of charter attendance operates through peer quality, these treatment effects do not reflect the impact of changing the overall composition of the charter student body. Specifically, any peer effects are built into the parameters α_k^0, so these parameters will change as peer quality changes if peer effects are important.

For two reasons, however, peer effects seem unlikely to an important part of the achievement effects generated by charter schools. First, Table 1 shows that charter applicants' baseline test scores are about 0.2σ above the Boston average. For this difference in peer quality to explain the causal effects reported in Table 2, peer effects would have to be roughly 2σ to 3σ per standard deviation of peer quality. In a summary of the peer effects literature, Sacerdote (2011) reports a wide range of estimates, almost all of which are substantially less than these magnitudes. This suggests that charter schools produce achievement gains mostly through channels other than peer effects. Second, students who experience larger increases in peer quality as a result of charter attendance do not experience larger achievement gains. Appendix Figure A1 plots lottery-specific reduced form effects on peer quality against effects on test scores. Peer quality is defined as the average 4th-grade test score of the peers with whom a student attends 6th grade. The figure shows that there is essentially no relationship between the effect of winning a charter lottery on peer quality and the effect on academic achievement. This suggests that peer effects are not a primary driver of charter effectiveness, so the α_k^0 likely reflect factors other than peer quality.[21]

7.2 Charter Expansion Effects

7.2.1 Additional Assumptions

The treatment effects discussed above describe the effects of modifying the student population served by Boston's charter system in various ways (such as replacing charter students with current non-charter students, for example). An alternative approach to quantifying the policy importance of demand-side behavior is to use estimates of charter demand to predict the efficacy of charter school expansion. Predicting charter expansion effects requires assumptions about the supply side of the charter market. A full model of charter supply is outside the scope of this paper; I instead make simplifying assumptions that allow me to interpret out-of-sample predictions as plausible upper bounds on the effects of charter expansion due to demand-side behavior.

The supply side of the charter market is defined by a set of charter schools, with each school characterized by a location, an admission probability π_k, an average utility γ_k^0, and a mean achievement parameter α_k^0. To choose locations for the first six expansion schools, I use the addresses of new campuses that opened through 2013 (see Appendix Table A1). For further expansions, I choose locations using predictions from a probit model of the probability that a charter is located within a zip code as a function of average share non-white, share subsidized lunch, and average baseline MCAS scores in the zip code.[22] Each expansion school is placed sequentially in the center of the zip code with the highest predicted probability among those that do not already contain a charter.[23]

Charter admission probabilities are assumed to adjust endogenously to equate the demand for charter enrollment among admitted students with the supply of charter seats. I take charter school seating capacities as exogenously given, and solve for a Subgame Perfect Nash Equilibrium in which charters optimally set admission probabilities to maximize enrollment subject to their capacity constraints. Capacities for new schools are chosen randomly from the distribution of capacities for existing schools. Appendix C describes the details of the equilibrium and the methods used to compute counterfactual admission probabilities.

To choose the average utility and test score parameters, I assume that the charter school sector exhibits constant

[21] Angrist et al. (2013a) present additional evidence that the effects of Boston's charter high schools do not seem to be driven by peer quality.

[22] Marginal effects from the probit model are reported in Appendix Table A6.

[23] Anecdotally, location decisions for Boston charter schools are often determined by the availability of vacant buildings, such as empty churches (Roy 2010). In an alternative set of simulations, I chose charter locations randomly from a grid of half-mile by half-mile blocks covering Boston. The results were qualitatively similar to those reported below.

returns to scale (CRTS). Specifically, I treat the vectors $(\gamma_k^0, \alpha_k^0)'$ as independent and identically distributed draws from a fixed distribution $F(\gamma^0, \alpha^0)$. Each new school is assigned a draw from the estimated joint distribution of parameters listed in Table 6. There are several reasons that this assumption may fail to hold in practice. If teachers, principals, or other inputs are supplied inelastically, it may be difficult for new charters to replicate the production technology used by existing campuses (Wilson 2008). Public schools may also respond to charter competition, though existing evidence suggests that the effects of charter entry on traditional public school students are small (Imberman 2011). If peer effects do play a role in charter effectiveness, these effects will be diluted in expansions that draw in less positively selected students. Together, these factors seem likely to reduce the efficacy of charter schools at larger scales. The simulation results should therefore be viewed as upper bounds on the effects of charter expansion. The CRTS assumption allows me to describe demand-driven limits on the effects of expansion in a best-case scenario for charter supply.

7.2.2 Simulation Results

Figure 4 summarizes the counterfactual simulations. The outcomes of interest are school choices, charter oversubscription, average 8th-grade test scores, and charter school treatment effects. In each panel, a vertical black line indicates the existing number of charter schools, and a red line indicates the size of Boston's planned charter expansion. Panel A of Figure 4 shows how charter application and attendance rates change as the charter sector expands. Panel B shows effects on admission probabilities and seating capacity utilization. Panel C shows effects on average math and ELA scores. Panel D shows the effect of treatment on the treated in each simulation. Table 8 shows numerical results for choice behavior and test scores in a subset of the simulations.

The simulations imply that charter schools have had a significant impact on the distribution of test scores in Boston. This can be seen in the second row of Table 8, which shows the simulated effects of closing all charter schools. Without charter schools, the gap in average test scores between Boston and the rest of Massachusetts would widen by 0.043σ in math and 0.033σ in ELA. This implies that the existence of charter schools has reduced average test score gaps between Boston and the rest of the state by 12 percent in math and 6 percent in ELA.

The next row of Table 8 shows that the recently authorized charter expansion is predicted to further raise average math and ELA scores by 8 percent and 6 percent, respectively. As shown in Figure 4, opening additional charters is predicted to continue to boost average test scores. Furthermore, panel D of Figure 4 shows that with constant returns to scale on the supply side, the efficacy of charter schools will *increase* as the charter sector expands. The TOTs associated with Boston's planned expansion are 4 percent and 7 percent larger than the TOT for current charter students in math and ELA. TOTs for an expansion that raises the number of charter schools to 20 are 7 percent and 12 percent larger than the current TOT. This reflects the pattern of selection discussed in Section 6: at the margin, charter expansion draws in students with weaker tastes for charter schools, who receive larger achievement gains. As a result, charter school expansion has the potential to produce large gains for marginal applicants.

However, the simulation results also imply that demand for charter schools in Boston is limited, especially among the students with the largest potential achievement gains. Panel B of Figure 4 shows that charter expansion is predicted to reduce oversubscription: admission probabilities rise quickly with the number of schools, and the share of seats left empty also increases. In a setting with 20 charter schools, 87 percent of charter applicants are admitted, so a student who wishes to attend a charter is almost guaranteed the opportunity to do so. Nevertheless, less than half of students apply to a charter, 21 percent attend one, and 17 percent of charter seats are empty. Moreover, the ATEs in Table 7 are well above the TOTs in all simulations, suggesting that many of the students for whom charters are most effective choose to remain in public schools. This finding is driven by the large application cost and negative average utilities reported in tables 4 through 6, together with comparatively weak preferences for charters among disadvantaged and low-achieving students. The simulation results suggest that without efforts to induce charter applications from the highest-benefit students, the effects of realistic charter expansions may be limited by weak demand.

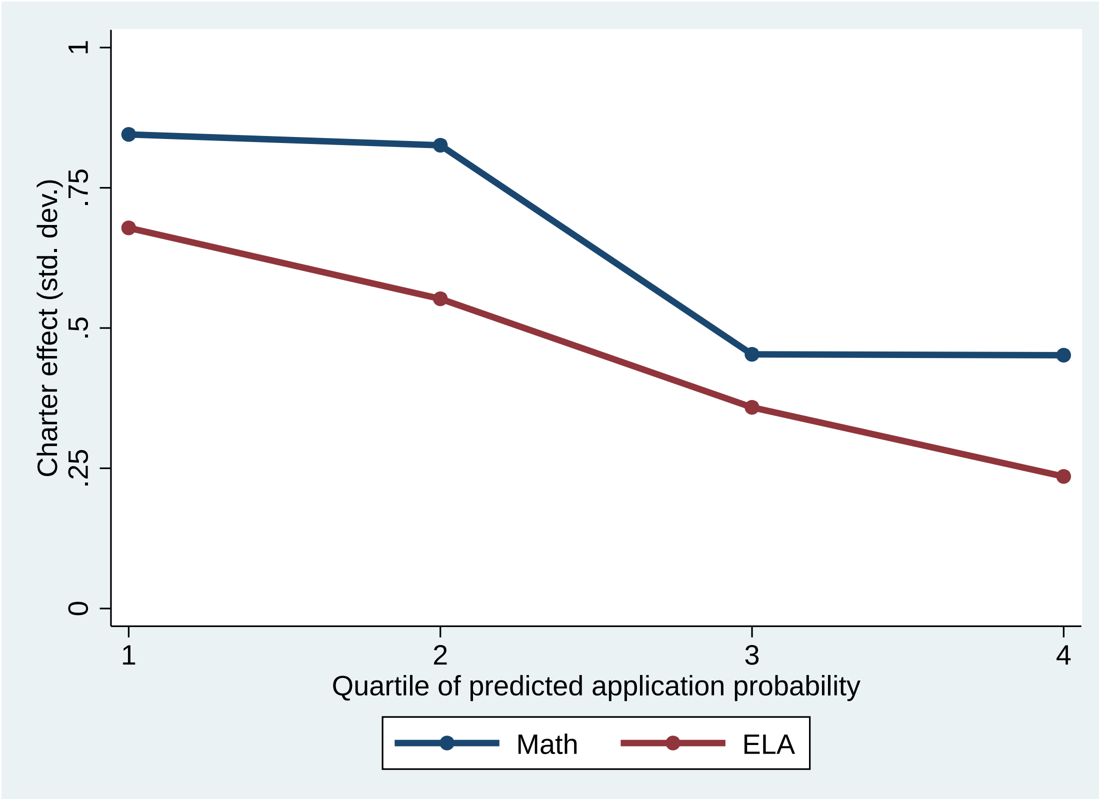
8 Conclusion

This paper develops a structural model of charter school applications, attendance decisions, and academic achievement to analyze patterns of absolute and comparative advantage in school choice. Estimates of the model reveal that tastes for charter schools are inversely related to achievement gains: low-achievers, poor students, and those with weak unobserved tastes for charters gain the most from charter attendance, but are unlikely to apply. Charter school choices are therefore inconsistent with sorting based on comparative advantage. As a consequence, counterfactual simulations show that charter effectiveness is increasing in the size of the charter sector, as expansions draw in students with weaker preferences who receive larger gains. At the same time, demand for charters among the highest-benefit students is weak, so the effects of charter expansion may be limited by weak demand even in the best-case scenario for charter supply.

This pattern is surprising – the canonical Roy (1951) selection model predicts that students with more to gain from charter attendance will be more likely to apply. However, the "reverse Roy" pattern described here is consistent with the possibility that effective charter schools compensate for differences in human capital investments across families, but parents who invest more on other dimensions are more likely to enroll their children in charter schools. This pattern is also consistent with a growing body of evidence suggesting that lower-income students are less likely to choose high-quality schools in a variety of settings (Buter et al., 2013; Brand and Xie, 2010; Dillon and Smith, 2013; Hastings et al., 2009; Hoxby and Avery, 2012). This constellation of findings has broader implications for the design of school-choice programs. An increase in the availability of high-quality schools without commensurate outreach efforts may not induce disadvantaged students to attend these schools, even if the benefits from doing so are especially large for such students.

These findings raise the further question of whether parents who forgo large potential achievement gains are truly uninterested in achievement, or simply unaware of differences in effectiveness across schools. The model estimated in this paper does not distinguish between these two possibilities. If the lack of demand for charter schools among disadvantaged students reflects a lack of information, the demand for charters may shift as parents become more informed. In related work, Hastings and Weinstein (2008) show that providing test score information leads parents to choose schools with higher test scores, suggesting that informational frictions may play a role. Changes in recruitment practices may also change the pattern of selection into charter schools; recent legislation authorizing charter expansion in Massachusetts requires schools to take efforts to recruit applicants who are demographically similar to students in the local district. In future work, I plan to use data from Boston's expansion to validate the model estimated here, and to study changes in the demand for charter schools as the city's charter sector expands.

Figure 1: Relationship Between Charter Effects and Predicted Application Probabilities



P-values for equal effects across quartiles: Math = 0.038, ELA = 0.089

Notes: This figure plots coefficients from 2SLS regressions of test scores on interactions of charter attendance and dummies for quartiles of the predicted probability of applying to a charter school, instrumenting with interactions of the lottery offer and quartile dummies. Predicted application probabilities are computed in the full sample using the probit model in column (3) of Table 1. The sample is restricted to lottery applicants before quartiles are defined. Models also control for quartile main effects, lottery fixed effects, and grade effects. The sample stacks scores in grades six through eight.

Figure 2: Sequence of Events

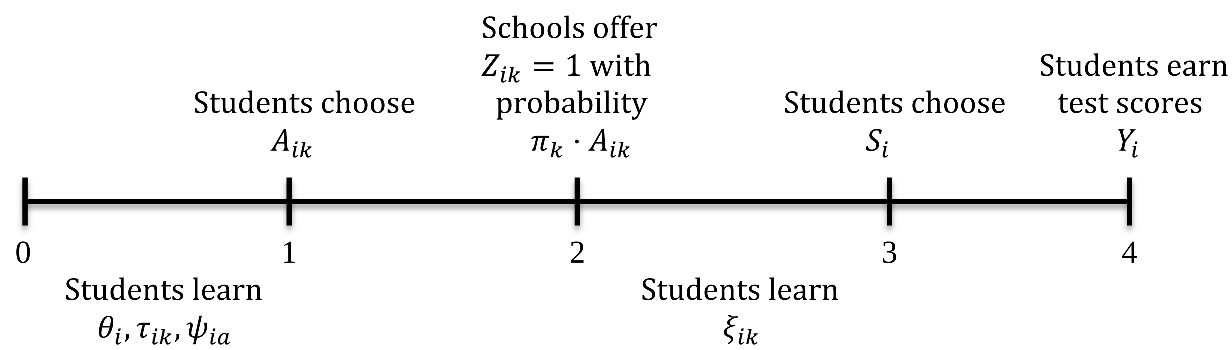
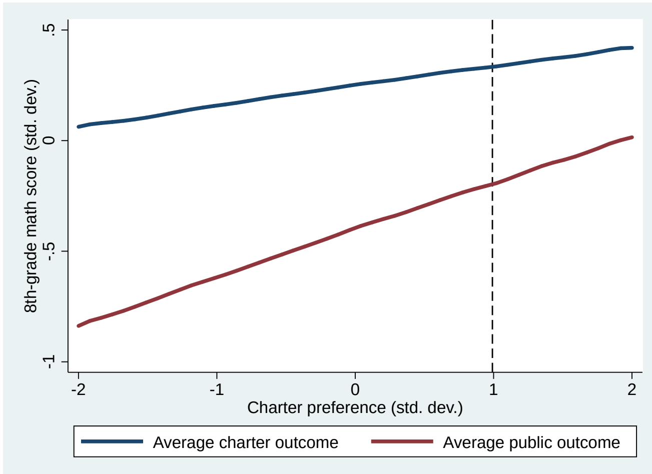
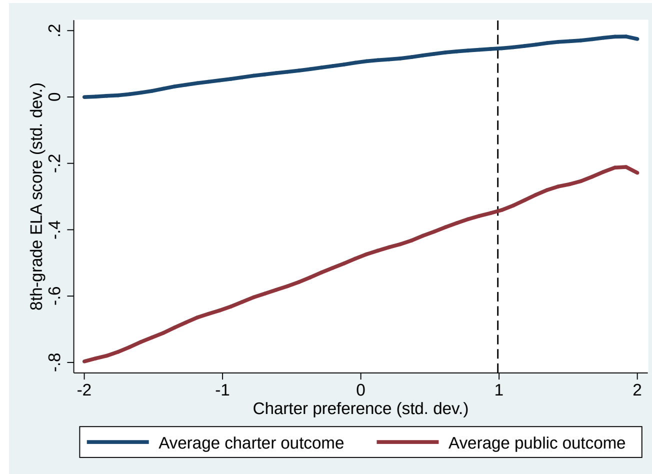


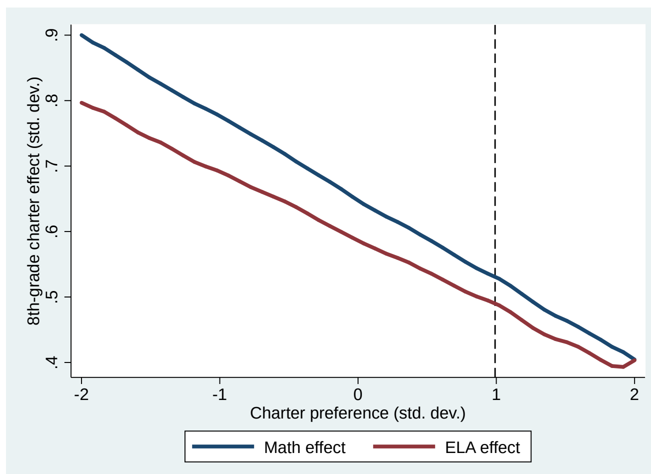
Figure 3: Absolute and Comparative Advantage in Charter School Choice



A. Math potential outcomes



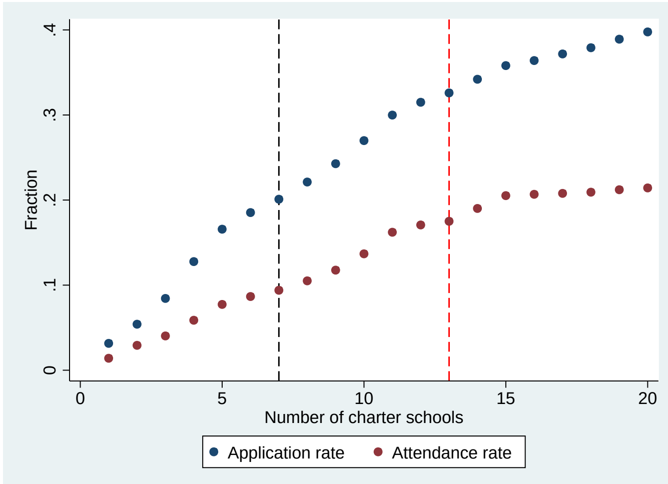
B. ELA potential outcomes



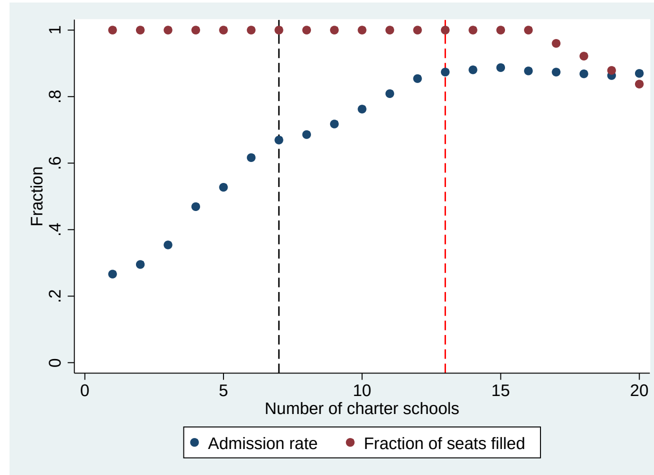
C. Charter effects

Notes: This figure plots relationships between preferences for charter schools and potential 8th-grade test scores in charter and traditional public schools. The figure is produced by simulating 100,000 observations from the structural model using the empirical distribution of observed student characteristics. A charter preference, an average charter potential outcome, and a public potential outcome are then computed for each observation. Charter outcomes are weighted averages of school-specific outcomes, with weights proportional to school capacity. Panels A and B plot coefficients from local linear regressions of potential outcomes on charter preferences in this simulated data. Panel C is a corresponding plot for the causal effect of charter attendance, which is the difference between charter and public potential outcomes. Local linear regressions use triangle kernels with bandwidths of 0.1. Charter preferences are in standard deviation units. Dashed lines show the average preference for current applicants.

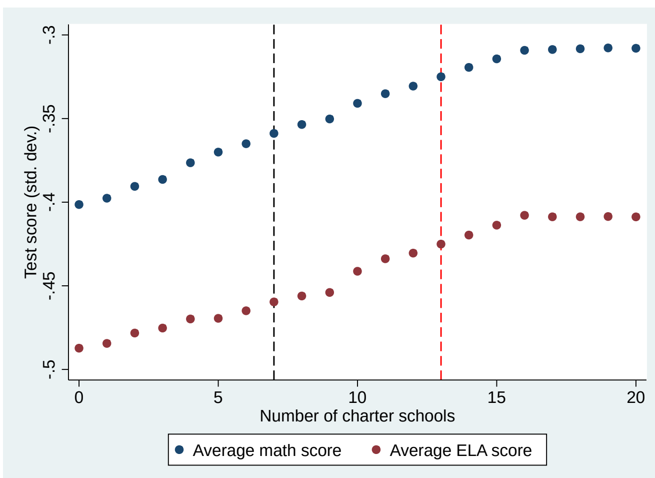
Figure 4: Simulated Effects of Charter School Expansion



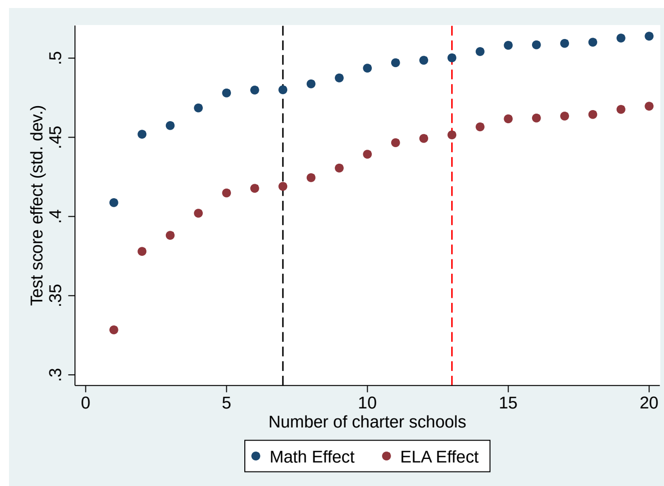
A. Applications and attendance



B. Oversubscription



C. Average test scores in Boston



D. Effects of treatment on the treated

Notes: This figure displays simulated effects of charter school expansion. The black dashed line in each panel corresponds to the existing number of charter schools, while the red dashed line corresponds to Boston's planned expansion. Simulated statistics are produced using 100 simulations per observation in the data set.

Table 1: Descriptive Statistics

	Means		
	All students	Charter applicants	Probit marginal effect
	(1)	(2)	(3)
Applied to charter school	0.177	-	-
Applied to more than one charter school	0.048	0.272	-
Received charter offer	0.126	0.709	-
Attended charter school	0.112	0.593	-
Female	0.483	0.480	-0.008
			(0.008)
White	0.147	0.166	-0.037***
			(0.011)
Hispanic	0.394	0.310	-0.046***
			(0.009)
Subsidized lunch	0.818	0.721	-0.094***
			(0.012)
Special education	0.231	0.178	-0.026***
			(0.010)
Limited English proficiency	0.210	0.132	-0.048***
			(0.010)
Miles to closest charter school	2.110	1.889	-0.033***
			(0.003)
Value-added of public schools in zip code	-0.051	-0.050	-0.120
			(0.087)
4th grade math score	-0.545	-0.347	0.014***
			(0.005)
4th grade ELA score	-0.660	-0.433	0.010*
			(0.005)
N	10122	1794	10122

Notes: This table shows descriptive statistics for students attending 4th grade at traditional public schools in Boston between 2006 and 2009. The sample excludes students with missing middle school test scores. Column (1) shows means for the full sample, while column (2) shows means for charter applicants. Column (3) shows marginal effects from a probit model where the dependent variable is an indicator for applying to any charter. Marginal effects are evaluated at the sample mean of all variables. Robust standard errors are in parentheses. *significant at 10; **significant at 5%; ***significant at 1%

Table 2: Lottery-based Estimates of Charter School Effects

	Pooled sample		2SLS estimates for subgroups					
			Subsidized lunch		Below-median baseline score		Non-white	
	First stage	2SLS	Yes	No	Yes	No	Yes	No
Subject	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	0.677***	0.587***	0.673***	0.390***	0.847***	0.468***	0.681***	0.223
	(0.021)	(0.072)	(0.085)	(0.125)	(0.088)	(0.070)	(0.081)	(0.141)
p-value			0.060		0.001		0.006	
ELA	0.677***	0.407***	0.471***	0.264**	0.652***	0.217***	0.507***	0.021
	(0.021)	(0.074)	(0.087)	(0.122)	(0.095)	(0.065)	(0.084)	(0.143)
p-value			0.161		0.000		0.004	
N (scores)	5108		3699	1409	2537	2571	4268	840
N (students)	1794		1294	500	894	900	1497	297

Notes: This table reports 2SLS estimates of the effects of attendance at Boston charter schools on test scores for lottery applicants. The sample stacks test scores in grades 6 through 8. The endogenous variable is a dummy for attending any charter school after the lottery and prior to the test. The instrument is a dummy for receiving a lottery offer from any charter school. Column (1) reports coefficients from regressions of charter attendance on the offer variable in the sample of all applicants. Column (2) reports corresponding 2SLS estimates for math and ELA scores. Columns (3)-(8) show 2SLS coefficients from models that interact the endogenous variable with the characteristic in the column heading, adding an interaction of this characteristic and the lottery offer as an instrument. *P*-values are from tests of the hypothesis that the coefficients for both groups are the same. Sample sizes in columns (7) and (8) are for math. All models control for lottery fixed effects and grade effects. Standard errors are robust to heteroskedasticity and are clustered at the student level.
*significant at 10; **significant at 5%; ***significant at 1%

Table 3: The Distance Instrument

	Balance check: 4th grade scores		2SLS comparison			
	Math	ELA		First stage	Math 2SLS	ELA 2SLS
Controls	(1)	(2)	Instrument	(3)	(4)	(5)
None	0.041***	0.053***	Lottery	0.677***	0.587***	0.407***
	(0.009)	(0.010)		(0.021)	(0.072)	(0.074)
N (scores)					5108	
N (students)	10122				1794	
Baseline characteristics	0.012	0.021***	Distance	-0.025***	0.600***	0.203
	(0.008)	(0.008)		(0.003)	(0.204)	(0.198)
N (scores)					29009	
N (students)	10122				10122	

Notes: Columns (1) and (2) show regressions of 4th-grade test scores on miles to the closest charter middle school. The first row includes no controls, while the second controls for student characteristics, including sex, race, free lunch status, special education status, limited English proficiency, and value-added of public schools in the zip code. Columns (3) through (5) show 2SLS results for middle school test scores using the lottery and distance instruments. The lottery estimates are reproduced from Table 2. The distance models control for student characteristics and 4th grade test scores. Standard errors are clustered at the student level.

*significant at 10; **significant at 5%; ***significant at 1%

Table 4: Estimates of Utility Parameters

Parameter	Description	Estimate (1)	Standard error (2)	Marginal effect (3)
γ^0	Mean charter utility	-1.306***	0.132	-
γ^x	Female	-0.027	0.080	-0.001
	Black	0.307***	0.137	0.027
	Hispanic	0.045	0.139	0.005
	Subsidized lunch	-0.740***	0.111	-0.071
	Special education	-0.329***	0.115	-0.027
	Limited English proficiency	-0.390***	0.113	-0.032
	Baseline math score	0.168***	0.060	0.014
	Baseline ELA score	0.111*	0.060	0.008
	Value-added of public schools in zip code	-0.839	0.909	-0.076
γ^d	Distance	-0.174***	0.009	-0.006
γ^a	Application cost	0.978***	0.040	-
σ_0	Standard deviation of charter school tastes	2.235***	0.129	-
λ_τ	Scale of school-specific tastes	0.046	0.039	-
λ_ψ	Scale of application-specific tastes	0.205***	0.010	-
N	Sample Size		9156	

Notes: This table reports maximum simulated likelihood estimates of the parameters of the structural school choice model. The sample includes all students with observed 8th-grade test scores. The likelihood is evaluated using 300 simulations per observation. Column (1) reports parameter estimates, while column (2) reports standard errors. The constant is the average of school-specific mean utilities, evaluated at the sample mean of the covariate vector X. Column (3) reports average marginal effects of observed characteristics on the probability of applying to at least one charter school. Marginal effects for discrete variables are differences between average simulated application probabilities with the relevant characteristic set to 1 and 0 for all observations. Marginal effects for continuous variables are average simulated numerical derivatives of the application probability. Marginal effects are evaluated using 100 simulations per observation. The marginal effect for distance is the effect of a one-mile increase in distance to a school on the probability of applying to that school, averaged across schools.

*significant at 10; **significant at 5%; ***significant at 1%

Table 5: Estimates of Achievement Parameters

Parameter	Description	Charter school Estimate (1)	Charter school Standard error (2)	Traditional public school Estimate (3)	Traditional public school Standard error (4)	Charter effect Estimate (5)	Charter effect Standard error (6)
		Panel A. Math					
α^0_m	Mean potential outcome	0.272***	0.080	-0.367***	0.009	0.639***	0.081
α^x_m	Female	0.032	0.040	0.025	0.016	0.057	0.043
	Black	0.048	0.066	-0.166***	0.026	0.214***	0.071
	Hispanic	0.135***	0.068	-0.093***	0.026	0.228***	0.073
	Subsidized lunch	0.046	0.049	-0.120***	0.023	0.165***	0.054
	Special education	-0.288***	0.052	-0.374***	0.019	0.086	0.055
	Limited English proficiency	-0.025	0.065	0.076***	0.020	-0.101	0.068
	Baseline math score	0.358***	0.030	0.476***	0.011	-0.118***	0.032
	Baseline ELA score	0.046	0.029	0.065***	0.01	-0.019	0.030
	Value-added of public schools in zip code	0.479	0.430	0.943***	0.171	-0.464	0.462
$\alpha^\theta_m \times \sigma_\theta$	Taste for charter schools (std. dv. units)	0.010	0.023	0.100***	0.012	-0.090***	0.026
		Panel B. ELA					
α^0_e	Mean potential outcome	0.138	0.088	-0.424***	0.010	0.562***	0.089
α^x_e	Female	0.164***	0.043	0.183***	0.016	-0.019	0.046
	Black	0.101	0.068	0.058**	0.028	0.159***	0.074
	Hispanic	0.177**	0.070	-0.031	0.028	0.209***	0.076
	Subsidized lunch	0.029	0.051	-0.116***	0.025	0.146***	0.056
	Special education	-0.265***	0.051	-0.398***	0.018	0.133**	0.055
	Limited English proficiency	-0.023	0.063	0.044**	0.020	-0.067	0.066
	Baseline math score	0.119***	0.030	0.164***	0.011	-0.044	0.032
	Baseline ELA score	0.284***	0.031	0.366***	0.010	-0.083**	0.033
	Value-added of public schools in zip code	0.634	0.466	0.893***	0.185	-0.259	0.501
$\alpha^\theta_e \times \sigma_\theta$	Taste for charter schools (std. dv. units)	-0.037	0.026	0.027**	0.013	-0.064**	0.029
N	Sample size			9156			

Notes: This table reports maximum simulated likelihood estimates of the parameters of the 8th-grade achievement distribution. Panel A shows estimates for math, while Panel B shows estimates for ELA. The likelihood is evaluated using 300 simulations per observation. Mean potential outcomes are evaluated at the sample mean of the covariate vector X. The mean potential outcome for charter schools is a weighted average of school-specific means, with weights proportional to school capacity. The sample includes all students with observed 8th-grade test scores. The likelihood is evaluated using 300 simulations per observation.
*significant at 10; **significant at 5%; ***significant at 1%

Table 6: Estimates of School-specific Parameters

School	Average utility (1)	Admission probability (2)	Test score effects: Math (3)	Test score effects: ELA (4)
Charter school 1	-0.695***	0.516***	0.492***	0.577***
	(0.139)	(0.064)	(0.116)	(0.125)
Charter school 2	-0.615***	0.390***	0.471***	0.540***
	(0.136)	(0.057)	(0.101)	(0.118)
Charter school 3	-1.308***	0.653***	0.543***	0.510***
	(0.140)	(0.039)	(0.123)	(0.137)
Charter school 4	-1.638***	0.706***	0.771***	0.618***
	(0.135)	(0.051)	(0.119)	(0.139)
Charter school 5	-0.801***	0.394***	0.551***	0.344***
	(0.132)	(0.074)	(0.115)	(0.126)
Charter school 6	-2.203***	0.824***	0.682***	0.834***
	(0.145)	(0.055)	(0.123)	(0.149)
Charter school 7	-1.883***	0.875***	0.968***	1.007***
	(0.158)	(0.039)	(0.171)	(0.174)

Notes: This table reports maximum simulated likelihood estimates of the school-specific parameters from the structural model. The likelihood is evaluated using 300 simulations per observation. The admission probabilities in column (2) are averages for 2006-2009. Average utilities and test score effects are computed at the population mean of the covariate vector X.
*significant at 10; **significant at 5%; ***significant at 1%

Table 7: Charter School Treatment Effects

Subject	TOT (1)	ATE (2)	TNT (3)
Math	0.480	0.639	0.680
ELA	0.419	0.562	0.578

Notes: This table reports model-predicted effects of treatment on the treated (TOT), average treatment effects (ATE), and effects of treatment on the non-treated (TNT) for Boston's charter schools in 8th grade. The TOT is the effect of charter attendance on students who attend charter schools. The ATE is the average potential effect of charter schools on all students. The TNT is the average potential effect of charter attendance on students who do not attend charter schools. Treatment effects are weighted averages of school-specific effects, with weights proportional to school capacity.

Table 8: Simulated Effects of Charter School Expansion

	Charter demand				Test scores			
Policy change	Application rate (1)	Attendance rate (2)	Admission probability (3)	Fraction of seats filled (4)	Avg. math score (5)	Math TOT (6)	Avg. ELA score (7)	ELA TOT (8)
None (7 charter schools)	0.189	0.094	0.670	1.000	-0.359	0.493	-0.460	0.419
All charter schools close	-	-	-	-	-0.402 (-11.6%)	-	-0.493 (-6.1%)	-
Boston's planned expansion (expand to 13 schools)	0.326 (72.43%)	0.175 (86.47%)	0.874 (30.52%)	1.000 (0.00%)	-0.329 (-8.47%)	0.513 (4.10%)	-0.431 (-6.29%)	0.452 (7.76%)
Expand to 20 schools	0.398 (110.37%)	0.214 (128.22%)	0.870 (29.92%)	0.838 (-16.24%)	-0.308 (-14.18%)	0.527 (6.88%)	-0.409 (-11.05%)	0.470 (12.08%)

Notes: This table reports simulated effects of changing Boston's charter school network on charter demand and test scores. Numbers in parentheses are percentage changes relative to the existing charter system. Simulated statistics are produced using 100 simulations per observation in the data set.

References

1. Abadie, A. (2002). "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models." *Journal of the American Statistical Association* 97(457).

2. Abdulkadiroglu, A., Angrist J., Dynarski, S., Kane, T., and Pathak, P. (2011). "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *The Quarterly Journal of Economics* 126(2).

3. Ajayi (2013). "School Choice and Educational Mobility: Lessons from Secondary School Applications in Ghana." Mimeo, Boston University.

4. Angrist, J., Cohodes, S., Dynarski, S., Pathak, P., and Walters, C. (2013a). "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice." NBER Working Paper no. 19275.

5. Angrist, J., Pathak, P., and Walters, C. (2013b). "Explaining Charter School Effectiveness." *American Econometric Journal: Applied Economics* 5(4).

6. Angrist, J., Dynarski, S., Kane, T., Pathak, P., and Walters, C. (2012). "Who Benefits from KIPP?" *Journal of Policy Analysis and Management* 31(4).

7. Angrist, J., and Imbens, G. (1995). "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90(430).

8. Arcidiacono, P., Aucejo, E., and Hotz, V. (2013). "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." Mimeo, Duke University.

9. Berry, S., Levinsohn, J., and Pakes, A. (1995). "Automobile Prices in Market Equilibrium." *Econometrica* 63(4).

10. Booker, K., Sass, T., Gill, B., and Zimmer, R. (2011). "The Effects of Charter High Schools on Educational Attainment." *Journal of Labor Economics* 29(2).

11. Boston Municipal Research Bureau (2008). "The Status of Charter Schools in Boston." http://www.bmrb.org /content/upload/Charter608.pdf.

12. Brand, J., and Xie, Y. (2010). "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75(2).

13. Butler, J., Carr, D., Toma, E., and Zimmer, R. (2013). "Choice in a World of New School Types." *Journal of Policy Analysis and Management*, forthcoming

14. Card, D. (1993). "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." NBER Working Paper no. 4483.

15. Chade, H., Lewis, G., and Smith, L. (2009). "A Supply and Demand Model of the College Admissions Problem." Mimeo, Arizona State University.

16. Chade, H., and Smith, L. (2006). "Simultaneous Search." *Econometrica* 74(5).

17. Commonwealth of Massachusetts (2010). "An Act Relative to the Achievement Gap: Turning Around Low-Performing Schools and Promoting Innovation for All." https://www.mass.gov/Eoedu/docs/ legislation_policy/20100125_ed_law_fact_sheet.pdf.

18. Center for Research on Education Outcomes (CREDO, 2013). "National Charter School Study." Stanford University.

19. Curto, V., and Fryer, R. (2011). "Estimating the Returns to Urban Boarding Schools: Evidence from SEED." NBER Working Paper no. 16746.

20. Dahl, G. (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70(6).

21. Deming, D., Hastings, J., Kane, T., and Staiger, D. (2014). "School Choice and College Attendance: Evidence from Randomized Lotteries." *American Economic Review*, forthcoming.

22. Dillon, E., and Smith, J. (2013). “The Determinants of Mismatch Between Students and Colleges.” NBER Working Paper no. 19286.

23. Dobbie, W., and Fryer, R. (2011). “Are High Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children’s Zone.” *American Economic Journal: Applied Economics* 3(3).

24. Dobbie, W., and Fryer, R. (2013). “Getting Beneath the Veil of Effective Schools: Evidence from New York City.” *American Economic Journal: Applied Economics* 5(4).

25. Dubin, J., and McFadden, D. (1984). “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption.” *Econometrica* 52(2).

26. Epple, D., and Romano, R. (1998). “Competition between Private and Public Schools, Vouchers, and Peer-Group Effects.” *The American Economic Review* 88(1).

27. Epple, D., Romano, R., and Seigh, H. (2003). “Peer Effects, Financial aid and Selection of Students into Colleges and Universities: An Empirical Analysis.” *Journal of Applied Econometrics* 18(5).

28. Epple, D., Romano, R., and Sieg, H. (2006). “Admission, Tuition, and Financial Aid Policies in the Market for Higher Education.” *Econometrica* 74(4).

29. Epple, D., Romano, R., Sarpca, S., and Sieg, H. (2013). “The US Market for Higher Education: A General Equilibrium Analysis of State and Private Colleges and Public Funding Policies.” NBER Working Paper no. 19298.

30. Ferreyra, M., and Kosenok, G. (2011). “Charter School Entry and School Choice: The Case of Washington, D.C.” Mimeo, Carnegie Mellon University.

31. Fryer, R. (2011). “Creating ’No Excuses’ (Traditional) Public Schools: Preliminary Evidence from an Experiment in Houston.” NBER Working Paper no. 17494.

32. Gleason, P., Clark, M., Tuttle, C., and Dwoyer, E. (2010). “The Evaluation of Charter School Impacts: Final Report.” Washington, D.C.: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

33. Hastings, J., Kane, T., and Staiger, D. (2009). “Heterogeneous Preferences and the Efficacy of Public School Choice.” Mimeo, Yale University.

34. Hastings, J., and Weinstein, J. (2008). “Information, School Choice, and Academic Achievement: Evidence from Two Experiments.” *The Quarterly Journal of Economics* 123(4).

35. Hausman, J., and Wise, D. (1978). “A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences.” *Econometrica* 46(2).

36. Heckman, J. (1979). “Sample Selection as a Specification Error.” *Econometrica* 47(1).

37. Heckman, J. (1990). “Varieties of Selection Bias.” *The American Economic Review* 80(2).

38. Heckman, J., Tobias, J., and Vytlacil, E. (2001). “Four Parameters of Interest in the Evaluation of Social Programs.” *Southern Economic Journal* 68 (2).

39. Hoxby, C., and Avery, C. (2012). “The Missing ’One-Offs:’ The Hidden Supply of High-Achieving, Low-Income Students.” NBER Working Paper no. 18586.

40. Hoxby, C., and Murarka, S. (2009). “Charter Schools in New York City: Who Enrolls and How They Affect Their Students’ Achievement.” NBER Working Paper no. 14852.

41. Hoxby, C., and Rockoff, J. (2004). “The Impact of Charter Schools on Student Achievement.” Harvard Institute of Economic Research Working Paper Series.

42. Imbens, G., and Angrist, J. (1994). “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62(2).

43. Imberman, S. (2011). “The Effect of Charter Schools on Achievement and Behavior of Public School Students.” *Journal of Public Economics* 95(7-8).

44. Kirkeboen, L., Leuven, E., and Mogstad, M. (2014). "Field of Study, Earnings, and Self-Selection." Mimeo, University of Chicago.

45. Lee, L. (1983). "Some Approaches to the Correction of Selectivity Bias." *Review of Economic Studies* 49(3).

46. Levitz, J. (2013). "New Front in Charter Schools." *The Wall Street Journal*, March 10.

47. Massachusetts Department of Elementary and Secondary Education (2011). "Understanding Charter School Tuition Reimbursements." http://www.doe.mass.edu/charter/finance/tuition/Reimbursements.html.

48. Massachusetts Department of Elementary and Secondary Education (2012a). "Application for a Massachusetts Horace Mann Public Charter School." http://www.doe.mass.edu/charter/app/HM_full.pdf.

49. Massachusetts Department of Elementary and Secondary Education (2012b). "Charter School News Archives." http://www.doe.mass.edu/charter/default.html?section=archive.

50. Mehta, N. (2011). "Competition in Public School Districts: Charter School Entry, Student Sorting, and School Input Determination." Mimeo, University of Pennsylvania.

51. Nevo, A. (2000). "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics and Management Strategy* 9(4).

52. Pathak, P., and Sonmez, T. (2008). "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review* 98(4).

53. Rothstein, R. (2004). Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap. Washington, D.C.: Economic Policy Institute.

54. Roy, A. (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2).

55. Roy, A. (2010). "Public Charter School Sets Sights on Vacant South End Church." *South End Patch*, November 10.

56. Sacerdote, B. (2011). "Peer Effects in Education: How Might They Work, How Big are They, and How Much Do We Know Thus Far?" In: Hanushek, E., Marchin, S. and Woessmann, L., eds., Handbook of the Economics of Education, Volume 3. Amsterdam: Elsvier.

57. Train, K. (2003). Discrete Choice Models with Simulation. New York: Cambridge University Press.

58. Walters, C. (2013). "A Structural Model of Charter School Choice and Academic Achievement." Mimeo, MIT.

59. Wilson, S. (2008). "Success at Scale in Charter Schooling." American Enterprise Institute Future of American Education Project, Working Paper no. 2008-02.

For Online Publication

Appendix A: Relationship to Roy Model

This appendix shows that equations (8) through (12) nest a Roy model of selection in which students seek to maximize achievement and have private information about their test scores in charter and public schools. For simplicity, I omit application costs and preferences for distance. Achievement for student i at charter school k is given by

$$Y_i(k) = \alpha_k^0 + X_i'\alpha_c^x + \eta_{ic} + \nu_{ik},$$

while public school achievement is

$$Y_i(0) = \alpha_0^0 + X_i'\alpha_0^x + \eta_{i0} + \nu_{i0}$$

where $E[\nu_{ik}|X_i, \eta_{ic}, \eta_{i0}] = 0$. Assume that students know the parameters of these equations, their own characteristics X_i, and private signals of their achievement in charter and public schools η_{ic} and η_{i0}. Also assume that $(\eta_{ic}, \eta_{i0})'$ follows a bivariate normal distribution with $E[\eta_{i\ell}|X_i] = 0$ and $Var(\eta_{i\ell}) = \sigma_\ell^2$ for $\ell \in \{c, 0\}$, and $Cov(\eta_{ic}, \eta_{i0}) = \sigma_{c0}$. The ν_{ik} represent random fluctuations in test scores unknown to the student.

Suppose that students choose schools to maximize expected achievement. Then student utility can be written

$$u_{ik} = \alpha_k^0 + X_i'\alpha_c^x + \eta_{ic},$$

$$u_{i0} = \alpha_0^0 + X_i'\alpha_0^x + \eta_{i0}.$$

Subtracting u_{i0} from u_{ik}, student preferences can be equivalently represented by the utility functions

$$U_{ik} = \gamma_k^0 + X_i'\gamma^x + \theta_i,$$

where

$$\gamma_k^0 = \alpha_k - \alpha_0,$$

$$\gamma^x = \alpha_c^x - \alpha_0^x,$$

$$\theta_i = \eta_{ic} - \eta_{i0},$$

and $U_{i0} \equiv 0$. These preferences are a special case of equation (8) with $\gamma^d = \gamma^a = 0$ and $Var(\upsilon_{ik}) = Var(\psi_{ia}) = 0$.

Returning to the test score equation, we have

$$E(Y_i(k)|X_i, \theta_i] = \alpha_k^0 + X_i'\alpha_c^x + \alpha_c^\theta \cdot \theta_i,$$

$$E(Y_i(0)|X_i, \theta_i] = \alpha_0^0 + X_i'\alpha_0^x + \alpha_0^\theta \cdot \theta_i,$$

where

$$\alpha_c^\theta = \frac{\sigma_c^2 - \sigma_{c0}}{\sigma_c^2 + \sigma_0^2 - 2\sigma_{c0}},$$

$$\alpha_0^\theta = \frac{\sigma_{c0} - \sigma_0^2}{\sigma_c^2 + \sigma_0^2 - 2\sigma_{c0}}.$$

This implies that potential test scores are given by

$$Y_i(k) = \alpha_k^0 + X_i'\alpha_c^x + \alpha_c^\theta \cdot \theta_i + \epsilon_{ik},$$

$$Y_i(0) = \alpha_0^0 + X_i'\alpha_0^x + \alpha_0^\theta \cdot \theta_i + \epsilon_{i0},$$

where $E[\varepsilon_{ik}|X_i, \theta_i] = 0$, which is the specification for achievement in equations (11) and (12).

Finally, note that the Roy framework implies that $\alpha_c^\theta > 0$, $\alpha_0^\theta < 0$, and $\alpha_c^\theta - \alpha_0^\theta = 1$. If students choose schools to maximize academic achievement, then charter preferences will be positively related to scores in charter schools, negatively related to scores in public schools, and the causal effect of charter attendance will increase with charter preferences.

Appendix B: Identification of Preference Coefficients

This appendix uses a simplified version of the structural model to demonstrate identification of the coefficients on the charter preferences θ_i in equations (11) and (12). Suppose there is a single charter school, and the utilities of charter and public school attendance are given by

$$U_{i1} = \gamma^0 + \gamma^d \cdot D_i + \theta_i + \upsilon_i - \gamma^a \cdot A_i,$$

$$U_{i0} = -\gamma^a \cdot A_i,$$

where D_i is the distance to the charter school, A_i indicates a charter application, $\theta_i \sim N(0, \sigma_\theta^2)$ is observed prior to the application decision, and $\upsilon_i \sim N(0,1)$ is observed after the application decision.[24] The charter school holds a lottery for applicants with acceptance probability π.

The expected utility of applying to the charter school is

$$\pi \cdot E[\max\{\gamma^0 + \gamma^d \cdot D_i + \theta_i + \upsilon_i, 0\}|\theta_i] - \gamma^a,$$

while not applying yields utility of zero with certainty. It is optimal to apply if

$$\psi(\gamma^0 + \gamma^d \cdot D_i + \theta_i) > \frac{\gamma^a}{\pi},$$

where $\psi(t) \equiv \Phi(t) \cdot (t + \phi(t))$. It is straightforward to show that $\psi(\cdot)$ is strictly increasing, so the application rule can be written

$$A_i = 1\{\theta_i > \theta^*(D_i)\},$$

where

$$\theta^*(D) = \psi^{-1}\left(\frac{\gamma^a}{\pi}\right) - \gamma^0 - \gamma^d \cdot D.$$

Note that with $\gamma^d < 0$, we have $\frac{d\theta^*}{dD} > 0$: students who live further from the charter school must have stronger tastes for charter attendance to justify incurring the application cost.

Let $S_i(z)$ indicate charter attendance as a function of Z_i. Rejected applicants cannot attend, so $S_i(0) = 0 \forall i$. Attendance for admitted applicants is given by

$$S_i(1) = 1\{\gamma^0 + \gamma^d \cdot D_i + \theta_i + \upsilon_i > 0\}.$$

Lottery applicant compliers choose to apply and have $S_i(1) = 1$. Compliers are therefore characterized by

$$(A_i = 1) \cap (S_i(1) > S_i(0)) \Longleftrightarrow \theta_i > \max\{\theta^*(D_i), -\gamma^0 - \gamma^d \cdot D_i - \upsilon_i\}.$$

The model for potential outcomes in charter and public school is

[24] I use a normal distribution rather than an extreme value distribution for υ_i because it allows me to obtain analytic formulas in the calculations to follow.

$$Y_i(1) = \alpha_1^0 + \alpha_1^\theta \cdot \theta_i + \epsilon_{i1},$$

$$Y_i(0) = \alpha_0^0 + \alpha_0^\theta \cdot \theta_i + \epsilon_{i0},$$

with $E[\epsilon_{i\ell}|\theta_i, D_i] = 0$ for $\ell \in \{0,1\}$. It is straightforward to show that average potential outcomes for compliers who live a distance D from charter schools are given by

$$E[Y_i(\ell)|A_i = 1, S_i(1) > S_i(0), D_i = D] = \alpha_\ell^0 + \alpha_\ell^\theta \cdot \mu_\theta^c(D),$$

where

$$\mu_\theta^c(D) = \sigma_\theta \cdot \Phi\left(\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right) \cdot \lambda\left(\frac{\theta^*(D)}{\sigma_\theta}\right)$$

$$+\sigma_\theta \cdot (1 - \Phi\left(\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right)) \cdot \int \lambda\left(\frac{-\gamma^0 - \gamma^d \cdot D - \upsilon_i}{\sigma_\theta}\right) dF\left(\upsilon_i|\upsilon_i < -\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right)$$

Here $\lambda(t) \equiv \frac{\phi(t)}{1-\Phi(t)}$ is the inverse Mills ratio.

The inverse Mills ratio is an increasing function, so $\mu_\theta^c(D)$ is increasing in D. Applicant compliers who apply to charters from further away therefore have stronger preferences for charters, and comparisons of potential outcomes for lottery compliers who live different distances from charter schools identify the relationship between preferences and achievement. Specifically, for $D_1 \neq D_0$, we have

$$\frac{E[Y_i(\ell)|A_i = 1, S_i(1) > S_i(0), D_i = D_1] - E[Y_i(\ell)|A_i = 1, S_i(1) > S_i(0), D_i = D_0]}{\mu_\theta^c(D_1) - \mu_\theta^c(D_0)} = \alpha_\ell^\theta$$

for $\ell \in \{0,1\}$. The numerator of the left-hand side of this equation can be computed using the methods described in Abadie (2002) for estimating marginal mean counterfactuals for compliers. The denominator is non-zero because complier preferences vary with distance; it can be calculated with knowledge of the parameters of the student utility function, which are identified from charter application and attendance behavior. The selection parameters α_ℓ^θ are therefore identified.

Appendix C: Equilibrium Admission Probabilities

Description of the Game

This appendix describes the determination of equilibrium admission probabilities for use in counterfactual simulations. These probabilities are determined in a Subgame Perfect Nash Equilibrium (SPE) in which students make utility-maximizing choices as described in Section 4, and schools set admission probabilities to fill their capacities, or come as close as possible to doing so.

The time of the game follows Figure 2. Strategies in each stage of the game are as follows:

1. Students choose applications.

2. Schools observe students' application choices, and choose their admission probabilities.

3. Offers are randomly assigned among applicants.

4. Students observe their offers and make school choices.

To simplify the game, I assume that the distribution of students is atomless, so schools do not change their admission probabilities in the second stage in response to the application decisions of individual students in the first stage. Students therefore act as "price takers" in the first stage, in the sense that they do not expect schools to react to their application choices. This implies that the game can be analyzed as if applications and admission probabilities are chosen simultaneously. I analyze the static Nash equilibria of this simultaneous-move game, which are equivalent to Subgame Perfect equilibria of the dynamic game described above.

Definition of Equilibrium

An equilibrium of the game requires an application rule for each student, a vector of admission probabilities π^*, and a rule for assigning school choices that satisfy the following conditions:

1. The probability that student i chooses application bundle a is given by $Q_{ia}(\theta_i, \tau_i; \pi^*)$, where Q_{ia} is defined as in Section 4 and now explicitly depends on the vector of admission probabilities students expect to face in each lottery

2. For each k, π^*_k is chosen to maximize enrollment subject to school k's capacity constraint, taking student application rules as given and assuming that other school choose π^*_{-k}, which denotes the elements of π^* excluding the k-th.

3. After receiving the offer vector z, student i chooses school k with probability $P_{ik}(z, \theta_i, \tau_i)$ as in Section 4.

School Problem

I begin by deriving a school's optimal admission probability as a function of students' expected admission probabilities and the actions of other schools. Let Λ_k denote the capacity of school k, which is the maximum share of students that can attend school k. Suppose that students anticipate the admission probability vector π^e when making application decisions in the first stage of the model. Their application decisions are described by

$Q_{ia}(\theta_i, \tau_i; \pi^e)$. In addition, suppose that schools other than k admit students with probability π_{-k}. If school k admits students with probability π_k in the second stage, its enrollment is given by

$$e_k(\pi_k, \pi_{-k}, \pi^e) = E\left[\sum_{a\in\{0,1\}^K}\sum_{z\in\{0,1\}^K} Q_{ia}(\theta_i, \tau_i; \pi^e) f(z|a; \pi_k, \pi_{-k}) P_{ik}(z, \theta_i, \tau_i)\right].$$

School k choose π_k to solve

$$\max_{\pi_k\in[0,1]} e_k(\pi_k, \pi_{-k}, \pi^e) \quad s.t. \quad e_k(\pi_k, \pi_{-k}, \pi^e) \leq \Lambda_k. \tag{15}$$

The best response function $\pi_k^{BR}(\pi_{-k}, \pi^e)$ is the solution to problem (15). The optimal admission probability sets school k's enrollment equal to its capacity if possible. The following equation implicitly defines π_k^{BR} at interior solutions:

$$E\left[\sum_a\sum_z Q_{ia}(\theta_i, \tau_i; \pi^e) f(z|a; \pi_k^{BR}, \pi_{-k}) P_{ik}(z, \theta_i, \tau_i)\right] = \Lambda_k.$$

Noting that $P_{ik}(z) = 0$ when $z_k = 0$ (since school k is not in student i's choice set if she does not receive an offer) and setting $f_k(1|a_k; \pi_k) = a_k\pi_k$, this equation can be rewritten as

$$E\left[\sum_{a:a_k=1}\sum_{z:z_k=1} Q_{ia}(\theta_i, \tau_i; \pi^e) f_{-k}(z_{-k}|a_{-k}; \pi_{-k}) \cdot \pi_k^{BR} \cdot P_{ik}(z, \theta_i, \tau_i)\right] = \Lambda_k,$$

where z_{-k}, a_{-k}, and f_{-k} are z, a and f excluding the k-th elements. An interior solution for π_k^{BR} therefore satisfies

$$\pi_k^{BR} = \frac{\Lambda_k}{E\left[\sum_{a:a_k=1}\sum_{z:z_k=1} Q_{ia}(\theta_i, \tau_i; \pi^e) f_{-k}(z_{-k}|a_{-k}; \pi_{-k}) P_{ik}(z, \theta_i, \tau_i)\right]}$$

$$\equiv \Gamma_k(\pi_{-k}, \pi^e).$$

If the denominator of Γ_k is sufficiently small, it may exceed one, in which case school k cannot fill its capacity. In this case, the optimal action is to set $\pi_k = 1$ and fill as many seats as possible. This implies that the best response function is given by

$$\pi_k^{BR}(\pi_{-k}, \pi^e) = \min\{\Gamma_k(\pi_{-k}, \pi^e), 1\}.$$

Existence of Equilibrium

Let $\pi^{BR} : [0,1]^K \to [0,1]^K$ be the vector-valued function defined by

$$\pi^{BR}(\pi) \equiv \left(\pi_1^{BR}(\pi_{-1}, \pi), \ldots, \pi_K^{BR}(\pi_{-K}, \pi)\right)'.$$

A vector of admission probabilities supports a Nash equilibrium if and only if it is a fixed point of $\pi^{BR}(\pi)$. The following theorem shows that an equilibrium of the game always exists.

Theorem: *There exists a $\pi^* \in [0,1]^K$ such that $\pi^{BR}(\pi^*) = \pi^*$.*

Proof: Note that $Q_{ia}(\theta_i, \tau_i; \pi)$ is continuous in π and strictly positive, $P_{ik}(z, \theta_i, \tau_i)$ is strictly positive when $z_k = 1$, and $f_{-k}(z_{-k}|a_{-k}; \pi_{-k})$ is continuous in π_{-k} and sums to one for each a_{-k}, so the denominator of Γ_k is always non-zero and continuous in π. π_k^{BR} is therefore a composition of continuous functions, and is continuous. Then π^{BR} is a continuous function that maps the compact, convex set $[0,1]^K$ to itself. Brouwer's Fixed Point Theorem immediately applies and π^{BR} has at least one fixed point in $[0,1]^K$.

Uniqueness of Equilibrium

I next give conditions under which the equilibrium is unique. Define the functions

$$\Delta_k(\pi) \equiv \pi_k - \min\{\Gamma_k(\pi_{-k}, \pi), 1\}$$

and let $\Delta(\pi) \equiv (\Delta_1(\pi), \ldots, \Delta_K(\pi))'$. A vector supporting an equilibrium satisfies $\Delta(\pi^*) = 0$. A sufficient condition for a unique equilibrium is that the Jacobean of $\Delta(\pi)$ is a positive dominant diagonal matrix. This requires the following two conditions to hold at every value of $\pi \in [0,1]^K$:

$$\text{1a.} \quad \frac{\partial \Delta_k}{\partial \pi_k} > 0 \quad \forall k$$

$$\text{2a.} \quad \left|\frac{\partial \Delta_k}{\partial \pi_k}\right| \geq \sum_{j \neq k} \left|\frac{\partial \Delta_k}{\partial \pi_j}\right| \quad \forall k$$

To gain intuition for when a unique equilibrium is more likely, note that in any equilibrium, admission probabilities must be strictly positive for all schools; an admission rate of zero guarantees zero enrollment, while expected enrollment is positive and less than Λ_k for a sufficiently small positive π_k. When $\pi_k > 0$, we can write Γ_k as

$$\Gamma_k(\pi_{-k}, \pi) = \frac{\Lambda_k \pi_k}{e_k(\pi_k, \pi_{-k}, \pi)}$$

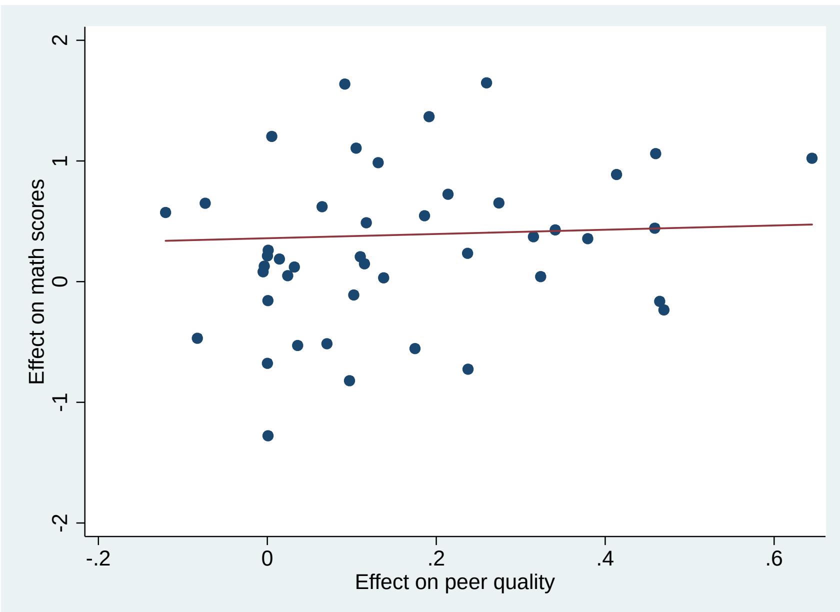
.It follows that conditions 1a and 2a are equivalent to the following conditions on the model's enrollment elasticities:

$$\text{1b.} \quad \frac{\partial \log e_k}{\partial \log \pi_k} > \left(\frac{\Lambda_k - e_k}{\Lambda_k}\right) \quad \forall k$$

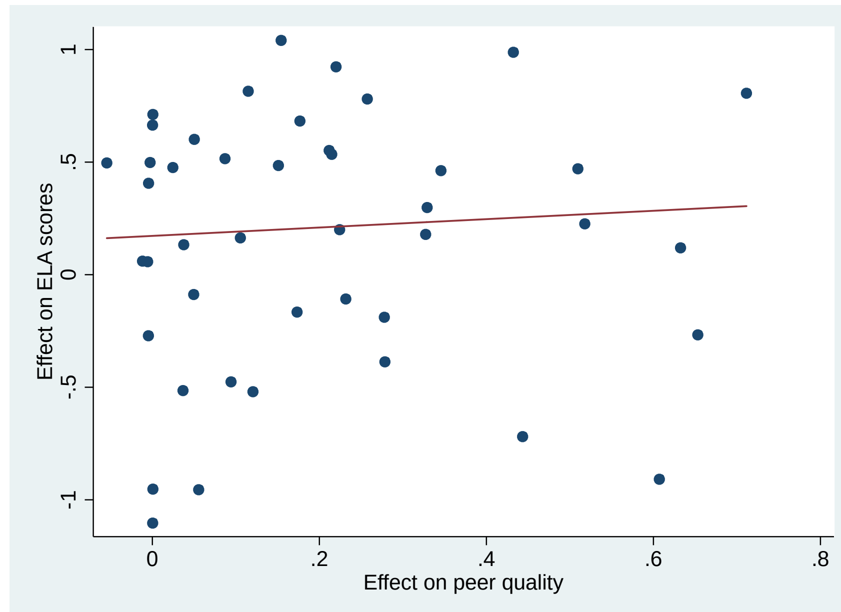
$$\text{2b.} \quad \frac{\partial \log e_k}{\partial \log \pi_k} \geq \sum_{j \neq k} \frac{\pi_k}{\pi_j} \cdot \left|\frac{\partial \log e_k}{\partial \log \pi_j}\right| + \left(\frac{\Lambda_k - e_k}{\Lambda_k}\right) \quad \forall k$$

Condition 1b necessarily holds in the neighborhood of an equilibrium since the elasticity of school k's enrollment with respect to its own admission probability is positive and $\Lambda_k \approx e_k$. This condition is more likely to hold throughout the parameter space when demand for charter schools is strong, so that $e_k(\pi_k, \pi_{-k}, \pi) > \Lambda_k$ at most values of π. Condition 2b is also more likely to hold in these circumstances, and when the cross elasticities of enrollment at school k with respect to other schools' admission probabilities are small. This occurs when charter demand is more segmented. If preferences for distance are strong enough, for example, each student will consider only the closest charter school, and the cross elasticities are zero, leading to a unique equilibrium. To compute equilibria in the counterfactual simulations, I numerically solved for fixed points of the best response vector $\pi^{BR}(\pi) \equiv \left(\pi_1^{BR}(\pi, \pi_{-1}), \ldots, \pi_K^{BR}(\pi, \pi_{-K})\right)'$. I never found more than one equilibrium in any counterfactual.

Figure A1: Relationship Between Effects on Peer Quality and Effects on Test Scores



A. Math



B. ELA

Notes: This figure plots coefficients from regressions of 6th-grade test scores on lottery offers against coefficients from regressions of peer quality on offers, lottery by lottery. Lotteries are defined as combinations of application cohorts and schools applied. Peer quality for a given student is defined as the average 4th-grade test score of the students with whom he or she attends 6th grade. The red lines are from OLS regressions of test score effects on peer quality effects, weighting by sample size. The slope are 0.18 (s.e. = 0.43) for math and 0.19 (s.e. = 0.34) for ELA.

Table A1: Boston Charter Middle Schools

School name (1)	Grade coverage (2)	Years open (3)	Records available (4)	Oversubscribed cohorts (5)	Linked schools (6)
		Panel A. Schools open before 2011			
Academy of the Pacific Rim	5-12	1997-	Yes	2006-2009	-
Boston Collegiate	5-12	1998-	Yes	2006-2009	-
Boston Preparatory	6-12	2004-	Yes	2006-2009	-
Edward Brooke	K-8 (with 5th entry)	2002-	Yes	2007-2009	-
Excel Academy	5-8	2003-	Yes	2008-2009	-
MATCH Middle School	6-8	2008-	Yes	2007-2009	-
Smith Leadership Academy	6-8	2003-	No	-	-
Roxbury Preparatory	6-8	1999-	Yes	2006-2009	-
Uphams Corner	5-8	2002-2009	No	-	-
		Panel B. Expansion schools			
Dorchester Preparatory	5-12	2012-	-	-	Roxbury Preparatory
Edward Brooke II	K-8 (with 5th entry)	2011-	-	-	Edward Brooke
Edward Brooke III	K-8 (with 5th entry)	2012-	-	-	Edward Brooke
Excel Academy II	5-12	2012-	-	-	Excel Academy
Grove Hall Preparatory	5-12	2011-	-	-	Roxbury Preparatory
KIPP Academy Boston	5-8	2012-	-	-	KIPP Academy Lynn

Notes: This table lists charter middle schools serving traditional student populations in Boston, Massachusetts. Schools are included if they accept students in 5th or 6th grade. Panel A lists schools open between the 2007-2008 and 2011-2012 school years, while Panel B lists expansion school opened for 2011-2012 and 2012-2013. Column (3) lists the opening and (where relevant) closing year for each school. Column (4) indicates whether applicant records were available for cohorts attending 4th grade between 2006 and 2009, and column (5) lists the cohorts for which lotteries were held during this period. For expansion schools, column (6) lists existing Massachusetts charter schools operated by the same organization.

Table A2: Covariate Balance

Variable	Differential (1)
Female	-0.023
	(0.031)
White	-0.013
	(0.020)
Hispanic	-0.003
	(0.026)
Subsidized lunch	0.009
	(0.027)
Special education	-0.010
	(0.024)
Limited English proficiency	-0.005
	(0.019)
Miles to closest charter school	-0.036
	(0.066)
Value-added of public schools in zip code	0.002
	(0.003)
4th grade math score	-0.044
	(0.060)
4th grade ELA score	0.034
	(0.062)
Joint *p*-value	0.846
N	1794

Notes: This table reports coefficients from regressions of baseline characteristics on a lottery offer dummy, controlling for lottery fixed effects. The *p*-value is from a test that the coefficients in all regressions are zero.

Table A3: Attrition

	Follow-up rate	Differential
	(1)	(2)
All students	0.889	-
Lottery applicants	0.869	-0.001
		(0.015)
N (scores)	35849	6417

Notes: This table reports the fraction of follow-up test scores in grades 6 through 8 observed for students attending 4th grade in Boston between 2006 and 2009. A student is coded as observed in a grade if both her math and ELA scores are recorded. The sample stacks grades, and includes observations for all scores that should be observed assuming normal academic progress after 4th grade. Column (1) shows the follow-up rate, while column (2) shows the difference in follow-up rates for charter lottery winners and losers. This differential is computed from a regression that controls for lottery fixed effects. The standard error is robust to heteroskedasticity and is clustered at the student level.

*significant at 10; **significant at 5%; ***significant at 1%

Table A4: School Practices

Practice	School 1 (1)	School 2 (2)	School 3 (3)	School 4 (4)	School 5 (5)	School 6 (6)	School 7 (7)	Other MA (8)
Instruction time								
Days per year	190	190	190	180	185	193	190	185
Length of school day (hours:minutes)	8:25	7:00	8:30	7:56	9:00	7:33	7:14	7:17
School philosophy (5 pt. scale)								
No Excuses	4	4	4	5	5	5	5	2.76
Emphasize traditional reading and math	5	5	5	5	5	5	4	3.86
Emphasize discipline/comportment	5	5	5	5	5	5	5	3.33
Emphasize measurable results	5	5	5	5	5	5	5	3.62
School practices (1 or 0 for yes/no)								
Parent and student contracts	1	1	1	0	1	1	1	0.67
Uniforms	1	1	1	1	1	1	1	0.74
Merit/demerit system	1	1	1	1	0	1	1	0.30
Classroom techniques (5 pt. scale)								
Cold calling	3	5	5	5	5	3	5	2.48
Math drills	2	4	5	5	5	5	5	3.33
Reading aloud	4	5	5	4	4	5	4	3.14

Notes: This table shows school practices at Boston charter middle schools, measured from a survey of school administrators. Columns (1)-(7) show practices for the 7 schools used to estimate the structural model, while column (8) shows an average for other charter middle schools in Massachusetts.

Table A5: Distance to Charter Schools Among Applicants

Applicants choosing:	Fraction (1)	Mean distance (2)	Extra distance (3)
Closest charter	0.405	1.91	0.00
2nd closest	0.22	2.94	1.12
3rd closest	0.16	4.17	2.39
4th closest	0.09	5.09	3.11
5th closest	0.081	6.70	4.70
6th closest	0.037	8.50	6.48
7th closest	0.006	11.73	9.84

Notes: This table shows the fractions of applicants who applied to each possible choice by distance. Column (1) shows fractions of applicants whose closest chosen school had a given rank in the set of school-specific distances. Column (2) shows mean distance among students who made each choice. Column (3) shows extra distance relative to the closest charter school.

Table A6: Determinants of Charter School Locations

Variable	Probit marginal effect (1)
Share non-white	1.12
	(1.02)
Share subsidized lunch	2.54
	(1.65)
Average MCAS score	1.55*
	(0.83)
N (zip codes)	22

Notes: This table reports marginal effects from a probit model for charter school location decisions. Each observation is a zip code. The dependent variable is an indicator equal to one if a charter school is located in the zip code. Marginal effects are evaluated at the sample mean.
*significant at 10; **significant at 5%; ***significant at 1%