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, the Becker Applied Economics Seminar at the University of Chicago, Cornell University, Duke University, the Federal Reserve Bank of St. Louis, Mathematica Policy Research, MIT, Northeastern University, the Spencer Foundation,the Stanford Graduate School of Business, UC Berkeley, UC Davis, UCLA, the University of Chicago Harris School, the University of Michigan, the University of Virginia, Warwick University, and WEAI provided useful comments. I also thank Aviva Aron-Dine, David Card, David Chan, Sarah Cohodes, William Darrity Jr., Will Dobbie, Susan Dynarski, Maria Ferreyra, James Heckman, Patrick Kline, Michal Kurlaender, Bridget Terry Long, Magne Mogstad, 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.

The Demand for Effective Charter Schools
Christopher R. Walters
NBER Working Paper No. 20640
October 2014, Revised November 2015
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 framework describing charter school application and attendance decisions with a selection model linking student preferences to the achievement gains generated by charter attendance. The model is estimated using instruments derived from randomized entrance lotteries and distance to charter schools. The results show that charter schools reduce achievement gaps between high- and low-achieving groups, so poorer, lower-achieving students have a comparative advantage in the charter sector. Higher-income students and students with high prior test scores have stronger demand for charter schools, however, implying that preferences for charters are inversely related to potential achievement gains. The structural estimates show a similar pattern of selection on unobservables: test score gains are smaller for students with stronger unobserved tastes for charter schools. These findings imply that students do not sort into charter schools on the basis of comparative advantage in academic achievement; instead, potential achievement gains are larger for students who choose not to apply to charter schools. Simulations of an equilibrium school choice model indicate that efforts to target students who are unlikely to apply could substantially boost the effects of charter school expansion.

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 expand the scope for school choice are an increasingly common phenomenon in US public school districts. Examples include charter schools, vouchers, and district-wide choice plans allowing students to choose from a variety of traditional public schools. A central motivation for such reforms is that school choice may serve as an "escape hatch" for disadvantaged students with low-quality neighborhood schools, permitting exit to higher-quality schools and pressuring ineffective schools to improve. School choice also creates scope for improved allocative efficiency: students may sort into schools that are particularly good matches, increasing aggregate productivity through comparative advantage (Hoxby, 1998; 2003). On the other hand, school choice might widen educational inequality if richer families are more likely to choose high-quality schools, and competitive incentives may be weak if most parents choose based on factors other than school quality (Ladd, 2002; Rothstein, 2006; Barseghyan et al., 2014). The aggregate and distributional effects of school choice depend in large part on which students take advantage of opportunities to attend better schools.

The contemporary school choice debate centers on charter schools, a rapidly growing education reform. Charters are publicly funded, non-selective schools that operate outside traditional districts, allowing them freedom to set curricula and make staffing decisions. Previous studies of charter schools focus on the causal effects of these schools on the students who attend them. While evidence on the effects of non-urban charter schools is mixed,[1] studies based on entrance lotteries show that attendance at charters in Boston and New York's Harlem Children's Zone boosts academic achievement sharply (Abdulkadiroglu et al., 2011; Dobbie and Fryer, 2011). Angrist et al. (2012; 2013a; 2013b; forthcoming), Dobbie and Fryer (2013; 2015), Gleason et al. (2010), Hoxby and Murarka (2009), and Hoxby and Rockoff (2004) also report positive effects for urban charter schools.

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.[2] Existing studies typically restrict attention to samples of lottery applicants, among whom charter 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 essential both for interpreting existing evidence and for evaluating the efficacy of charter school expansions. Of particular interest is whether students sort into the charter sector on the basis of potential achievement gains. If gains are atypically large for charter applicants, local average treatment effects (LATE) derived from lottery-based instruments will overstate potential effects for non-applicants and provide a misleading picture of the impacts of charter expansion (Imbens and Angrist, 1994; Rothstein, 2004).[3] If students

[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. See Epple et al. (2015) for a recent review of research on charter schools.

[2]Exceptions include Hastings et al. (2009), who study preferences submitted to a school choice mechanism in Charlotte NC, and Ferreyra and Kosenok (2011) and Mehta (2011), who develop equilibrium models of charter school entry and student sorting. None of these papers model the link between decisions to apply to school choice programs and causal effects; the latter two do not use application or lottery data. Other related studies look at selection in higher education (Hoxby and Avery, 2012; Arcidiacono et al., 2013; Buter et al., 2013; Dillon and Smith, 2013) and in educational programs outside the US (Ajayi, 2013; Kirkeboen et al., 2014).

[3]Rothstein (2004, p.82) offers a version of this view. He writes of the Knowledge is Power Program (KIPP), a high-performing

with large potential benefits are unlikely to apply, on the other hand, reforms that draw non-applicants into the charter sector may generate substantial returns.

This paper studies the demand for charter middle schools in Boston, with a focus on absolute and comparative advantage in school choice. 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 analyze this process using a dynamic, unordered discrete choice model 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. This portfolio-choice model of charter demand is combined with a flexible model for potential academic achievement in charter schools and traditional public schools. As in the canonical Heckman (1979) sample selection framework, the model allows a link between potential outcomes and the latent preferences that drive charter application and attendance decisions, thereby creating scope for selection according to absolute and comparative advantage. This allows generalization from effects for lottery applicants to causal parameters relevant for policies that expand charter schooling to new populations.

Identification of the model parameters is achieved by combining instruments derived from randomized charter entrance lotteries with a second set of instruments based on distance to charter schools. Intuitively, entrance lotteries identify causal effects for the selected sample of lottery applicants; distance generates variation in preferences that shifts the composition of the applicant sample, tracing out the relationship between unobserved tastes and causal effects. To manage the dimensionality of endogeneity in a dynamic discrete choice setting with many choices, I impose a "one factor" restriction that posits a one-dimensional unobserved preference driving endogenous selection into the charter sector. I present evidence that Boston's charters are relatively homogeneous and that students treat them as close substitutes, suggesting that a one-factor selection structure is reasonable.

Estimates of the model reveal that students do not sort into charter schools on the basis of comparative advantage. Instead, the results indicate a "reverse Roy" selection pattern in which students with the largest potential test score impacts have the weakest demand for charter schools.[4] This pattern can be traced to selection on both observed and unobserved student characteristics. Richer, higher-achieving students are more likely to apply to charter schools, but charters boost scores more for poor students and low-achievers; similarly, test score gains are largest for students with weaker unobserved tastes for charter schools. As a result, existing estimates *understate* the potential achievement effects of charter schools for non-applicants. The model estimates imply that the average potential effect of charter schools on non-charter students (the effect of treatment on the non-treated, TNT) is roughly 40 percent larger than the average effect for enrolled

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."

[4]This terminology is a reference to the canonical Roy (1951) selection model, which assumes pure positive selection on outcome gains. The results here indicate that charter students are instead negatively selected on test score gains.

charter students (the effect of treatment on the treated, TOT). These results are consistent with the possibility that high-performing 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 seek out seats at effective schools.

I quantify the policy implications of these results by simulating charter expansion effects in an equilibrium school choice model. The simulations show that while effects for marginal applicants are substantial, the effects of charter expansion may nevertheless be limited by weak demand among the highest-benefit students: students act as if charter application costs are high, and many prefer to attend traditional public schools even when charters offering guaranteed admission are located in close proximity. This finding motivates simulations of structural reforms that make charter attendance more attractive rather than merely adding more seats. The results show larger impacts for reforms that draw in new populations than for reforms that expand the scale of the charter system but do not alter its structure. This implies that policies targeting students who are unlikely to apply have the potential to boost overall productivity in the charter sector. Importantly, reforms that target students based on observables are shown to be less effective than policies that change patterns of selection on unobservables, indicating the importance of attention to self-selection in the design of school choice programs.

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 outlines the structural model of charter demand and academic achievement, and Section 4 discusses identification and estimation of the model. Section 5 reports the parameter estimates. Section 6 summarizes patterns of selection and comparative advantage in charter school choice, and compares these patterns to what might be learned from an atheoretical extrapolation based on lottery applicants. Section 7 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.[5] 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 (Mas-

[5]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. See Abdulkadiroglu et al. (2014) for a recent analysis of Horace Mann charters.

sachusetts 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 an educational 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 and other outcomes (Angrist et al., 2013b; Curto and Fryer, 2011; Dobbie and Fryer, 2013, 2015; 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, 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.

The demand for charters in Boston is 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.[6] 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

[6]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.

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 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 here come from three sources. Demographics, school attendance, and test scores are derived from an administrative database provided by the Massachusetts Department of Elementary and Secondary Education (DESE). Spatial locations are coded from data on student addresses provided by the BPS district. Finally, information on charter school applications and lottery offers comes from 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. Students must also have non-missing 4th grade demographics and test scores, as well as school attendance information and test scores in 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. Students are coded as charter enrollees if they attend a charter school at any time prior to the relevant test.

Student addresses are merged with 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, coded as great-circle distance in miles.[7]

The DESE and address data are matched 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 this period, so the admissions data provides a nearly-complete picture of charter application decisions. 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.[8] Of the two schools without available records, one closed prior to the 2010-2011 school year; the

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

[8] I classify charter schools as middle schools if they accept applicants in 5th or 6th grade. Two Boston charter schools accept

other declined to provide records. In the analysis below, I treat these schools as equivalent to traditional public schools. Lottery records are matched 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.

2.3 Descriptive Statistics

The final analysis sample includes 9,156 students who attended BPS schools in 4th grade between 2006 and 2009. Descriptive statistics for this sample are reported in Table 1. As shown in column (1), 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 describes the quality of nearby traditional public school options. Students can choose between many schools in Boston's centralized traditional public school assignment mechanism, so the characteristics of the relevant fallback 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 schools. I proxy for the quality of nearby schools using estimates from a value-added regression of middle school test scores on school indicators, with controls for demographics and lagged scores.[9] Though value-added estimates of this type may be biased by student sorting, they provide a rough guide to local public school quality. 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 0.52σ and 0.64σ 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. Together, the statistics in Table 1 show that Boston's

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.

[9] The value-added sample stacks test scores in grades through eight and regresses these on 4th grade scores, the other demographic variables from Table 1, and indicators for sixth grade school. The value-added for a given school is the coefficient on the relevant school indicator in this regression. The value-added procedure is jackknifed so that a student's own score does not contribute to her own measure of mean local value-added. Value-added in a student's zip code is the average of math and ELA value-added measures for schools in this area.

charter applicants are less disadvantaged and higher-achieving than the general Boston population on several dimensions. I next outline a model of selection into the applicant sample that can be used to predict charter effects for other groups of students.

3 Modeling Charter School Attendance

3.1 Setup

I model charter application choices as a random utility optimal portfolio choice problem. Figure 1 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 this school. In the third stage, students choose schools denoted $S_i \in \{0, 1...K\}$, 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.

3.2 Student Choice Problem

3.2.1 Preferences

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

$$U_{ik} = \gamma_k^0 + X_i'\gamma^x + \gamma^d D_{ik} + \theta_i + \upsilon_{ik} - c_i(A_i) \quad \text{for } k > 0, \tag{1}$$

where X_i is a vector of observed characteristics for student i and D_{ik} measures distance to school k. The utility of public school attendance is

$$U_{i0} = \upsilon_{i0} - c_i(A_i). \tag{2}$$

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 include the disutility of filling out application forms and the opportunity cost of time spent attending information sessions and lotteries. These costs may also capture frictions associated with learning about charter schools.[10] The application cost function is parameterized as

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

[10] 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).

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 captures any latent factors that influence students to opt out of traditional public school in favor of charter schools, such as the perceived average 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, θ_i is the random coefficient on a charter school indicator (see, e.g., Hausman and Wise, 1978; Berry et al., 1995; and Nevo, 2000). The presence of θ_i implies that charter schools are closer substitutes for each other than for traditional public schools. θ_i is assumed to follow 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 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.[11]

While equations (1) and (2) do not explicitly include academic achievement, they implicitly allow for the possibility that school choices are partially or entirely determined by expected test scores. If a student know her average score gain from attending charters relative to her best available traditional public school option, for example, this information may be included in θ_i. Similarly, information about idiosyncratic achievement benefits at specific schools may enter through υ_{ik}. Appendix A clarifies this point by showing that the model described here is compatible with a classical Roy (1951)-style selection model in which students make choices based on expected achievement. On the other hand, the model does not impose that students choose schools to maximize test scores; it is also compatible with the possibility that school choices are determined by factors unrelated to, or even negatively correlated with, potential achievement.

3.2.2 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 the application vector A_i is

$$f(Z_i|A_i;\pi) = \prod_k [A_{ik} \cdot (\pi_k Z_{ik} + (1-\pi_k)(1-Z_{ik})) + (1-A_{ik}) \cdot (1-Z_{ik})]. \quad (3)$$

[11]In other words, ξ_{ik} follows a standard Gumbel distribution, which provides the scale normalization for the model.

The admission probabilities π_k are allowed to vary by application cohort. If school k is undersubscribed and hence does not hold a lottery for a particular cohort, π_k equals one for that cohort.

3.2.3 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 determines the available school choices. Consider the decision facing a student at stage 3 in Figure 1. 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

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

Write

$$\tilde{U}_{ik}(\theta_i, \tau_{ik}) = \gamma_k^0 + X_i'\gamma^x + \gamma^d D_{ik} + \theta_i + \tau_{ik}$$

for the component of utility at charter k known prior to application, excluding application costs. The corresponding public school utility is $\tilde{U}_{i0}(\theta_i, \tau_{i0}) = \tau_{i0}$. Student i's optimal school choice at stage 3 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

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

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

where ν is Euler's constant.

Students choose charter applications to maximize expected utility, anticipating offer probabilities and their own subsequent attendance choices. Consider the application decision facing a student at stage 1 in Figure 1. 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 $c_i(A_i)$. Let

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

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

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

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

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.

3.3 Academic Achievement

Students are tested after application and attendance decision have been made. Let $Y_i(k)$ denote student i's *potential* test score if she enrolls in school k. These potential outcomes are parameterized as

$$Y_i(k) = \alpha_k^0 + X_i'\alpha_c^x + \alpha_c^\theta\theta_i + \epsilon_{ik} \quad \text{for } k > 0, \tag{4}$$

$$Y_i(0) = \alpha_0^0 + X_i'\alpha_0^x + \alpha_0^\theta\theta_i + \epsilon_{i0}. \tag{5}$$

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. The residual ϵ_{ik} follows a normal distribution with mean zero and variance σ_k^2. This variance is assumed to be the same across charter schools but possibly different between charter and traditional public schools. The observed score for student i is the potential score associated with her optimal school choice: $Y_i = Y_i(S_i)$.

The causal effect of attending charter k relative to traditional public school for student i is $Y_i(k) - Y_i(0)$. The coefficients α_c^θ and α_0^θ govern comparative and absolute advantage in charter school choice with respect to unobservables. If $\alpha_c^\theta > 0$, students with stronger unobserved 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 as in a standard Roy (1951) selection model. Choices are inconsistent with such sorting if the difference is negative.

Together, equations (1) through (5) provide a complete description of charter demand and potential academic achievement. It is worth noting some of the key modeling assumptions implicit in these equations. The model emphasizes differences between charter and traditional public schools, while placing some restrictions on heterogeneity across charter schools. Variation 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 (4) implies that the relationship between the unobserved taste θ_i and student achievement is the same at every charter school. Heterogeneity in mean preferences and potential 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. This emphasis on differences between charters and traditional public schools mirrors the approach to identification described in the next section, which emphasizes selection into the charter sector rather than across charter schools.

4 Identification and Estimation

4.1 Exclusion Restriction

Identification of the parameters of equations (4) and (5) is based on the following exclusion restriction:

$$E[\epsilon_{ik}|X_i, Z_i, D_i, \theta_i, \upsilon_i, \psi_i] = 0, \tag{6}$$

where D_i, υ_i, and ψ_i are vectors of the D_{ik}, υ_{ik} and ψ_{ia}, respectively. Equation (6) embeds three assumptions. First, the lottery offer vector Z_i is excluded from the potential achievement equations. This requires that offers have no direct affect on test scores, a standard assumption in the charter lottery literature. Second, distance to charter schools D_i is excluded from these equations. Finally, the school- and application-specific taste shocks υ_{ik} and ψ_{ia} are excluded. I next discuss the latter two assumptions in detail and provide suggestive evidence in support of them.

4.2 Exclusion of Distance

Equation (6) implies that lottery offers and distance to charter schools are excluded instruments that may be used to identify the effects of charter attendance. Lottery offers identify local average treatment effects for selected samples of charter applicants (Imbens and Angrist, 1994; Angrist, Imbens and Rubin 1996). Going from LATE to more general causal effects requires identification of the relationship between unobserved tastes and causal effects, which in turn necessitates another instrument that shifts the composition of the applicant pool (Heckman and Vytlacil, 2005). Under the restriction in (6), distance serves this purpose. Application rates are likely to be high in the immediate neighborhood of a charter school, which implies that the sample of applicants in this neighborhood is relatively unselected. Observably similar students who are willing to travel long distances to apply to a charter will on average have higher values of the unobserved taste θ_i. If distance is as good as random conditional on X_i, comparisons of mean outcomes by distance therefore identify relationships between θ_i and potential outcomes. Appendix B uses a simplified example to show analytically how the combination of lottery and instruments identifies the selection parameters α_c^θ and α_0^θ, permitting extrapolation outside the lottery sample.[12] This section discusses potential threats to the validity of the distance instrument.

The use of distance as an instrument for charter application 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 the observed covariate vector X_i. A sufficient condition for exclusion is that charter school leaders make choices between neighborhoods on the basis of averages of the variables in X_i. This is 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. It's also worth noting that in contrast to traditional public schools, students who live in the immediate vicinity of a charter school do not receive priority in admission over other Boston students. This mitigates the incentive for students to relocate in order to attend charter schools.

Columns (1) and (2) of Table 2 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 indicates that observed demographic characteristics capture much of the relationship between charter locations and academic achievement, lending 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 2 report separate estimates using lottery offers and distance as instruments for charter attendance in equations for eighth-grade test scores. The lottery estimates come from two-stage least squares (2SLS)

[12] Appendix B looks at a model with one school. In the general seven-school model estimated here, additional information about the selection parameters is derived from the composition of students' application portfolios. For example, the model attributes a stronger unobserved preference to a student who applies to more than one charter school than to an observably-similar student who applies to only one.

models using a lottery offer indicator as an instrument for a charter attendance indicator, controlling for lottery risk sets.[13] The distance models include the full sample and control for student characteristics and baseline scores. 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 2.6 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 0.43σ and 0.28σ in math and ELA, compared to lottery estimates of 0.57σ and 0.48σ. 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.[14]

4.3 Exclusion of School-Specific Preferences

The model estimated here is dynamic, and the set of possible choice sequences is large. As a result, it is not feasible to allow unrestricted dependence of potential outcomes on all the selection errors (θ_i, v_i, ψ_i). There are two common approaches to reducing the dimensionality of endogeneity in multinomial sample selection models (see Bourguignon et al., 2007). First, as in Dubin and McFadden (1984), one can impose a parametric assumption on the joint distribution of outcome and selection errors and use this assumption to derive an expression for the conditional distribution of selected outcomes. The model can then be estimated by either maximum likelihood or a semiparametric two-step procedure. Second, as in Dahl (2002), one can place restrictions on the relationship between mean potential outcomes and selection probabilities while leaving the marginal distribution of outcomes unspecified. This semiparametric approach does not require correct specification of potential outcome distributions, but typically involves ad hoc assumptions about the relationship between outcomes and choice probabilities.

I use a version of the first approach, combining a parametric assumption with exclusion restrictions to manage endogeneity in charter application and attendance decisions. Equations (4) through (6) imply that the conditional mean of ϵ_{ik} is linear in the charter preference θ_i. Since both of these variables are normally distributed, this is equivalent to assuming that θ_i and $Y_i(k)$ follow a joint normal distribution as in the canonical Heckman (1979) sample selection (Heckit) model. I estimate the model by maximum likelihood, so the approach taken here is similar to full information maximum likelihood (FIML) estimation of a Heckit model with a two-stage portfolio choice problem in place of the usual static probit selection equation.

To further limit the dimensionality of the problem, I exclude v_i and ψ_i from equations (4) and (5). This restriction implies that selection on unobservables has a "one factor" form: endogenous selection into the charter

[13] Appendix Table A2 verifies the construction of the lottery offer instrument by comparing baseline characteristics of lottery winners and losers within risk sets. The results show that observed characteristics for these two groups are similar, suggesting random assignment was successful. Appendix Table A3 investigates attrition for the full and sample and by lottery offer status. Followup rates are high for the full sample and for lottery applicants (85 and 81 percent for 8th grade outcomes), and the difference in followup rates between winners and losers is small and statistically insignificant.

[14] 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.

sector depends only on students' overall tastes for charter schools relative to traditional public schools, not on school- or portfolio-specific tastes. This assumption allows students to make choices based on heterogeneity in the average effects of charter schools (captured by γ_k and α_k^0) and their own average effects of charter attendance (captured by θ_i), but it rules out the possibility that choices *between* charters are related to idiosyncratic treatment gains across schools. Intuitively, a student may know which charter schools are better overall, and she may also know whether she is well-suited to the charter treatment in general (because she dislikes her neighborhood traditional public school, for example). I assume that she does not know whether a particular charter is *especially* good for her.

Two pieces of evidence suggest this assumption is reasonable. First, inputs and practices are highly homogeneous across Boston's charter middle schools. This can be seen in Appendix Table A4, which reports responses to a survey on school practices for the seven charter schools studied here along with other charter schools in Massachusetts. 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; these practices are less common elsewhere in the state. With homogeneous school practices, it seems less likely for parents to know that the environment at a specific school will lead to a particularly large idiosyncratic benefit.

Second, Table 3 shows that application portfolio choices among charter applicants are determined mostly by distance. In the model outlined above, the decision to choose one charter school over another is explained 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 conditional on distance. Forty-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 conditional on distance.

4.4 Estimation

I estimate the parameters of the model by maximum simulated likelihood (MSL). Let Ω denote the parameters of equations (1) through (5). The likelihood contribution of a student with outcome variables (A_i, Z_i, S_i, Y_i) is

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

$$\times \frac{1}{\sigma_{s(i)}}\phi\left(\frac{Y_i - \alpha^0_{s(i)} - X_i'\alpha^x_{s(i)} - \alpha^\theta_{s(i)}\theta}{\sigma_{s(i)}}\right) dF(\theta, \tau | X_i, D_i, \Omega). \quad (7)$$

Here the subscript $a(i)$ denotes the application bundle chosen by student i, while $s(i)$ denotes her school choice.[15] The integral in equation (7) is evaluated by simulation. Let θ_i^r and τ_i^r be draws of θ and τ for student i in simulation r, and define the simulation-specific likelihood contribution

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

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. The results were not sensitive to increasing the number of draws beyond around 100. Standard errors are calculated from the average outer product of the score of the simulated likelihood.

5 Structural Estimates

Estimates of the model parameters are reported in tables 4, 5 and 6. These results use 8th grade test scores as outcomes and assume a bivariate normal distribution for the ϵ_{ik} across subjects. Estimates for earlier grades are reported in Appendix tables A5 and A6; patterns of results for these grades are very similar to the 8th grade estimates discussed in detail below.[16] Appendix C provides further analysis showing that the model fits key patterns of heterogeneity in choices and outcomes present in the data.

5.1 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

[15] $s(i)$ 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 $\alpha^x_{s(i)}$.

[16] Models for 7th and 8th grade treat students as charter enrollees if they attend a charter at any time after the lottery and prior to the test. Roughly 20 percent of students initially enrolled in charter schools are no longer enrolled by eighth grade. The results should therefore be interpreted as effects of initial charter enrollment, inclusive of any downstream school switching that results from this enrollment. The consistency of the estimates across grades suggests that post-lottery school switching is not an important factor driving the results.

average marginal effects of observed characteristics on the probability of applying to at least one charter school.[17] The covariate vector X_i is de-meaned in the estimation sample, so the intercept (reported as the average of the γ_k^0) is the average utility of charter attendance. The estimated intercept is negative and statistically significant, implying 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. 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. 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, a 41 percent reduction relative to the mean application rate of 17.5 percentage points.

The bottom half of Table 5 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 terms, a one-standard-deviation increase in θ_i is equivalent to a 13-mile increase in distance to 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).[18] 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 a consequence of the pattern displayed in Table 3, which indicates little heterogeneity in school choices among applicants conditional on distance. The lack of variation in school-specific tastes is further evidence that a one factor selection model is reasonable.

5.2 Achievement Parameters

Estimates of the parameters of 8th-grade math and ELA achievement distributions appear in panels A and B of Table 5. 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.

[17] 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.

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

Columns (2), (4), and (6) report standard errors. The reported charter intercept is the mean of the school-specific intercepts α_k^0. The intercept in column (5) can therefore be interpreted as an average of school-specific population average treatment effects (ATE).

The estimates in Table 5 reveal that charters have larger effects on students with characteristics that predict weak public school performance. 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. This finding is consistent with previous lottery-based estimates showing larger charter impacts for poorer and lower-achieving applicants (Abdulkadiroglu et al., 2011; Angrist et al., 2013b). Patterns of effect heterogeneity among applicants might stem from differences in unobserved characteristics or differences in effectiveness across charter schools, possibilities that were not investigated in previous work. The model estimated here shows that these patterns are due to technological differences in effects across groups rather than differences in unobservables or application portfolios.

Estimates of the selection 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.

5.3 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 suggests that choices between charter schools are not primarily based on differences in test scores effects.

6 Absolute and Comparative Advantage in Charter School Choice

6.1 Selection and Charter School Effects

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 charter preferences define the index

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

$\mathcal{P}_i$ captures student i's average 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 public school outcomes is summarized by the function

$$\alpha_0(p) \equiv E\left[Y_i(0)|\mathcal{P}_i = p\right].$$

$\alpha_0(p)$ is the average potential traditional public school outcome for students with preference p. Similarly, average potential charter achievement can be summarized as $\alpha_c(p) \equiv \sum_{k=1}^{K} w_k E\left[Y_i(k)|\mathcal{P}_i = p\right]$. I set the weights w_k proportional to charter enrollment shares, which implies that $E\left[\alpha_c(\mathcal{P}_i)|S_i > 0\right]$ matches the average outcome for charter students. The average achievement benefit generated by charter attendance is $\beta(p) \equiv \alpha_c(p) - \alpha_0(p)$. This function captures the relationship between charter preferences and the causal effects of charter attendance.[19]

Figure 2 characterizes patterns of absolute and comparate advantage in charter school choice. Panel A plots the conditional expectation functions $\alpha_c(p)$ and $\alpha_0(p)$ (left axis) along with $\beta(p)$ (right axis) for 8th grade math, computed via local linear regressions fit to data simulated from the model. Since the charter preference $\mathcal{P}_i$ has no natural scale it is standardized to have mean zero and standard deviation one in the population. The dashed vertical line shows the mean preference for charter enrollees and the dotted line displays the average preference for traditional public students. The intersections of these lines with the mean potential outcome and charter effect curves can be read as average outcomes and causal effects for charter and non-charter students.

The results in Figure 2 reveal that students with stronger demand for charter schools have absolute advantages in both the charter and traditional public sectors, but students with weaker preferences have a comparative advantage in the charter sector. The mean potential outcomes $\alpha_c(p)$ and $\alpha_0(p)$ rise with charter preferences, a pattern that is driven both by observed characteristics (since disadvantaged students and those with low past scores have weaker tastes for charters and lower outcomes) and unobserved characteristics (since

[19] The $\beta(p)$ function is closely related to the marginal treatment effect (MTE) concept of Heckman and Vytlacil (1999; 2005). MTEs characterize treatment effect heterogeneity as a function of the unobserved cost of treatment participation expressed as a uniformly distributed random variable. In a static two-school model with charter participation equation $S_i = 1\left\{h(X_i, Z_i) + \theta_i > 0\right\}$ and $\theta_i \sim N(0, \sigma_\theta^2)$, MTEs are defined as $MTE(x, u) \equiv E\left[Y_i(1) - Y_i(0)|X_i = x, \Phi(-\theta_i/\sigma_\theta) = u\right]$. The $\beta(p)$ function is an adaptation of this idea to a dynamic setting with many choices, relating average preferences to average causal effects as a function of both observed and unobserved characteristics.

high-θ_i students have stronger tastes for charters and higher outcomes). The slope of $\alpha_c(p)$ is less steep than the slope of $\alpha_0(p)$, however, so the effect of charter attendance $\beta(p)$ falls sharply as charter preference rise.

In contrast to the standard Roy (1951) model of sorting on comparative advantage, this implies a "reverse Roy" pattern in which students with larger treatment gains are less likely to select into charter schools. As a result, potential charter impacts are larger for students who do not attend charter schools than for charter enrollees. Specifically, the effect of treatment on the treated (TOT), given by $E\left[\beta(\mathcal{P}_i)|S_i > 0\right]$, is roughly 0.5σ. The effect of treatment on the non-treated (TNT), defined as $E\left[\beta(\mathcal{P}_i)|S_i = 0\right]$, is over 0.7σ, which represents a 40 percent increase over the TOT.

Panel B of Figure 2 decomposes this pattern into components due to observed and unobserved characteristics. Specifically, this panel plots separate average treatment effects conditional on the observed part of charter preferences ($X_i'\gamma^x$) and the unobserved part (θ_i). These two components are normalized separately in standard deviation units. The results show that selection on unobservables is more extreme than selection on observables, in the sense that the gap in unobserved preferences for charter vs. non-charter students is larger than the gap in observed preferences. The "reverse Roy" pattern is evident for both observables and unobservables, however, and the charter treatment effect slopes down more steeply as a function of observed tastes. As a result, the gap between TOT and TNT is explained in roughly equal measure by observables and unobservables: the difference in treatment effects for charter and non-charter students due to observables is roughly 0.1σ, and the difference due to unobservables is also about 0.1σ.[20] Note that differences in average effects across observed subgroups are identified by differences in lottery estimates for these groups, while differences in effects with respect to unobservables are identified by variation in effects by distance (as discussed in Section 4.3). This shows that the model generates a consistent pattern of negative selection on treatment gains using two different sources of variation.

One possible explanation for this pattern is that 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 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 2.

The selection patterns documented here may be due to "true" tastes or information. In other words, parents may choose not to apply to charter schools either because they do not like the No Excuses charter model, or because they do not know about charter schools. Students may also choose not to apply to charter schools because of logistical barriers in the application process, which appear to be important in other educational contexts (Bettinger et al., 2012). The utility parameters estimated here will capture differences in choice probabilities resulting from any of these possibilities. There is not necessarily a clear distinction between se-

[20] X_i and θ_i are independent by definition, so the overall difference between TOT and TNT equals the sum of differences due to these two components.

lection driven by tastes vs. information; some parents may not know about available schooling options because they are not interested in school quality and therefore have not invested in learning about it. Nonetheless, these possibilities have different implications for charter school expansion, as selection patterns may change if parents gain more information as the system expands. In the counterfactual simulations to follow I consider some reforms that modify preferences to approximate the effects of learning about charter schools.

6.2 Alternative Approaches to Extrapolation

To highlight the value of the structural selection model estimated here, it is worth comparing the model's predicted causal parameters to atheoretical predictions derived from lottery estimates of the type commonly reported in the previous literature. Abdulkadiroglu et al. (2011) use lotteries to estimate the effects Boston charter schools, which show large average charter effects and larger impacts for poorer and lower-achieving students. This section compares the results of reduced-form extrapolation based on these and other key covariates to the insights gleaned from the structural model.

A standard covariate-based approach to extrapolation reweights experimental or quasi-experimental treatment effect estimates to match the distribution of observed characteristics in a new population (see, e.g., Angrist and Fernandez-Val, 2010). This approach can be operationalized through estimation of 2SLS models with second stage

$$Y_i = \alpha + \beta C_i + \varphi_{a(i)} + \epsilon_i, \tag{8}$$

and first stage

$$C_i = \delta + \tau C_i + \kappa_{a(i)} + \eta_i,$$

where $C_i = 1\{S_i > 0\}$ is a charter school indicator, $\tilde{Z}_i = 1\{Z_i \neq 0\}$ is an indicator for an offer at any charter, and $\varphi_{a(i)}$ and $\kappa_{a(i)}$ are dummy variables for lottery portolios. Lottery-based studies of charter schools typically report estimates from models of this form (Abdulkadiroglu et al., 2011; Dobbie and Fryer, 2013). Simple covariate-based predictions of the TNT and TOT are $\sum_g \beta_g Pr[G_i = g|S_i = 0]$ and $\sum_g \beta_g Pr[G_i = g|S_i = 0]$, where G_i indicates an exclusive and exhaustive set of covariate-based groups and β_g is the coefficient from estimation of (8) within group g. I estimate these parameters by plugging in 2SLS estimates of β_g and empirical group probabilities, then compare them to corresponding estimates derived from the structural model.

Covariate-based and model-based predictions of treatment parameters are compared in Table 7. Column (1) replicates the basic 2SLS estimates from Table 2, which equal 0.573 in math and 0.484 in ELA. These estimates are somewhat larger than model-based estimates of the TOT, reported in column (5). This discrepancy reflects the fact that 2SLS estimation of (8) generates a particular weighted average of effects across lotteries and schools that may not be interpretable as an effect for any specific subpopulation of economic interest.[21]

[21] Appendix D shows that the 2SLS estimand in equation (8) is given by $\beta_{2SLS} = \sum_a \omega_a \beta_a$, where β_a is a portfolio-specific IV estimate, the weights ω_a sum to one and

$$\omega_a \propto Pr[A_i = a]\, Pr\left[\tilde{Z}_i = 1|A_i = a\right]\left(1 - Pr\left[\tilde{Z}_i = 1|A_i = a\right]\right)\left(Pr\left[C_i = 1|A_i = a, \tilde{Z}_i = 1\right] - Pr\left[C_i = 1|A_i = a, \tilde{Z}_i = 0\right]\right).$$

Columns (2), (3) and (4) show that reweighting 2SLS estimates based on subsidized lunch status, terciles of baseline test score, or interactions of these covariates with race and special education status tends to raise the implied estimate of TNT relative to TOT. This is a consequence of larger impacts for lower-achieving groups combined with lower charter enrollment probabilities for these groups.

This qualitative pattern is similar to the structural results discussed above. The predicted magnitudes generated by the reduced form and structural approaches are very different, however. Covariate-based extrapolation suggests relatively small gaps between the TNT and TOT (0.035σ and 0.055σ in math and ELA), while the structural approach generates large predicted gaps (0.22σ and 0.18σ). This is driven by the link between unobserved preferences and treatment gains uncovered by the structural model. The model estimates imply that the lottery applicant sample is selected on unobservables in addition to observables, so estimates based on observables in this sample generate inaccurate predictions for the unselected population. In the Boston charter context, extrapolating from lottery-based quasi-experiments to more general policy-relevant causal parameters requires accounting for the selection process that generates the quasi-experimental sample.

7 Counterfactual Simulations

The selection model estimated here shows 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 in general: 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 by simulating the impacts of changes to the Boston charter landscape.

The simulations predict the effects of drawing new populations into charter schools by expanding the set of available schools and (in some cases) changing student preferences. I report on four sets of counterfactual simulations. The first, a "baseline" charter school expansion, adds charters in locations already designated as sites for new schools, then in areas with high predicted probabilities of charter entry based on historical location patterns. The second, a "geographic targeting" counterfactual, adds schools in areas with high predicted test score impacts given the observable characteristics of students nearby. The third "reduced cost" expansion maintains the geographically targeted school locations but modifies preferences to eliminate application costs and make charter and traditional schools equally desirable on average. This counterfactual approximates the effects of expanding information and eliminating logistical barriers under the assumption that the demand-side patterns reported here are largely driven by lack of information rather than true preferences. The final "altered preference" simulation reduces the disutility of charter enrollment for students who are currently unlikely to attend, which may be viewed as an outreach effort that specifically targets low-demand groups. This simulation gives a sense of the potential effects of policies that change the pattern of self-selection into

The parameter β_a is a lottery-specific local average treatment effect, approximately equal to the within-lottery TOT because few students attend charters without offers (there are essentially no "always takers"). This argument shows that 2SLS generates a weighted average of lottery-specific effects with weights that depend on lottery offer probabilities and first-stage shifts in charter attendance, which in general does not correspond to the overall TOT or any other standard parameter.

charter schools.

7.1 Additional Assumptions

To focus attention on demand and student selection I make a set of simplifying assumptions about the supply side of the charter school market that allow simulation of counterfactuals. The supply side 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 in the baseline simulation, I use the addresses of new campuses that opened through 2013, after the application data used here were collected (see Appendix Table A1). Locations for further expansions are based on 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; estimates of this model appear in Appendix Table A8. 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. Locations in the other three simulations are chosen based on zip code averages of $X_i'(\alpha_c^x - \alpha_0^x)$, the component of charter school effects explained by observables.

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 set equal to the mean capacity for existing schools. Appendix E describes the details of the equilibrium and the methods used to compute counterfactual admission probabilities.

The average test score and utility parameters for new schools are set equal to the estimated means of γ_k^0 and α_k^0 from Table 6. This is effectively a constant returns to scale assumption implying that charter schools will remain equally productive as the system expands. There are several reasons 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 play a role in charter effectiveness, these effects may be diluted in expansions that draw in less positively selected students.[22] Together, these factors seem likely to reduce the efficacy of charter schools at larger scales. In this case the simulation results may be viewed as upper bounds on the effects of charter expansion determined by demand-side behavior.

[22]Existing evidence suggests that peer effects are not an important part of the achievement effects generated by charter schools. Angrist et al. (2013a) show that charter lottery impacts are unrelated to the mean change in peer quality resulting from lottery admission. This finding is replicated in Appendix Figure A2 for the data used here.

7.2 Charter Expansion Effects

Figure 3 summarizes the counterfactual simulations. The outcomes of interest are school choices, charter oversubscription, average 8th-grade math 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 shows how charter application and attendance rates change as the charter sector expands in the baseline simulation, while Panel B displays effects on admission probabilities and school capacity utilization. Panel C reports the effect of treatment on the treated in each simulation.

To focus on marginal students drawn into the charter sector by expansion, Panel D also plots a variant of the Marginal Treatment Effect (MTE) parameter of Heckman and Vytlacil (1999) for students approximately indifferent between charter and traditional public schools. For students receiving at least one charter offer, let

$$k^*(i) = \arg \max_{k \in \mathcal{C}(Z_i), k \neq 0} U_{ik}$$

denote student i's preferred charter school among those from which she received offers. Define

$$MTE(\Delta) = E\left[Y_i(k^*(i)) - Y_i(0) \,\middle|\, |U_{ik^*(i)} - U_{i0}| \leq \Delta, \mathcal{C}(Z_i) \neq \{0\}\right]. \quad (9)$$

For small Δ, $MTE(\Delta)$ gives the average effect of charter attendance for students approximately indifferent between charter and traditional public schools.[23] Since receiving an offer requires applying to a charter school, MTE captures causal effects for students who are willing to apply to charters and are on the margin of deciding to attend. Figure 3 reports MTEs setting Δ equal to a tenth of a standard deviation of the difference $|U_{ik^*(i)} - U_{i0}|$ among applicants in the current charter system.

The results for the baseline simulation imply that demand for charter schools in Boston may be quickly exhausted as the system expands. Panel B 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 when the number of schools moves beyond 15. In a setting with 20 charter schools, almost all 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 13 percent of charter seats are empty. This pattern is driven by the large application cost and negative average utilities reported in tables 4 through 6. Panels C and D shows that average and marginal treatment effects increase with the size of the charter sector, a consequence of the "reverse Roy" pattern documented in Section 6: expansion draws in students with weaker tastes for charter schools, who experience larger achievement gains. This implies that charter expansion produces large effects for marginal students, but the combination of a rising MTE and weak demand indicates that many high-benefit students choose to remain in traditional public schools even when charter seats are widely available.

Results from the geographic targeting counterfactual show that choosing new charter locations based on observed characteristics of nearby students is unlikely to meaningfully alter the effectiveness of charter

[23] See Heckman et al. (forthcoming) for discussion of related treatment effect concepts in dynamic discrete choice models.

expansion. Attendance rates in the baseline and geographically targeted counterfactuals are indistinguishable, and treatment effects are also very similar. This is a consequence of the fact that targeting locations based on observables does not change the pattern of self-selection. Opening a new charter school in a location near students with high predicted benefits draws in a few of these students, but does not induce attendance from students with weak unobserved tastes. As a result, most high-benefit students continue to remain in traditional public schools.

Counterfactuals that alter the pattern of self-selection into charter schools generate more profound changes. The reduced cost simulation eliminates application costs and average attendance disutilities, increasing overall charter demand by construction. Treatement effects are substantially larger in this counterfactual than the baseline counterfactual for all sizes of the charter sector. More students are willing to attend charter schools when the mean utility of doing so is higher, leading to less severe self-selection and therefore higher average test score gains. This finding suggests that policies that boost overall demand, such as providing information about charter schools more widely, are likely to boost average charter achievement effects as well.

Finally, Panel D shows that expansions targeting students with weak preferences would further increase charter productivity. In addition to eliminating application and mean utility costs, the altered preference counterfactual truncates the distribution of the charter preference $\mathcal{P}_i$ from above at the median, inducing students who currently dislike charter schools to behave like the median student. The results here may be viewed as the effects of outreach efforts attracting students who are especially unlikely to attend; for example, if the unobserved taste θ_i partly captures differences in knowledge across households, this policy may approximate the effects of targeting information specifically to families that lack it. TOTs in this counterfactual are larger than corresponding effects for the reduced cost counterfactual. MTEs are even larger, a consequence of weaker average tastes among marginal students than among inframarginal charter enrollees. Marginal students in the 20-school expansion gain nearly 0.7σ, an effect only slightly smaller than the effect of treatment on non-treated students in the current system. Together, the findings in Figure 3 suggest that reforms aimed at changing self-selection into charter schools have the potential to boost achievement impacts substantially more than reforms that merely add more seats or target locations for new schools based on observables.

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 among Boston students 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 best-case scenarios 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 broad implications for the design of school-choice programs. The introduction of a high-quality educational program without commensurate outreach efforts may not induce disadvantaged students to participate, even if the benefits from doing so are especially large for such students. In Boston, New York and many other cities, decentralized charter school application systems require parents to take steps outside of the usual school choice process, a possible source of logistical barriers for some families. Integrating charter schools into centralized school choice plans (as is done in Denver and New Orleans, for example) may reduce these barriers. More generally, my results suggest that efforts to target students who are otherwise unlikely to participate in school choice programs may yield high returns.

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 here 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. The mechanisms through which parents form preferences over schools are an important topic for future work.

Figure 1: Sequence of Events

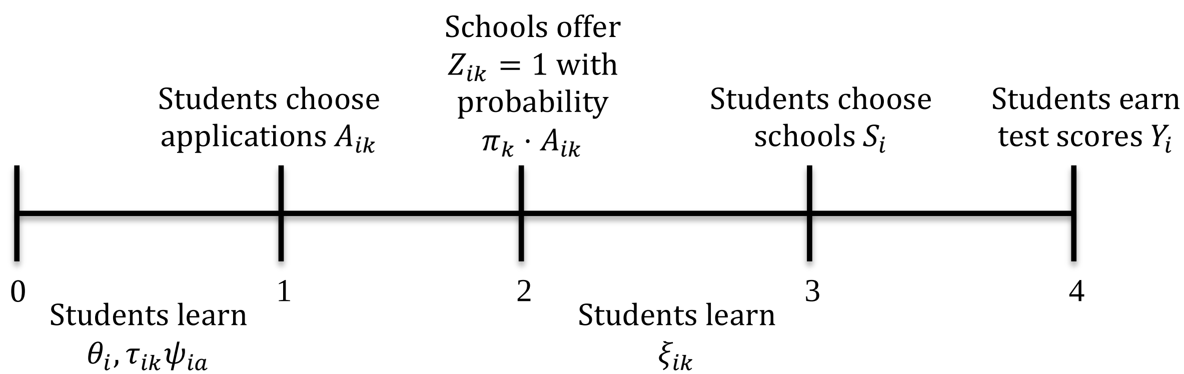
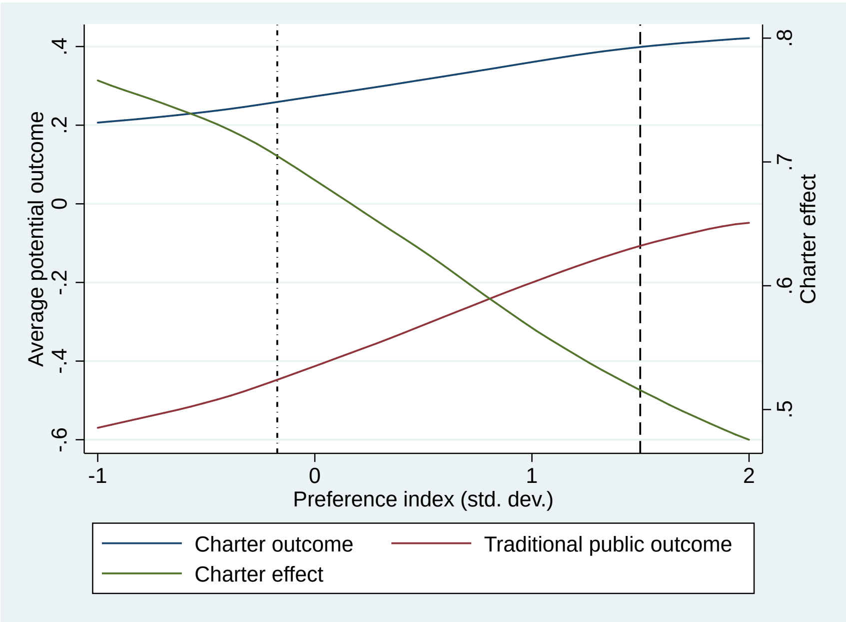
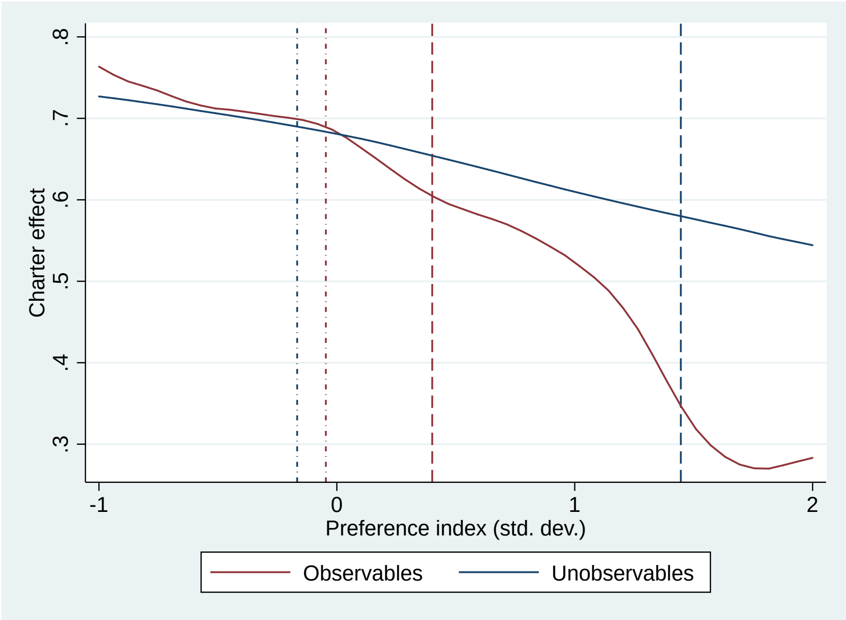


Figure 2: Absolute and Comparative Advantage in Charter School Choice



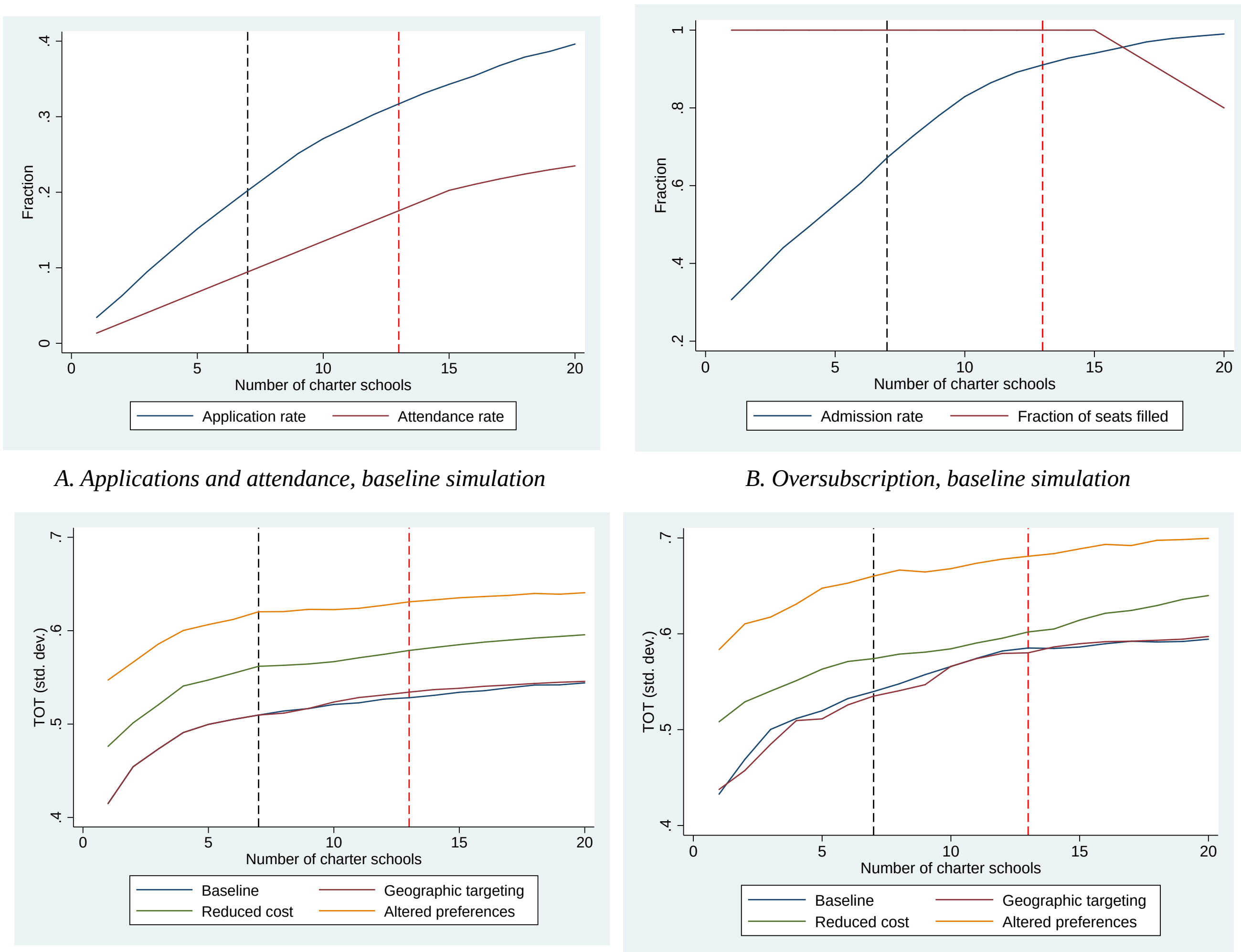
A. Selection on levels and gains



B. Selection on observables and unobservables

Notes: This figure displays relationships between preferences for charter school attendance and outcome levels and gains. Panel A plots conditional expectations of potential outcome levels in charter and traditional public schools (left axis) and causal effects of charter attendance (right axis) as functions of a charter preference index that combines observable and unobservable student characteristics. Panel B plots conditional expectations of charter effects as functions of two separate preference indices based on observable and unobservable characteristics. Preference indices are standardized to have mean zero and standard deviation one in the population. All conditional expectations are estimated via local linear regressions in a data set of 10,000 individuals simulated from the structural model, using a Gaussian kernel and rule-of-thumb bandwidth. Covariates and spatial locations are obtained by sampling with replacement from the observed data. Dashed lines indicate mean preference indices for students attending charter schools, and dotted lines show mean indices for traditional public students. Blue and red lines in Panel B correspond to mean observable and unobservable preference indices by charter attendance.

Figure 3: Simulated Effects of Charter School Expansion



A. Applications and attendance, baseline simulation

B. Oversubscription, baseline simulation

C. Effects of treatment on the treated

D. Marginal treatment effects

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. The baseline simulation adds charter schools at planned locations, then in zip codes with high predicted charter entry probabilities. The geographic targeting simulation adds new charter schools in zip codes with the highest predicted average treatment effects as a function of observables. The reduced cost simulation sets application costs and the charter mean utility to zero. The altered preference counterfactual truncates charter preferences from above at the median. Effects of treatment on the treated are average effects of charter attendance for students who attend charter schools in each counterfactual. Marginal treatment effects are average effects of charter attendance for students indifferent between attending and not attending charter schools in each counterfactual. Results are based on 5,000 simulations of the structural model.

Table 1: Descriptive Statistics

	All students (1)	Charter applicants (2)
Applied to charter school	0.175	-
Applied to more than one charter school	0.046	0.265
Received charter offer	0.125	0.718
Attended charter school	0.112	0.600
Female	0.492	0.490
White	0.142	0.165
Hispanic	0.398	0.317
Subsidized lunch	0.821	0.723
Special education	0.226	0.170
Limited English proficiency	0.212	0.136
Miles to closest charter school	2.105	1.859
Value-added of public schools in zip code	-0.051	-0.049
4th grade math score	-0.520	-0.314
4th grade ELA score	-0.636	-0.413
N	9156	1601

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 without 8th grade test scores. Column (1) shows means for the full sample, while column (2) shows means for charter applicants.

Table 2: Instruments for Charter School Applications and Attendance

		Balance check: 4th-grade scores		2SLS comparison			
		Math	ELA		First stage	Math 2SLS	ELA 2SLS
Controls		(1)	(2)	Instrument	(3)	(4)	(5)
None		0.038***	0.048***	Lottery	0.674***	0.573***	0.484***
		(0.010)	(0.010)		(0.022)	(0.078)	(0.083)
	N	9156				1601	
Baseline characteristics		0.012	0.020**	Distance	-0.026***	0.427*	0.282
		(0.008)	(0.008)		(0.003)	(0.237)	(0.239)
	N	9156				9156	

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 lottery offers and distance to the closest charter as instruments for charter attendance in equations for 8th-grade test scores. The lottery models are restricted to applicants and control for application portfolio indicators. The distance models include all students and control for student characteristics and 4th grade test scores. Outcomes are 8th grade test scores. Robust standard errors in parentheses.
*significant at 10; **significant at 5%; ***significant at 1%

Table 3: 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 4: Utility Parameter Estimates

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 students with 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 covariates. 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: Achievement Parameter Estimates

		Charter school		Traditional public school		Charter effect	
		Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Parameter	Description	(1)	(2)	(3)	(4)	(5)	(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 covariates. The mean potential outcome for charter schools is a weighted average of school-specific means. 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: School-specific Parameter Estimates

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: Comparison of Reduced Form and Structural Approaches to Extrapolation

		2SLS estimate	Covariate-based extrapolation			Model-based extrapolation
			Subsidized lunch	Baseline score	Interacted covs.	
Subject	Parameter	(1)	(2)	(3)	(4)	(5)
Math	TOT	0.573	0.584	0.600	0.606	0.505
	TNT	0.573	0.622	0.632	0.641	0.723
ELA	TOT	0.484	0.493	0.465	0.486	0.421
	TNT	0.484	0.520	0.528	0.541	0.608

Notes: This table compares charter school treatment effects obtained by covariate-based reweighting of lottery estimates vs. prediction from the structural model. Columns (1) through (4) are based on 2SLS models estimated in the lottery sample. These models interact charter school attendance with observed covariates, instrumenting with interactions of the lottery offer and covariates and controlling for main effects and application portfolio indicators. The TOT rows use the resulting coefficients to predict effects for charter students, and the TNT rows predict effects for non-charter students. Column (1) reports results with no interactions, while column (2) uses subsidized lunch status and column (3) uses terciles of baseline test score. Column (4) interacts charter attendance with all combinations of income, race, baseline score tercile and special education status. Column (5) reports predicted effects based on 5,000 simulations of the structural model.

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. Abdulkadiroglu, A., Angrist, J., Hull, P., and Pathak, P. (2014). "Charters Without Lotteries: Testing Takeovers in New Orleans and Boston." NBER Working Paper no. 20792.

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

5. Angrist, J. (1998). "Estimating the Labor Market Impact of Voluntary Military Service using Social Security Data on Military Applicants." *Econometrica* 66.

6. 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.

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

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

9. Angrist, J., and Fernandez-Val, I. (2010). "ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework." In: D. Acemoglu, M. Arellano, and E. Dekel, eds., Advances in Economics and Econometrics. Cambridge, UK: Cambridge University Press.

10. 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).

11. Angrist, J., Imbens, G., and Rubin, D. (1996). "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434).

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

13. Barseghyan, L., Clark, D., and Coate, S. (2014). "Public School Choice: An Economic Analysis." NBER Working Paper no. 20701.

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

15. Bettinger, E., Long, B., Oreopolous, P., and Sanbonmatsu, L. (2012). "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment." *Quarterly Journal of Economics* 127 (3).

16. 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).

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

18. Bourguignon, F., Fournier, M., and Gurgand, M. (2007). "Selection Bias Corrections Based on the Multinomial Logit Model: Monte Carlo Comparisons." *Journal of Economic Surveys* 21.

19. 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).

20. 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

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

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

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

24. 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.

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

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

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

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

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

30. 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).

31. 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).

32. Dobbie, W., and Fryer, R. (2015). "The Medium-Term Impacts of High-Achieving Charter Schools." *Journal of Political Economy* 123(5).

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

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

35. 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).

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

37. 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.

38. Epple, D., Romano, R., and Zimmer, R. (2015). "Charter Schools: A Survey of Research on Their Characteristics and Effectiveness." NBER Working Paper no. 21256.

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

40. Fryer, R. (2011). "Creating 'No Excuses' (Traditional) Public Schools: Preliminary Evidence from an Experiment in Houston." NBER Working Paper no. 17494.

41. 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.

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

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

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

45. Heckman, J. (1979). "Sample Selection as a Specification Error." *Econometrica* 47(1).

46. Heckman, J. (1990). "Varieties of Selection Bias." *The American Economic Review* 80(2).

47. Heckman, J., Humphries, J., and Veramendi, G. (forthcoming). "Dynamic Treatment Effects." *Journal of Econometrics.*

48. Heckman, J., and Vytlacil, E. (1999). "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences* 96(8).

49. Heckman, J., and Vytlacil, E. (2005). "Structural Equations, Treatment Eects, and Econometric Policy Evaluation." *Econometrica* 73.

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

51. Hoxby, C. (1998). "The Economics of School Reform." NBER Reporter.

52. Hoxby, C. (2003). "School Choice and School Productivity: Could School Choice Be a Tide that Lifts All Boats?" In: Hoxby, C., ed., The Economics of School Choice. NBER conference volume. Chicago, IL: University of Chicago Press.

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

54. 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.

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

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

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

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

59. Ladd, H. (2002). Market-Based Reforms in Urban Education. Washington, DC: Economic Policy Institute.

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

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

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

63. 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.

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

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

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

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

68. 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.

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

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

71. 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.

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

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

74. 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 (1) through (5) 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 (1) 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 (4) and (5).

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 (4) and (5). 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

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

[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.

The model for potential outcomes in charter and public school is

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

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

with $E[\epsilon_{ik}|\theta_i, D_i] = 0$ for $k \in \{0,1\}$. 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

$$\frac{E[Y_i(k)|A_i = 1, S_i(1) > S_i(0), D_i = d_1] - E[Y_i(k)|A_i = 1, S_i(1) > S_i(0), D_i = d_0]}{\mu_\theta^c(d_1) - \mu_\theta^c(d_0)} = \alpha_k^\theta$$

for $k \in \{0,1\}$. 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 conditional application and attendance probabilities. The selection parameters α_k^θ are therefore identified.

Appendix C: Model Fit

The model estimated here fits the data well. This can be seen in Table A7 and Figure A1, which compare observed and model-predicted patterns of heterogeneity across choices, outcomes, and schools. Panel A of Figure 3 splits the sample into vingtiles based on the model-predicted probability of applying to at least one charter school as a function of observed characteristics and distance. The horizontal axis plots mean predicted application probabilities in these cells, while the vertical axis displays empirical application probabilities. These points lie mostly along the 45 degree line, indicating that the model accurately reproduces differences in application probabilities across groups; the predicted probabilities range from 0.12 to 0.35, implying that the model captures a substantial amount of heterogeneity in preferences explained by observables. There is slight visual evidence of nonlinearity and an F-test marginally rejects the null hypothesis that all points lie exactly on the line ($p = 0.07$), but in general the model appears to provide a relatively good fit.

To assess whether the model captures heterogeneity in outcomes, Panel B of Figure A1 compares model-predicted and observed mean test scores in vingtiles of model predictions, separately for charter and non-charter students. Model-predicted outcomes are expected eighth-grade math scores conditional on a student's observed characteristics and choices, which implicitly incorporates heterogeneity on both observed and unobserved dimensions. The fit here is very good: most points lie close to the 45-degree line and an F-test does not reject the hypothesis that the model fits perfectly. Predicted scores exhibit substantial dispersion and there is significant overlap between predictions for charter and non-charter students, suggesting that the model captures a lot of variation in outcomes.

Finally, Table A7 explores the model's capacity to match cross-school heterogeneity in choices and treatment effects. Panel A reports model-based and observed application probabilities for each school while Panel B displays differences in outcomes for lottery winners and losers by school. The model slightly over-predicts the overall application rate and under-predicts scores for lottery winners, but these differences are small and the predicted patterns of heterogeneity across schools appear to accurately reflect the observed differences. As in Panel A of Figure A1 the hypothesis that the model fits all moments perfectly is rejected, but overall the model appears to generate a parsimonious yet accurate description of heterogeneity along many dimensions.

Appendix D: 2SLS Weights

This appendix derives the estimand in 2SLS models of the type estimated by Abdulkadiroglu et al. (2011) and other lottery-based studies of school choice programs. Consider the system

$$Y_i = \alpha + \beta C_i + \varphi_{a(i)} + \epsilon_i,$$

$$C_i = \delta + \tau \tilde{Z}_i + \kappa_{a(i)} + \eta_i,$$

where Y_i is a test score, C_i is a charter attendance dummy, $\tilde{Z}_i$ is a lottery offer dummy, and $\varphi_{a(i)}$ and $\kappa_{a(i)}$ are application portfolio (risk set) fixed effects. The reduced form corresponding to this system is

$$Y_i = \zeta + \rho \tilde{Z}_i + \iota_{a(i)} + u_i.$$

The reduced form and first stage are OLS regressions of test scores and charter attendance on the lottery offer with saturated portfolio controls. These equations therefore generate inverse-variance weighted averages of within-portfolio mean differences (Angrist, 1998). Specifically, let ρ_a and τ_a denote coefficients from regressions of Y_i and C_i on $\tilde{Z}_i$ within lottery portfolio a. Then

$$\rho = \sum_a \left(\frac{w_a}{\sum_{a'} w_{a'}} \right) \rho_a,$$

$$\tau = \sum_a \left(\frac{w_a}{\sum_{a'} w_{a'}} \right) \tau_a,$$

where

$$w_a = Pr\left[A_i = a\right] Pr\left[\tilde{Z}_i = 1 | A_i = a\right] \left(1 - Pr\left[\tilde{Z}_i = 1 | A_i = a\right]\right).$$

Since the 2SLS model is just-identified, the 2SLS estimand β is equal to the ratio of reduced form and first stage coefficients. This implies:

$$\begin{aligned}
\beta &= \frac{\rho}{\tau} \\
&= \frac{\sum_a w_a \rho_a}{\sum_a w_a \tau_a} \\
&= \frac{\sum_a (w_a \tau_a)(\rho_a / \tau_a)}{\sum_a w_a \tau_a} \\
&= \sum_a \left(\frac{\omega_a}{\sum_{a'} \omega_{a'}} \right) \beta_a,
\end{aligned}$$

where $\beta_a = (\rho_a / \tau_a)$ is a portfolio-specific IV coefficient and

$$\begin{aligned}
\omega_a &= w_a \tau_a \\
&= Pr\left[A_i = a\right] Pr\left[\tilde{Z}_i = 1 | A_i = a\right] \left(1 - Pr\left[\tilde{Z}_i = 1 | A_i = a\right]\right)
\end{aligned}$$

$$\times \left(E\left[C_i | \tilde{Z}_i = 1, A_i = a \right] - E\left[C_i | \tilde{Z}_i = 0, A_i = a \right] \right).$$

This argument shows that 2SLS estimation with application portfolio fixed effects generates a weighted average of portfolio-specific IV coefficients with weights proportional to sample size, the variance of the lottery offer, and the first stage shift in charter attendance resulting from the offer. These weights are similar to the weights derived in Angrist and Imbens (1995) for 2SLS models with saturated instrument-covariate interactions in the first stage. A saturated weighting scheme generates weights proportional to $w_a \tau_a^2$ rather than $w_a \tau_a$.

Appendix E: 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

$$\max_{\pi_k\in[0,1]} e_k(\pi_k, \pi_{-k}, \pi^e) \quad s.t. \quad e_k(\pi_k, \pi_{-k}, \pi^e) \leq \Lambda_k. \tag{10}$$

The best response function $\pi_k^{BR}(\pi_{-k}, \pi^e)$ is the solution to problem (10). 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 $\pi \in [0,1]^K$:

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

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

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

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

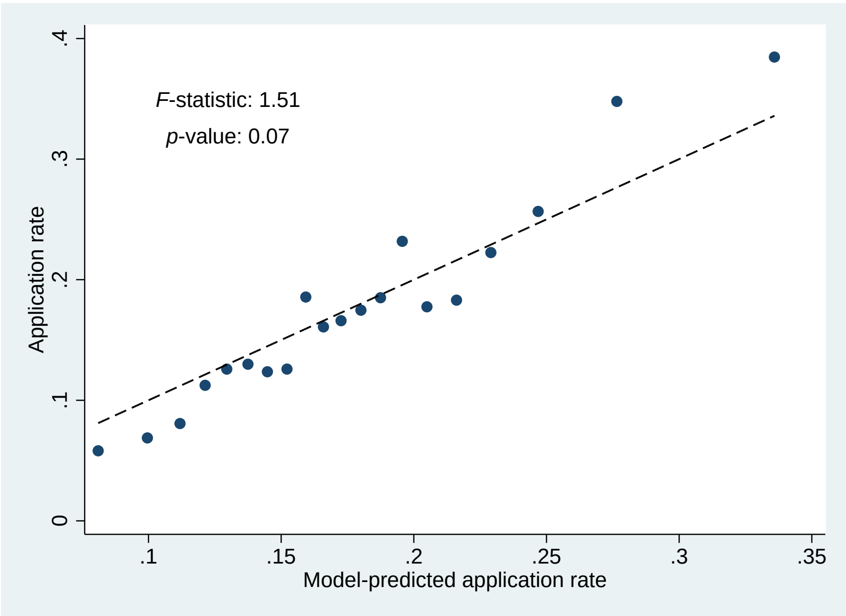
.It follows that conditions 1a and 2a are equivalent to the following conditions on the model's enrollment elasticities:

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

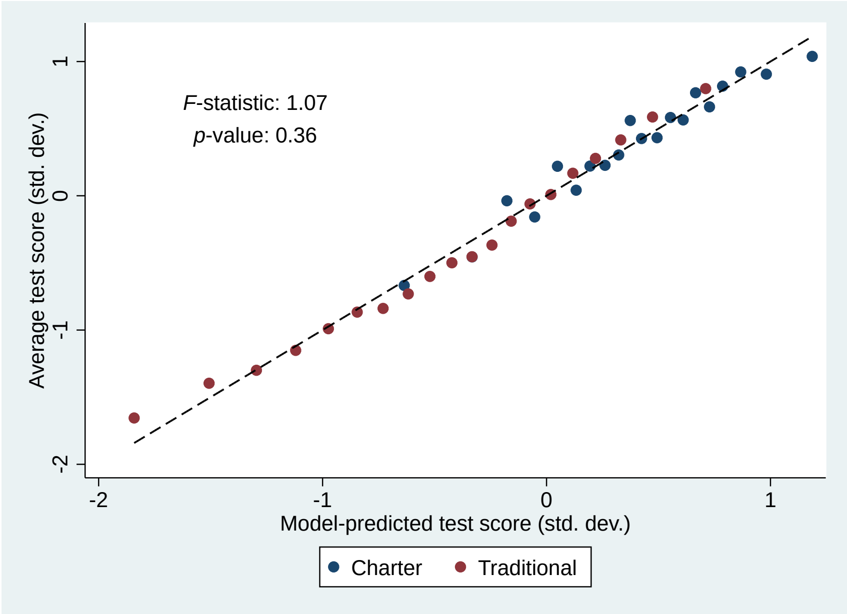
$$\text{2b. } \frac{\partial \log e_k}{\partial \log \pi_k} \geq \sum_{j \neq k} \frac{\pi_k}{\pi_j} \cdot \left|\frac{\partial \log e_k}{\partial \log \pi_j}\right| + \left(\frac{\Lambda_k - e_k}{\Lambda_k}\right) \quad \forall k$$

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: Model Fit



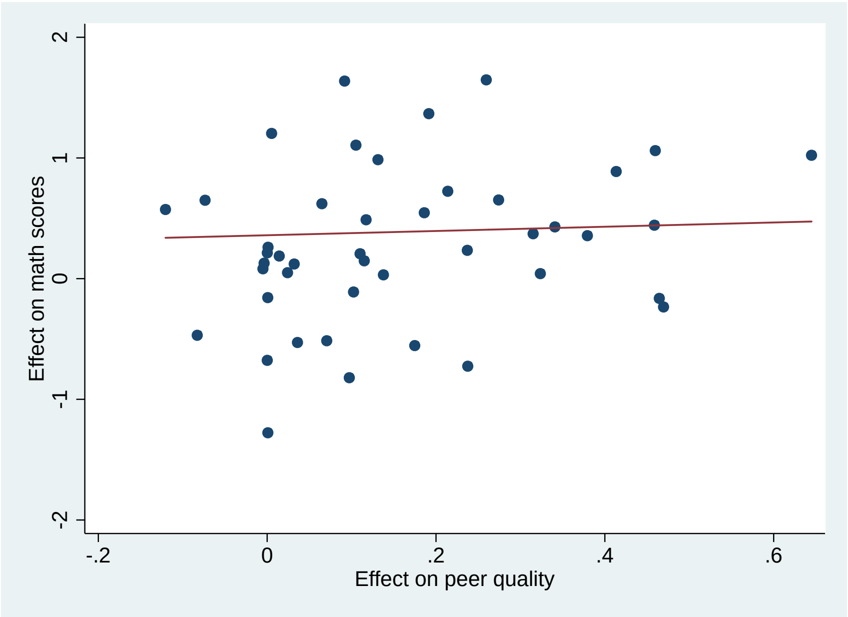
A. Application probability



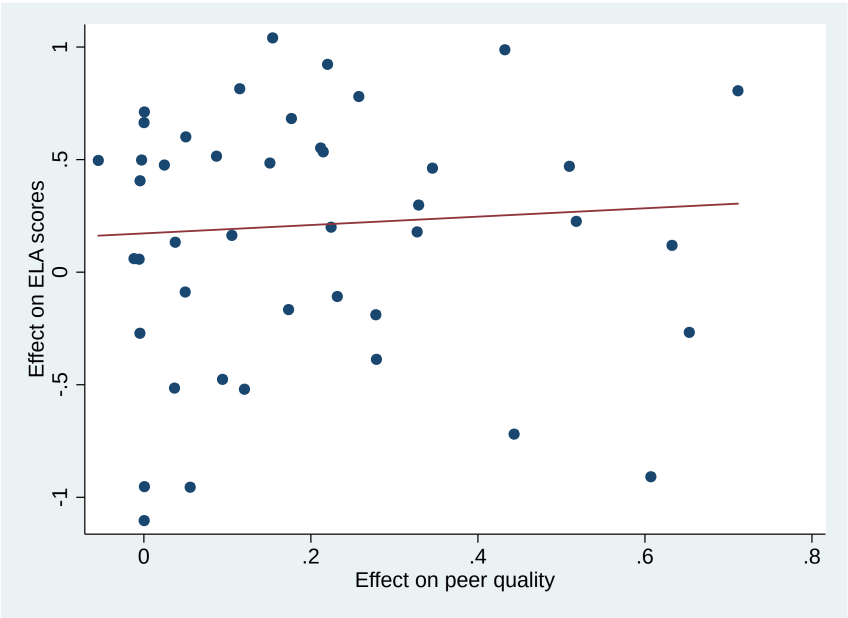
B. Test scores

Notes: This table compares charter application rates and mean test scores to predictions from the model. Panel A splits the sample into vingtiles of the model-predicted probability of applying to at least one charter school. Points on the vertical axis are mean observed application rates in these bins, while points on the horizontal axis are means of model-predicted rates. Panel B computes mean model-predicted 8th-grade math scores conditional on each student's observed school choice. The sample is split into vingtiles of this predicted score separately for charter and traditional public schools. Blue points plot mean observed scores against mean model predictions for charter schools, while red points plot corresponding observed and predicted means for traditional public schools. Predictions are averages over 5,000 simulations of the structural model, with covariates and spatial locations drawn with replacement from the joint distribution in the sample. Dashed lines show the 45-degree line. *F*-statistics and *p*-values are from tests of the hypothesis that this lines fit all points perfectly up to sampling error, treating the model predictions as fixed.

Figure A2: 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	-0.019
	(0.033)
White	-0.007
	(0.022)
Hispanic	0.008
	(0.028)
Subsidized lunch	-0.004
	(0.029)
Special education	-0.001
	(0.025)
Limited English proficiency	-0.006
	(0.020)
Miles to closest charter school	-0.033
	(0.070)
Value-added of public schools in zip code	0.002
	(0.003)
4th grade math score	-0.041
	(0.064)
4th grade ELA score	0.038
	(0.067)
Joint *p*-value	0.933
N	1601

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

	Full sample (1)	Lottery applicants (2)
Followup rate	0.848	0.806
Difference by predicted score	0.055***	0.062
	(0.021)	(0.047)
Difference by lottery win/loss	-	-0.016
		(0.046)
Interaction between win/loss and predicted score	-	0.015
		(0.053)
N	10797	1986

Notes: This table reports the fraction of follow-up test scores in 8th grade 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. Column (1) shows the follow-up rate for the ful sample as well as the difference in followup rates between students with above-median and below-median predicted 8th-grade math scores. Predicted scores are fitted values from regressions of 8th grade math scores on the baseline variables from Table 1. Column (2) shows the followup rate for lottery applicants along with coefficients from a regression of a followup indicator on the lottery offer, an indicator for an above-median predicted score, and the interaction of the two, controlling for risk set indicators. Robust standard errors in parentheses.
*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: Estimates of Math Achievement Parameters by Grade

		6th grade			7th grade			8th grade		
		Charter	Traditional	Charter effect	Charter	Traditional	Charter effect	Charter	Traditional	Charter effect
Parameter	Description	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
α^0	Mean potential outcome	0.028	-0.478***	0.506***	0.219***	-0.456***	0.676***	0.272***	-0.367***	0.639***
		(0.069)	(0.007)	(0.069)	(0.075)	(0.008)	(0.075)	(0.080)	(0.009)	(0.081)
α^x	Female	-0.006	-0.005	-0.001	0.116***	0.010	0.105***	0.032	0.025	0.057
		(0.034)	(0.013)	(0.036)	(0.038)	(0.015)	(0.040)	(0.040)	(0.016)	(0.043)
	Black	-0.104	-0.194***	0.090	-0.036	-0.181***	0.150**	0.048	-0.166***	0.214***
		(0.059)	(0.021)	(0.062)	(0.068)	(0.023)	(0.071)	(0.066)	(0.026)	(0.071)
	Hispanic	0.036	-0.100***	0.136***	0.033	-0.083***	0.116	0.135***	-0.093***	0.228***
		(0.062)	(0.022)	(0.066)	(0.070)	(0.025)	(0.074)	(0.068)	(0.026)	(0.073)
	Subsidized lunch	-0.022	-0.147***	0.125***	0.036	-0.150***	0.186***	0.046	-0.120***	0.165***
		(0.041)	(0.019)	(0.044)	(0.044)	(0.021)	(0.049)	(0.049)	(0.023)	(0.054)
	Special education	-0.362***	-0.354***	-0.008	-0.374***	-0.350***	-0.024	-0.288***	-0.374***	0.086
		(0.040)	(0.015)	(0.042)	(0.048)	(0.017)	(0.051)	(0.052)	(0.019)	(0.055)
	Limited English proficiency	-0.014	-0.049***	-0.066	0.005	0.083***	-0.078	-0.025	0.076**	-0.101
		(0.045)	(0.016)	(0.048)	(0.058)	(0.019)	(0.061)	(0.065)	(0.020)	(0.068)
	Baseline math score	0.402***	0.566***	-0.164***	0.367***	0.494***	-0.127***	0.358***	0.476***	-0.118***
		(0.024)	(0.009)	(0.026)	(0.028)	(0.010)	(0.029)	(0.030)	(0.011)	(0.032)
	Baseline ELA score	0.144***	0.101***	0.043**	0.076***	0.091***	-0.019	0.046	0.065***	-0.019
		(0.022)	(0.008)	(0.024)	(0.027)	(0.009)	(0.028)	(0.029)	(0.010)	(0.030)
	Value-added of public schools in zip code	0.359	0.953***	-0.594	0.422	0.928***	-0.506	0.479	0.943***	-0.464
		(0.412)	(0.162)	(0.449)	(0.421)	(0.168)	(0.457)	(0.430)	(0.171)	(0.462)
$\alpha^\theta \times \sigma_\theta$	Taste for charter schools (std. dv. units)	0.039	0.133***	-0.094***	0.029	0.061***	-0.032	0.010	0.100***	-0.090***
		(0.019)	(0.009)	(0.021)	(0.022)	(0.011)	(0.025)	(0.023)	(0.012)	(0.026)
N	Sample size		10122			9731			9156	

Notes: This table reports maximum simulated likelihood estimates of the parameters of the math achievement distribution. The likelihood is evaluated using 300 simulations per observation. Covariates are de-meaned in the estimation sample. The mean potential outcome for charter schools is a weighted average of school-specific means. The model is reestimated for each grade. The likelihood is evaluated using 300 simulations per observation.
*significant at 10; **significant at 5%; ***significant at 1%

Table A6: Estimates of ELA Achievement Parameters by Grade

Parameter	Description	6th grade Charter (1)	6th grade Traditional (2)	6th grade Charter effect (3)	7th grade Charter (4)	7th grade Traditional (5)	7th grade Charter effect (6)	8th grade Charter (7)	8th grade Traditional (8)	8th grade Charter effect (9)
α^0	Mean potential outcome	-0.085	-0.557***	0.472***	0.042	-0.517***	0.559***	0.138	-0.490***	0.627***
		(0.076)	(0.008)	(0.077)	(0.091)	(0.009)	(0.092)	(0.088)	(0.010)	(0.089)
α^x	Female	0.095**	0.158***	-0.063	0.170***	0.221***	-0.051	0.164***	0.183***	-0.019
		(0.038)	(0.013)	(0.040)	(0.041)	(0.015)	(0.044)	(0.043)	(0.016)	(0.046)
	Black	-0.078	-0.164***	0.084	0.003	-0.084***	0.087	0.101	-0.058**	0.159**
		(0.060)	(0.021)	(0.063)	(0.068)	(0.024)	(0.072)	(0.068)	(0.028)	(0.074)
	Hispanic	-0.052	-0.098***	0.047	0.088	-0.023	0.111	0.177**	-0.031	0.209***
		(0.062)	(0.022)	(0.066)	(0.073)	(0.025)	(0.077)	(0.070)	(0.028)	(0.076)
	Subsidized lunch	-0.063	-0.150***	0.087	0.029	-0.130***	0.159***	0.029	-0.116***	0.146***
		(0.045)	(0.020)	(0.049)	(0.051)	(0.022)	(0.055)	(0.051)	(0.025)	(0.056)
	Special education	-0.331***	-0.332***	0.001	-0.412***	-0.404***	-0.008	-0.265***	-0.398***	0.133***
		(0.042)	(0.015)	(0.045)	(0.052)	(0.016)	(0.054)	(0.051)	(0.018)	(0.055)
	Limited English proficiency	-0.009	-0.069***	0.060	-0.034	-0.014	-0.020	-0.023	0.044**	-0.067
		(0.048)	(0.016)	(0.051)	(0.061)	(0.019)	(0.064)	(0.063)	(0.020)	(0.066)
	Baseline math score	0.076***	0.183***	-0.107***	0.101***	0.180***	-0.078***	0.119***	0.366***	-0.083**
		(0.026)	(0.009)	(0.027)	(0.030)	(0.009)	(0.031)	(0.030)	(0.010)	(0.033)
	Baseline ELA score	0.488***	0.452***	0.036	0.331***	0.370***	-0.039	0.284***	0.164***	-0.044
		(0.025)	(0.008)	(0.026)	(0.029)	(0.009)	(0.030)	(0.031)	(0.011)	(0.032)
	Value-added of public schools in zip code	0.712	0.862***	-0.150	0.661	0.914***	-0.253	0.634	0.893***	-0.259
		(0.452)	(0.171)	(0.439)	(0.473)	(0.179)	(0.444)	(0.466)	(0.185)	(0.501)
$\alpha^\theta \times \sigma_\theta$	Taste for charter schools (std. dv. units)	-0.032	0.009	-0.041**	-0.025	-0.002	-0.024	-0.037	0.027**	-0.064**
		(0.021)	(0.009)	(0.023)	(0.027)	(0.012)	(0.030)	(0.026)	(0.013)	(0.029)
N	Sample size		10122			9731			9156	

Notes: This table reports maximum simulated likelihood estimates of the parameters of the ELA achievement distribution. The likelihood is evaluated using 300 simulations per observation. Covariates are de-meaned in the estimation sample. The mean potential outcome for charter schools is a weighted average of school-specific means. The model is reestimated for each grade. The likelihood is evaluated using 300 simulations per observation.

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

Table A7: Model Fit

A. Applications and attendance			*B. 8th grade math scores*				
				Lottery winners		Lottery losers	
	Model	Data		Model	Data	Model	Data
	(1)	(2)		(3)	(4)	(5)	(6)
Apply to charter	0.202	0.175	Charter applicants	0.169	0.236	-0.172	-0.151
Apply to more than one	0.034	0.046	Non-applicants	-	-	-0.450	-0.453
Offer takeup rate	0.781	0.836	Applicants to:				
Apply to:			Charter 1	0.264	0.331	-0.064	0.050
Charter 1	0.033	0.031	Charter 2	0.329	0.405	-0.121	0.110
Charter 2	0.039	0.036	Charter 3	0.091	0.131	-0.084	0.013
Charter 3	0.048	0.048	Charter 4	0.271	0.401	-0.281	-0.207
Charter 4	0.043	0.039	Charter 5	0.084	0.017	-0.103	-0.145
Charter 5	0.049	0.051	Charter 6	0.217	0.241	-0.101	-0.073
Charter 6	0.023	0.017	Charter 7	0.319	0.443	-0.088	0.088
Charter 7	0.022	0.016					
F-statistic:				1.68			
p-value:				0.04			

Notes: This table compares behavior and outcomes to predictions from the model. Panel A shows charter application and offer takeup rates. Panel B shows average test scores by application and offer status. Model predictions are averages over 5,000 simulations of the structural model, with covariates and spatial locations drawn with replacement from the joint distribution in the sample. The *F*-statistic and *p*-value come from a Wald test of the hypothesis that the model fits all observed moments up to sampling error.

Table A8: 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%