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, Aviva Aron-Dine, David Autor, 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, Parag Pathak, Jesse Rothstein, Stephen Ryan, Steven Stern, Xiao Yu Wang, Tyler Williams, and 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, The Stanford Institute for Theoretical Economics, UC Berkeley, UC Davis, UCLA, the University of Chicago Harris School, the University of Michigan, the University of Virginia, Warwick University, and WEAI for useful comments. 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 February 2017
JEL No. C25,I21,J2,J24

ABSTRACT

This paper studies the demand for charter schools in Boston, Massachusetts, with an emphasis on comparative advantage in school choice. I model charter school application and attendance decisions in a generalized Roy selection framework that links students' preferences to the achievement gains generated by charter attendance. I estimate the model using instruments based on randomized admission lotteries and distance to charter schools. The estimates show that students do not sort into charter schools on the basis of comparative advantage in academic achievement. Charter schools generate larger test score gains for disadvantaged, low-achieving students, but demand for charters is stronger among richer students and high achievers. Similarly, achievement benefits are larger for students with weaker unobserved preferences for charter schools. As a result, counterfactual simulations indicate that charter expansion is likely to be most effective when accompanied by efforts to target students who are currently unlikely to apply.

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 the U.S. education system. Examples include charter schools, vouchers, and district-wide choice plans allowing students to choose from menus 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 admission lotteries show that charters in Boston and New York boost academic achievement sharply (Abdulkadiroğlu et al., 2011; Dobbie and Fryer, 2011). Angrist et al. (2012, 2013, 2016), Dobbie and Fryer (2013, forthcoming), Gleason et al. (2010), Hoxby and Murarka (2009) and Hoxby and Rockoff (2004) also report positive effects of 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 selection into these programs.[2] Existing studies typically restrict attention to samples of lottery applicants, for whom admission offers are randomly assigned (see, e.g., Abdulkadiroğlu 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 (LATEs) derived from lottery-based instruments will overstate potential effects for non-applicants and provide a misleading picture of the impacts of charter expansion (Heckman et al., 2001).[3] If students with large potential

[1] Gleason et al. (2010) find that non-urban charters are no more effective than traditional public schools. Angrist et al. (2013) 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, and Ferreyra and Kosenok (2012) and Mehta (forthcoming), who develop equilibrium models of charter school entry and student sorting. Other related studies look at selection in higher education (Arcidiacono, 2005; Brand and Xie, 2010; Howell, 2010; Arcidiacono et al., 2016; Dillon and Smith, 2017) and in education programs outside the U.S. (Ajayi, 2013; Kirkebøen et al., 2016).

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

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

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 admission lottery at each charter. I analyze this process using a dynamic generalized Roy (1951) model that describes charter application portfolio choices, lottery offers, school attendance decisions, and test score outcomes.[4] The model is similar to the stochastic portfolio choice problems considered by Chade and Smith (2006) and Chade et al. (2014): students submit charter applications to maximize expected utility, taking account of admission probabilities and non-monetary application costs. As in Willis and Rosen's (1979) canonical analysis of education and self-selection, the model allows a link between outcomes and the unobserved preferences driving school choices, thereby creating scope for selection according to absolute and comparative advantage.

I estimate the model using instrumental variables (IVs) based on randomized charter admission lotteries and distance to charter schools. Lottery IV estimates identify local average treatment effects for selected sets of charter applicants, while distance shifts the composition of the applicant pool. I provide a semi-parametric identification argument showing that the combination of these two instruments allows generalization from lottery-based LATEs to causal parameters relevant for policies that expand charter schooling to new populations. Following Heckman (1979), I estimate the model using a two-step control function approach to correct for self-selection into charter application and enrollment.

Estimates of the model reveal that students do not sort into charter schools on the basis of comparative advantage in academic achievement. Instead, preferences for charter schools are weaker for students with larger test score benefits. 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 larger for students with weaker unobserved preferences for charter schools. I test and reject cross-equation restrictions implied by a model in which students choose schools to maximize test scores net of distance and application costs. These findings parallel results in the literature on female labor supply, which show negative associations between market wages and the propensity to work for some groups of women (Neal, 2004; Mulligan and Rubinstein, 2008). Recent studies of early childhood education programs also find negative selection on treatment effects (Cornelissen et al., 2016; Kline and Walters, 2016).

The results reported here imply that previous lottery-based studies understate the potential achievement effects of Boston's charter schools for non-applicants. Specifically, 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 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

[4] See Heckman et al. (2006), Heckman and Vytlacil (2007a), Heckman and Navarro (2007), and Heckman et al. (forthcoming) for analyses of static and dynamic generalized Roy models.

likely to seek out effective schools. I quantify the policy implications of this pattern by simulating charter expansion effects in an equilibrium school choice model. The simulations indicate substantial achievement impacts for marginal applicants, and show that charter expansion is likely to be most effective when targeted to students who are currently unlikely to apply.

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 model, and Section 4 discusses identification. Section 5 details the estimation procedure. Parameter estimates are reported in Section 6. Section 7 summarizes patterns of selection and comparative advantage in charter school choice, and compares these patterns to what might be learned from atheoretical extrapolation based on lottery applicants. Section 8 simulates the effects of counterfactual policies. Section 9 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 (Massachusetts Department of Elementary and Secondary Education, 2011). The Board of Education reviews each charter school's academic and organizational performance at five year intervals and decides whether charters should be renewed or revoked.

Enrollment in a Massachusetts charter school is open to all students who live in the local school district. If applications to a charter school exceed its seating capacity, the school must admit students by random lottery. Students interested in multiple charter schools submit a separate application 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 Sönmez, 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. Fourteen charter schools operated in Boston during the 2010-2011

[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 Abdulkadiroğlu et al. (2016) for a recent analysis of Horace Mann charters.

school year, 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. 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 admission 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, an emphasis 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., 2013; Dobbie and Fryer, 2013; Curto and Fryer, 2014; Fryer, 2014). Consistent with this evidence, Abdulkadiroğlu 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 boosts test scores for lottery applicants by 0.4 standard deviations (σ) in math and 0.2σ in reading.

The demand for charter schools in Boston is relevant to an ongoing policy debate. In recent years the growth of Massachusetts' charter sector has been slowed by the state's charter cap, a law that limits expenditures on charter tuition to 9 percent of the host district budget. The Board of Education stopped accepting proposals for new Boston charters after expenditure reached this cap in 2008 (Boston Municipal Research Bureau, 2008). A 2010 act of the Massachusetts legislature raised the cap to 18 percent of district spending for Boston and other low-performing districts. This reform led to the approval of six new charter middle schools (Commonwealth of Massachusetts, 2010). Massachusetts voters rejected a 2016 ballot measure that proposed a further increase in the charter school cap (Scharfenberg, 2016).

2.2 Data Sources and Sample Construction

The data used here come from three sources. Demographics, school attendance, and test scores are obtained 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 reading achievement tests. I begin by selecting from the database the four cohorts of students who attended a traditional BPS school in fourth grade between 2005-2006 and 2008-2009. Students must also have non-missing fourth grade demographics and test scores, as well as school attendance information and test scores in eighth grade. I retain information from the first time a student attempts a grade for students who repeat. Test scores are standardized to have mean zero and standard deviation one within each subject, year, and grade in Massachusetts. Students are coded as enrolled in a charter middle school if they attend the school at any time prior to the relevant test.[6]

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 fourth grade BPS address data. This restriction eliminates less than one percent of Boston fourth 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 of the nine Boston charter middle schools that operated between the 1997-1998 and 2010-2011 school years.[8] As shown in Appendix Table A1, the admissions data provide a complete record of applications to these seven schools for cohorts attending fourth grade between 2006 and 2009. Of the two schools without available records, one closed prior to the 2010-2011 school year; the other declined to provide records. The analysis below treats 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.[9] Not every charter school was oversubscribed in every year, so schools did not always hold lotteries. Column (5) of Table A1 shows that each of the seven sample schools held lotteries in at least two years. The analysis to follow sets admission probabilities to one for undersubscribed years.

2.3 Descriptive Statistics

The final analysis sample includes 9,156 students who attended BPS schools in fourth grade between 2006 and 2009. Descriptive statistics for this sample appear in Table 1. As shown in Panel A, 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. Panel B of Table 1 shows that 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. The last two rows of Table 1 report statistics for fourth grade math and reading test scores, normed to have mean zero and standard deviation one in the Massachusetts population. Boston fourth graders lag behind the state average by 0.52σ and 0.64σ in math and reading. Students who apply to charter schools score much higher than the overall Boston population: applicants' fourth grade scores exceed the city average by more than 0.2σ in

[6]School exit rates are similar for Boston traditional public schools and charter schools: the probability that a student remains in the same school from one middle school grade to the next is roughly 80 percent for both groups during my sample period.

[7]I also estimated models using travel times measured by Google Maps, obtained using the STATA *traveltime* command. Key estimates were similar for this alternative distance measure.

[8]Charter schools are classified as middle schools if they accept applicants in fifth or sixth grade. Two Boston charter schools accept students prior to fifth grade but serve grades six through eight. Since I restrict the analysis to students who attended traditional BPS schools in fourth grade, no students in the sample attend these schools.

[9]Most unmatched students are likely to be applicants who previously attended private schools and therefore lack earlier records in the state database. I exclude such students from the analysis by limiting the sample to students enrolled in BPS in fourth grade. Less than one percent of students in the remaining sample attend charter schools without an admission record, most likely because these students were unsuccessfully matched. These students are dropped in the analysis.

both subjects. Together, these statistics show that Boston's charter applicants are less disadvantaged and higher-achieving than other Boston students on several dimensions.

Panel C of Table 1 describes nearby middle school options for Boston students. The average student lives 2.1 miles from the nearest charter middle school and 0.5 miles from the nearest BPS district middle school. Charter applicants live closer to charter schools and farther from district schools than non-applicants, suggesting that distance may play a role in charter application decisions. The last row of Table 1 reports average value-added of the nearest BPS school, measured as the school average residual from a regression of sixth grade math scores on demographics and fourth grade scores for BPS students.[10] This metric may be viewed as a proxy for the quality of nearby traditional public school options. The average value-added of nearby BPS schools is slightly lower for charter applicants than for the full sample.

3 A Model of Charter School Choice and Academic Achievement

3.1 Setup

I model charter application choices as an optimal portfolio choice problem in which forward-looking students seek to maximize expected utility. Figure 1 explains the sequence of events described by the model. At stage one, students decide whether to apply to each of J charter schools, indexed by $j \in \{1 \ldots J\}$. The binary variable A_{ij} indicates that student i applies to school j, and the vector $A_i = (A_{i1}...A_{iJ})$ collects these indicators for all schools. In the second stage, charter school j randomly assigns offers to its applicants with probability π_j. The binary variable Z_{ij} indicates an offer for student i at school j, and $Z_i = (Z_{i1}...Z_{iJ})$ collects offers. Third, students choose schools denoted $S_i \in \{0, 1, ..., J\}$, where $S_i = 0$ indicates traditional public school attendance. Any student can attend a traditional public school, but student i can attend charter school j only if Z_{ij} equals one. Finally, students take achievement tests, with scores denoted Y_i.

3.2 Preferences

Students make application and attendance decisions to maximize expected utility net of application costs. Preferences for schools may depend on expected academic achievement. Let Y_{ij} denote the potential test score for student i if he or she enrolls in school j. These potential outcomes are given by

$$Y_{ij} = y_j (X_i, \epsilon_i), \tag{1}$$

where X_i is a vector of observed covariates and ϵ_i is unobserved academic ability. The utility associated with attending school j is

$$V_{ij} = U (Y_{ij}, X_i, D_{ij}, \omega_{ij}). \tag{2}$$

[10] The value-added calculation is jackknifed to remove the influence of a student's own score.

Here D_{ij} is distance to school j, and ω_{ij} includes unobserved attributes of school j as well as any unobserved characteristics of student i that determine valuations of school characteristics.

It will be convenient to normalize the utility of traditional public school attendance to zero and work with differences in utility between charter and public schools. Substituting (1) into (2) and differencing yields

$$V_{ij} - V_{i0} = U(y_j(X_i, \epsilon_i), X_i, D_{ij}, \omega_{ij}) - U\left(y_0(X_i, \epsilon_i), X_i, D_{i0}, \omega_{i0}\right)$$

$$\equiv u_j(X_i, D_{ij}, D_{i0}, \Psi_{ij}). \tag{3}$$

The variable Ψ_{ij} captures the influences of both academic ability ϵ_i and other unobserved factors ω_{ij} on preferences. The expression for utility in (3) does not explicitly include academic achievement or other school characteristics, but preferences for these attributes are embedded in the dependence of $u_j(\cdot)$ on X_i and Ψ_{ij}.

Throughout the analysis I maintain the following additive separability restriction:

$$u_j(X_i, D_{ij}, D_{i0}, \Psi_{ij}) = v_j(X_i, D_{ij}, D_{i0}) + \Psi_{ij}, \tag{4}$$

with Ψ_{ij} independent of X_i and $D_i = (D_{i0}...D_{iJ})$. Separable preferences of this sort are standard in analyses of dynamic discrete choice problems and treatment effects models (see, e.g., Cameron and Heckman, 1998, Heckman et al., 2016, and Vytlacil, 2002). Though Ψ_{ij} is presumed to be independent of covariates and distance in the population, (4) allows unobserved tastes to be correlated with observables conditional on charter application and enrollment choices.

Students are uncertain about their future preferences when making application decisions. The unobserved component of utility is decomposed as

$$\Psi_{ij} = \psi_{ij} + \xi_{ij}. \tag{5}$$

Here ψ_{ij} is a preference that is known at stage one and ξ_{ij} is a shock to preferences learned between stages two and three. The post-lottery shock ξ_{ij} explains why a student might apply to a charter school, receive an offer, and decline to attend.

Charter applicants also face application costs. Though submitting an application is nominally free, there is an opportunity cost of time spent filling out application forms and attending lotteries. Application costs may also capture frictions associated with learning about charter schools or school recruitment efforts.[11] Let $a = (a_1...a_J) \in \{0,1\}^J$ denote a possible charter application portfolio. The utility cost of submitting this portfolio for student i is $c(a, X_i, \eta_i)$, where η_i represents unobserved cost heterogeneity. These costs are known at the time of the application decision and are assumed to be independent of Ψ_{ij}. Students who choose not to apply to charter schools incur no costs, so $c(0, X_i, \eta_i) = 0$. A student who submits the application portfolio A_i and attends school j receives final net utility equal to $(V_{ij} - V_{i0}) - c(A_i, X_i, \eta_i)$.

[11] For example, Bergman and McFarlin (2016) show that some charter schools discourage applications from special education students.

3.3 Student Choices

3.3.1 Attendance choice

I derive students' optimal application and attendance rules by backward induction starting with stage three. At this point application costs are sunk, students know their charter offers, and there is no uncertainty about preferences. Student i can attend a traditional public school or any charter school that offers a seat. This student's set of school options is therefore

$$\mathcal{O}(Z_i) = \{0\} \cup \{j : Z_{ij} = 1\}.$$

Student i's optimal school choice at stage three is

$$S_i = \arg \max_{j \in \mathcal{O}(Z_i)} V_{ij} - V_{i0}. \tag{6}$$

The expected utility associated with this decision (before the realization of ξ_{ij}) is given by

$$w\left(Z_i|X_i, D_i, \psi_i\right) = E\left[\max_{j \in \mathcal{O}(Z_i)} V_{ij} - V_{i0}|X_i, D_i, \psi_i\right],$$

where $\psi_i = (\psi_{i1}...\psi_{iJ})$. Switching any element of Z_i from zero to one increases $w(Z_i|X_i, D_i, \psi_i)$, because an extra offer provides an option value at the school enrollment stage.

3.3.2 School lotteries

Schools hold independent lotteries in the second stage of the model. School j admits applicants with probability π_j. The probability mass function for offers Z_i conditional on the application portfolio A_i is

$$f(Z_i|A_i) = \prod_{j=1}^{J} [A_{ij}(\pi_j Z_{ij} + (1 - \pi_j)(1 - Z_{ij})) + (1 - A_{ij})(1 - Z_{ij})]. \tag{7}$$

I assume that students correctly forecast offer probabilities and therefore know this probability mass function when making application decisions.[12]

3.3.3 Application choice

Students choose charter application portfolios to maximize expected utility given the available information. At stage one student i knows X_i, D_i, ψ_i and η_i. The student does not know ξ_i, and her choice of A_i induces a lottery over Z_i at a cost of $c(A_i, X_i, \eta_i)$. The optimal portfolio choice is then

$$A_i = \arg \max_{a \in \{0,1\}^J} \sum_{z \in \{0,1\}^J} \left[f\left(z|a\right) w(z|X_i, D_i, \psi_i)\right] - c\left(a, X_i, \eta_i\right). \tag{8}$$

[12] In the empirical work the offer probabilities are allowed to vary by application cohort. The correlation in school admission rates from one year to the next is 0.61. Younger siblings of charter students are guaranteed admission, so π_j is set equal to one when a student has an older sibling at school j. Students are assumed to be siblings when they share an address.

Existing studies estimate charter school effects by comparing lottery winners and losers within charter application portfolios (Abdulkadiroğlu et al., 2011). Equation (8) provides a model-based description of how students choose to enter these quasi-experimental samples.

3.4 Academic Achievement

Since students choose schools optimally, the students enrolled in a particular school are not a random sample of the population. As in the Heckman (1979) sample selection framework, I model selection by allowing mean potential outcomes to depend on the unobserved preferences that determine school choices. Specifically, I assume:

$$E\left[Y_{ij}|X_i, D_i, Z_i, \psi_i, \xi_i, \eta_i\right] = \mu_j(X_i) + g_j\left(\psi_i\right), \tag{9}$$

where $\mu_j(X_i) \equiv E\left[Y_{ij}|X_i\right]$ is the conditional mean of Y_{ij} in the unselected population and $g_j(\cdot)$ is a function that satisfies $E\left[g_j(\psi_i)|X_i\right] = 0$.

Equation (9) combines four restrictions. First, the lottery offer vector Z_i is excluded from the potential achievement equations. This requires that lottery offers have no direct effects on test scores, a standard assumption in the school choice literature. Second, D_i is excluded from these equations, implying that distance is a valid instrument for charter school enrollment. Section 4.2 discusses this restriction. Third, application costs and post-lottery preference shocks are unrelated to potential outcomes. This implies that selection on unobservables operates through the latent preferences ψ_{ij}, which are known at the time of the application decision. The new information ξ_{ij} therefore reflects factors other than academic achievement. Finally, mean potential outcomes are assumed to be separable in observables and unobservables, a standard assumption in selection models. I next show that this specification nests a benchmark case in which students know their potential outcomes and choose schools to maximize academic achievement.

3.5 Restrictions Implied by Test Score Maximization

As noted by Willis and Rosen (1979), the theory of comparative advantage implies restrictions on the relationship between preferences and potential outcomes. It is instructive to consider a special case of the model in which students seek to maximize test scores net of distance and application costs. Suppose utility is given by

$$V_{ij} = \rho Y_{ij} + \varphi(X_i, D_{ij}) + \omega_{ij},$$

where $\varphi(\cdot)$ is a distance cost function satisfying $\varphi(x, 0) = 0 \;\; \forall x$. Assume potential outcomes are known at stage one and ω_{ij} is a random shock that occurs after the lottery.

Write potential outcomes as

$$Y_{ij} = \mu_j(X_i) + \epsilon_{ij},$$

with $E\left[\epsilon_{ij}|X_i\right] = 0$ by definition. Then the relative utility of attending charter school j is

$$V_{ij} - V_{i0} = v_j(X_i, D_{ij}, D_{i0}) + \psi_{ij} + \xi_{ij},$$

where $v_j(X_i, D_{ij}) = \rho[\mu_j(X_i) - \mu_0(X_i)] - [\varphi(X_i, D_{ij}) - \varphi(X_i, D_{i0})]$, $\psi_{ij} = \rho(\epsilon_{ij} - \epsilon_{i0})$, and $\xi_{ij} = \omega_{ij} - \omega_{i0}$.

This model implies

$$\mu_j(x) - \mu_0(x) = \frac{1}{\rho} \times v_j(x, 0, 0) \; \forall (x, j), \tag{10}$$

and

$$g_j(\psi_i) - g_0(\psi_i) = \frac{1}{\rho} \times \psi_{ij} \; \forall j. \tag{11}$$

Equation (10) states that differences in mean potential outcomes between a charter school and traditional public school should be proportional to the mean utility for the charter school after netting out the effects of distance. In other words, groups with larger average causal effects from attending a charter school should have stronger preferences for charter attendance. Equation (11) states that the test score gain generated by attending a charter school should be increasing in the unobserved taste for this school. These restrictions may be violated if parents cannot forecast potential outcomes or preferences for schools depend on factors other than academic achievement. In Section 7 I test whether charter application and attendance choices are consistent with equations (10) and (11).

4 Identification

4.1 Semi-parametric Identification

To analyze identification of the model, I consider a special case with one charter school and no unobserved application cost heterogeneity. In principle this analysis could be extended to the more general model with multiple schools and heterogeneous costs.[13] Appendix A establishes that the single-school model is a special case of the single-spell discrete duration model analyzed by Heckman and Navarro (2007), and applies their results to give precise conditions for semi-parametric identification of utility and potential outcome distributions. Here I offer intuition for how the combination of lottery and distance instruments is useful for identification of charter school effects.

As shown in Appendix A, the optimal application rule in a model with one charter school and no unobserved cost heterogeneity is

$$A_{i1} = 1\left\{\pi_1 h\left(v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1}\right) > c(X_i)\right\},$$

where $h(v)$ is a strictly increasing function derived in the appendix. Charter school attendance is given by

$$S_i = A_{i1} \times Z_{i1} \times 1\left\{v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \xi_{i1} > 0\right\}.$$

These two equations imply that preferences for students who apply and accept offers must satisfy

[13] Appendix A shows that when costs are heterogeneous the application choice model is non-separable in observables and unobservables even when the cost function itself is separable in X_i and η_i. The approach in Matzkin (2003) could be applied to analyze identification of this non-separable model.

$$v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \max\left\{h^{-1}\left(c(X_i)/\pi_1\right), \xi_{i1}\right\} > 0. \tag{12}$$

Students with these preferences apply to the lottery, then enroll in the charter school if and only if they receive a random offer. Such students are therefore "compliers" in a lottery-based instrumental variables model estimated on the sample of charter applicants (Angrist et al., 1996).

Define the population Wald (1940) instrumental variables estimand conditional on covariates and distance:

$$IV(x,d) \equiv \frac{E\left[Y_i|A_{i1} = Z_{i1} = 1, X_i = x, D_i = d\right] - E\left[Y_i|A_{i1} = 1, Z_{i1} = 0, X_i = x, D_i = d\right]}{E\left[S_i|A_{i1} = Z_{i1} = 1, X_i = x, D_i = d\right] - E\left[S_i|A_{i1} = 1, Z_{i1} = 0, X_i = x, D_i = d\right]}.$$

The arguments in Imbens and Angrist (1994) imply that $IV(x,d)$ identifies LATE, the average causal effect of charter attendance for compliers.[14] As a result, we have

$$\begin{aligned} IV(x,d) &= \mu_1(x) - \mu_0(x) \\ &+ E\left[g_1(\psi_{i1}) - g_0(\psi_{i1})|v_1(x,d_1,d_0) + \psi_{i1} + \max\left\{h^{-1}\left(c(x)/\pi_1\right), \xi_{i1}\right\} > 0\right]. \end{aligned} \tag{13}$$

The second term in (13) captures the effect of selection into application and offer takeup decisions on the instrumental variables estimand.

Semi-parametric identification of average treatment effects (ATE) in generalized Roy models is often secured with "identification at infinity" assumptions requiring instruments with large support (Heckman, 1990). A similar condition establishes identification here. Suppose there is a value of distance, d^*, that induces all students with $X_i = x$ to apply to the charter school and enroll when offered:

$$\lim_{d \to d^*} Pr\left[A_{i1} = 1|X_i = x, D_i = d\right] = \lim_{d \to d^*} Pr\left[S_i = 1|A_{i1} = Z_{i1} = 1, X_i = x, D_i = d\right] = 1.$$

For example, this may be satisfied for $d^* = (\bar{d}_0, 0)$ for $\bar{d}_0 > 0$, meaning that everyone in the immediate vicinity of a charter school applies and accepts the lottery offer when the closest public school is sufficiently far away.

If preferences have full support on the real line, condition (12) implies that $v_1(x, d_1, d_0)$ must approach infinity as d approaches d^* in order to drive both of these probabilities to one. Then

$$\begin{aligned} \lim_{d \to d^*} IV(x,d) &= \mu_1(x) - \mu_0(x) + E\left[g_1(\psi_{i1}) - g_0(\psi_{i1})|\psi_{i1} + \max\left\{h^{-1}\left(c(x)/\pi_1\right), \xi_{i1}\right\} > -\infty\right] \\ &= \mu_1(x) - \mu_0(x) + E\left[g_1(\psi_{i1}) - g_0(\psi_{i1})\right] \\ &= \mu_1(x) - \mu_0(x), \end{aligned}$$

where the last equality follows from the fact that $E\left[g_j(\psi_{i1})\right] = 0$.

This result shows that average treatment effects are semi-parametrically identified when distance is a sufficiently powerful predictor of charter preferences, because the lottery LATE equals the population ATE at

[14] Since students cannot enroll in a charter school without receiving an offer, the LATE in this simplified two-school model is also the effect of treatment on the treated, TOT (Bloom, 1984). In a model with multiple charter schools the estimand for the most commonly used IV estimator does not equal the TOT for the charter sector because IV captures a different weighted average across schools. This issue is discussed in Section 7.3.

distances that induce everyone to apply and accept offers. A similar calculation establishes identification of marginal mean potential outcomes: $\lim_{d \to d^*} E\left[Y_i | A_{i1} = 1, Z_{i1} = S_i = j, X_i = x, D_i = d\right] = \mu_j(x)$. Note that since seats are offered at random among applicants, identification of $\mu_0(x)$ does not require an additional value of distance that pushes the charter application probability to zero. The unselected mean public school outcome is revealed among lottery losers with D_i close to d^*. The lottery instrument therefore facilitates identification of average treatment effects under a weaker support condition than would be required if only distance were available.

An implication of this argument is that lottery compliers become increasingly selected as we move farther away from a charter school and the application and enrollment probabilities fall. This suggests that variation in lottery LATEs as a function of distance can be used to infer the relationship between preferences and test score gains. The mean charter preference for lottery compliers with $(X_i, D_i) = (x, d)$ is

$$\bar{\psi}_1(x,d) \equiv E\left[\psi_{i1} | v_1(x, d_1, d_0) + \psi_{i1} + \max\left\{h^{-1}\left(c(x)/\pi_1\right), \xi_{i1}\right\} > 0\right].$$

Consider two values of distance, d and d', such that $v_1(x, d_1, d_0) \neq v_1(x, d_1', d_0')$ and therefore $\bar{\psi}_1(x,d) \neq \bar{\psi}_1(x,d')$. In the model of test score maximization described in Section 3.5, we have

$$\frac{IV(x,d) - IV(x,d')}{\bar{\psi}_1(x,d) - \bar{\psi}_1(x,d')} = \frac{1}{\rho}.$$

With more than one value of X_i or two values of D_i, ρ is overidentified, permitting a test of the achievement maximization model or estimation of a more flexible relationship between preferences and achievement gains.

4.2 The Distance Instrument

The use of distance as an instrument for charter enrollment parallels the use of geographic instruments in previous research on college and school choice (see, e.g., Card, 1995, Neal, 1997, and Booker et al., 2011). Assumption (9) requires distance to have no direct effect on student performance, and also requires distance to be unrelated to potential outcomes conditional on X_i. A sufficient condition for the latter restriction is that charter school leaders choose between neighborhoods based on the distributions of these covariates. This seems plausible since X_i includes a rich set of student characteristics, including race, poverty, previous test scores, and the value-added of nearby public schools.

Table 2 explores the validity of the distance instrument by examining the relationship between distance to charter middle schools and test scores in elementary school. Columns (1) and (3) report coefficients from ordinary least squares (OLS) regressions of fourth grade math and reading scores on distance to the closest charter middle school, measured in miles. Columns (2) and (4) repeat this analysis using differential distance, defined as the difference between distance to the closest charter school and distance to the closest district school.[15] The estimates in the first row show that students who live farther from charter middle schools have significantly higher fourth grade test scores, suggesting that charter schools tend to systematically locate in

[15] Geweke et al. (2003) and Chandra and Staiger (2007) use similar differential distance instruments to study the causal impacts of hospitals on patient outcomes.

lower-achieving areas of Boston. Previous test scores are less strongly correlated with differential distance than with the level of distance to charter schools. This may reflect a tendency for charter schools to locate in denser areas where there are more schools of both types. The relationship between reading scores and differential distance is still statistically significant, however.

The second row of Table 2 shows that controlling for observed characteristics shrinks these imbalances considerably. Specifically, adding controls for sex, race, subsidized lunch, special education, limited English proficiency, and value-added of the closest district middle school renders the relationship between fourth grade math scores and distance insignificant. The coefficient for reading falls from 0.048 to 0.015, though it remains marginally statistically significant ($p = 0.06$). Corresponding estimates for differential distance are close to zero and statistically insignificant in both subjects. These results lend plausibility to the use of differential distance as an instrument in models that control for observed characteristics. The models estimated below parameterize preferences in terms of differential distance and control for these characteristics as well as fourth grade test scores.

4.3 Comparison of Lottery and Distance IV Estimates

To directly compare the two instruments used to estimate the selection model, Table 3 reports IV estimates using lottery offers and differential distance as instruments for charter attendance in equations for eighth grade test scores. The lottery estimates come from two-stage least squares (2SLS) models using a lottery offer indicator as an instrument for a charter attendance indicator in the sample of lottery applicants, controlling for lottery portfolio indicators.[16] The distance models use the full sample and control for fourth grade covariates.

As can be seen in column (1), both instruments generate strong first stage shifts in charter enrollment. A lottery offer increases the probability of charter attendance by 0.64, while a one-mile increase in differential distance decreases the probability of charter attendance by 2.6 percentage points. Columns (2) and (3) show that the two instruments produce roughly similar estimates of the effects of charter attendance, though the distance estimates are less precise. The distance instrument generates estimates of 0.45σ and 0.38σ in math and reading compared to lottery estimates of 0.55σ and 0.49σ.

The argument in Section 4.1 suggests that the interaction of lotteries and distance can be used to describe the nature of selection on unobservables. Figure 2 presents an empirical sketch of this idea by splitting the charter applicant sample into terciles of differential distance. Lottery estimates for these three groups show smaller test score gains for students who apply from farther away. The hypothesis that effects are equal across terciles is rejected at marginal significance levels in math ($p = 0.08$) though not in reading ($p = 0.33$). This pattern suggests that students who are willing to travel farther to attend charter schools experience smaller achievement gains, a finding that seems at odds with the model of test score maximization discussed in Section

[16] 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 eighth grade outcomes). Followup rates are slightly higher for higher-achieving students in the full sample, but there is no difference in followup rates for lottery winners and losers.

3.5. The analysis to follow uncovers a similar pattern using the full model.

5 Estimation

5.1 Functional Forms

To make estimation tractable I parameterize the preferences and potential outcomes introduced in Section 3. The mean utility of attending charter school j relative to public school is written

$$v_j(X_i, D_{ij}, D_{i0}) = \alpha_j + X_i'\beta - (D_{ij} - D_{i0}) \times (\varphi_0 + X_i'\varphi_x) - \left(D_{ij}^2 - D_{i0}^2\right) \times \varphi_d.$$

The parameter α_j allows for heterogeneity in average popularity across charter schools, while β measures variation in preferences for charter schools as a function of observed characteristics. This specification allows the effect of differential distance to depend on observables and includes a quadratic term to accommodate nonlinear responses to distance.

Unobserved preferences for charter schools are decomposed into a common component and a school-specific component:

$$\psi_{ij} = \theta_i + \tau_{ij}.$$

The variable θ_i, which appears in the utilities for all charter schools relative to public school, is the key unobservable driving selection into the charter sector. This preference captures any unobserved factors that influence students to opt out of traditional public school in favor of charters, such as the perceived average achievement gain from attending charter schools, attributes of the available traditional public school option, or parental motivation. The presence of θ_i implies that charter schools are closer substitutes for one another than for district schools. I allow flexible preferences for charter schools by assuming θ_i is drawn from a mixture of normal distributions. With an appropriate number of mass points, mixture models of this form can accurately approximate arbitrary distributions of unobserved heterogeneity (see Heckman and Singer, 1984 and Cameron and Heckman, 1998). I estimate the mean, variance, and probability associated with each component of the mixture, subject to the constraints that the probabilities sum to one and the overall mean satisfies $E[\theta_i] = 0$.

The τ_{ij} capture idiosyncratic tastes for particular charter schools. These tastes follow independent normal distributions with mean zero and variance σ_τ^2 conditional on θ_i. The post-lottery preference shocks ξ_{ij} follow independent standard logistic distributions. The latter assumption provides the scale normalization for the model.

Application costs are parameterized as

$$c(a, X_i, \eta_i) = 1\{|a| > 0\} \times \exp(\delta_0^f + X_i'\delta_x^f) + |a| \times \exp(\delta_0^m + X_i'\delta_x^m) + \sum_{j=1}^{J} a_j\eta_{ij},$$

where $|a| = \sum_j a_j$ is the number of charter school applications in portfolio a. The first two terms are fixed and marginal application costs, which vary with observed characteristics X_i. The unobserved cost η_{ij} is incurred for all portfolios that include school j. As in Howell (2010), this structure generates correlation between costs for portfolios with schools in common. The η_{ij} follow normal distributions with mean zero and variance σ_η^2, independent of all other variables in the model.

Finally, the outcome equations are

$$E\left[Y_{ij}|X_i, D_i, \theta_i, \tau_i\right] = \mu_j + X_i'\gamma_x^c + \gamma_\theta^c\theta_i + \gamma_\tau\tau_{ij},\ j = 1...J,$$

$$E\left[Y_{i0}|X_i, D_i, \theta_i, \tau_i\right] = \mu_0 + X_i'\gamma_x^0 + \gamma_\theta^0\theta_i.$$

This specification includes school-specific intercepts, as well as covariate effects that differ between charter and traditional public schools. The parameters γ_θ^0, γ_θ^c, and γ_τ describe selection on unobservables. γ_θ^0 measures selection on absolute advantage: if students with higher potential public school outcomes are more likely to select into charter schools, then $\gamma_\theta^0 > 0$. The difference $\left(\gamma_\theta^c - \gamma_\theta^0\right)$ measures selection on comparative advantage into the charter sector as a whole, and γ_τ measures selection on comparative advantage into individual charter schools. The model of test score maximization described in Section 3.5 implies $(\gamma_\theta^c - \gamma_\theta^0) = \gamma_\tau > 0$, but I do not impose this restriction in the estimation.

5.2 Estimation Procedure

I estimate the preference parameters by maximum simulated likelihood (MSL). Given the logistic assumption for ξ_{ij}, the school enrollment choice in (6) is a standard multinomial logit problem. The probability of choosing charter school j at this stage is

$$Pr\left[S_i = j|Z_i, X_i, D_i, \theta_i, \tau_i\right] = \frac{Z_{ij} \times \exp\left(v_j(X_i, D_{ij}, D_{i0}) + \theta_i + \tau_{ij}\right)}{1 + \sum_{j'=1}^{J} Z_{ij'} \times \exp\left(v_{j'}(X_i, D_{ij}, D_{i0}) + \theta_i + \tau_{ij'}\right)}$$

$$\equiv p(j|Z_i, X_i, D_i, \theta_i, \tau_i). \tag{14}$$

The probability of public school enrollment is one minus the sum of charter enrollment probabilities. The logit model implies that the expected utility resulting from the enrollment stage is

$$w(Z_i|X_i, D_i, \theta_i, \tau_i) = \log\left(1 + \sum_{j=1}^{J} Z_{ij} \times \exp\left(v_j(X_i, D_{ij}, D_{i0}) + \theta_i + \tau_{ij}\right)\right).$$

The portfolio decision in (8) does not yield closed forms for application choice probabilities. I approximate these probabilities with a logit kernel smoother (Train, 2003). For small λ, we have

$$Pr\left[A_i = a|X_i, D_i, \theta_i, \tau_i, \eta_i\right] \approx \frac{\exp\left(\left[\sum_z \left[f\left(z|a\right) w(z|X_i, D_i, \theta_i, \tau_i)\right] - c\left(a, X_i, \eta_i\right)\right]/\lambda\right)}{\sum_{a'} \exp\left(\left[\sum_z \left[f\left(z|a'\right) w(z|X_i, D_i, \theta_i, \tau_i)\right] - c\left(a', X_i, \eta_i\right)\right]/\lambda\right)}$$

$$\equiv q(a|X_i, D_i, \theta_i, \tau_i, \eta_i). \tag{15}$$

This expression can be interpreted as a multinomial logit choice probability from a model that adds an extreme value error with small variance to the expected utility associated with each application portfolio. I set λ equal to 0.05 in the estimation.

Combining (14) and (15), the likelihood of student i's application choice, lottery offers, and enrollment decision is

$$\mathcal{L}\left(A_i, Z_i, S_i|X_i, D_i\right) = \int q\left(A_i|X_i, D_i, \theta, \tau, \eta\right) \times f(Z_i|A_i) \times p(S_i|Z_i, X_i, D_i, \theta, \tau) dF\left(\theta, \tau, \eta|X_i, D_i\right).$$

I simulate this integral using 300 draws of (θ, τ, η) for each observation. For specifications with multiple mass points (types) for θ_i, the likelihood is a weighted average of type-specific likelihoods. The MSL estimator maximizes the sum of log simulated likelihoods for students in the sample.

Following Heckman (1979), I estimate the parameters of the outcome equations using a two-step control function approach. Define

$$\theta^*\left(A_i, Z_i, S_i, X_i, D_i\right) = E\left[\theta_i|A_i, Z_i, S_i, X_i, D_i\right],$$

$$\tau_j^*\left(A_i, Z_i, S_i, X_i, D_i\right) = E\left[\tau_{ij}|A_i, Z_i, S_i, X_i, D_i\right].$$

These functions are posterior means for unobserved preferences given observed choices, lottery offers, covariates and distances. I use the first-step choice model estimates to construct estimated posterior means by simulation, labeled $\hat{\theta}^*$ and $\hat{\tau}_j^*$. These estimates are then included in a second-step OLS regression:

$$\begin{aligned} Y_i = \mu_0 + X_i'\gamma_x^0 + \gamma_\theta^0\hat{\theta}^*\left(A_i, Z_i, S_i, X_i, D_i\right) + \sum_{j=1}^{J}(\mu_j - \mu_0)1\left\{S_i = j\right\} \\ +1\left\{S_i > 0\right\} \times \left[X_i'(\gamma_x^c - \gamma_x^0) + (\gamma_\theta^c - \gamma_\theta^0)\hat{\theta}^*\left(A_i, Z_i, S_i, X_i, D_i\right) + \gamma_\tau\hat{\tau}_{S_i}^*\left(A_i, Z_i, S_i, X_i, D_i\right)\right] + e_i. \end{aligned} \quad (16)$$

The posterior mean unobservables serve as control functions that correct for selection into school enrollment, allowing consistent estimation of the unselected potential outcome equations. I use methods described by Murphy and Topel (1985) to adjust inference for sampling error introduced by first-step estimation of the control functions.

6 Parameter Estimates

6.1 Preference Estimates

I report results from three increasingly flexible versions of the choice model. The first imposes a common value for the charter utility intercepts α_j, sets the variance of the idiosyncratic preferences τ_{ij} to zero, and models the charter taste θ_i with a single normal distribution. This is a model in which students view charter schools as homogeneous and choose between them only on the basis of distance and application costs. Analysis of this model is useful because it can be straightforwardly compared to the benchmark two-sector selection model commonly used in the literature (Heckman and Vytlacil, 2005; Heckman et al., 2006). The second model adds heterogeneity in α_j and τ_{ij}, thereby allowing students to have stronger preferences for specific charter schools.

The third model extends the second by replacing the single normal distribution for θ_i with a two-mass mixture of normals.

Table 4 displays the number of parameters and maximized log-likelihood for each preference model. Allowing charter school heterogeneity and a flexible distribution for θ_i dramatically improves the fit of the model. The homogeneous charter model includes 43 parameters and generates a log-likelihood value of -12,917. Allowing charter heterogeneity adds seven parameters and increases the log-likelihood by 854. A likelihood ratio test therefore rejects the model with homogeneous charter schools ($p < 0.01$). Likewise, the single normal model is decisively rejected in a test against the two-mass mixture model ($p < 0.01$). I therefore focus on estimates from the mixture model with heterogeneous charter schools, and report results for the other two models when these comparisons are useful. Appendix B provides further goodness of fit diagnostics for the mixture model.

Table 5 displays utility and application cost estimates from the mixture model. Column (1) shows estimates of the utility parameters α_j and β, while column (2) reports estimates of the distance cost parameters φ_0, φ_x and φ_d. The covariate vector X_i is de-meaned in the estimation sample so that main effects are effects at the mean. The intercept reported in column (1) is the average of school-specific utility intercepts; estimates of school-specific parameters appear in Appendix Table A5. The utility intercept is negative and statistically significant, implying that on average, students prefer to enroll in traditional public schools rather than charter schools even in the absence of distance and application costs. The estimated main effect in the distance cost function equals 0.18 with a standard error of 0.02, which indicates that distance has a significant negative effect on charter demand. The coefficient on distance squared is close to zero and insignificant, suggesting that the disutility of distance is roughly linear.

The coefficients for observables in column (1) are generally consistent with the demographic patterns reported in Table 1. Subsidized lunch status and special education are associated with weaker demand for charter schools, while higher fourth grade test scores are associated with stronger demand. I find no difference in charter preferences between males and females. Average preferences for charter schools are weaker for non-white students, but these students are also less sensitive to distance. Distance interaction effects for other groups are small.

Columns (3) and (4) of Table 5 display estimates of the natural logarithms of fixed and marginal application costs. Fixed and marginal costs equal $\exp(-2.1) = 0.12$ and $\exp(-0.66) = 0.52$ for an average student. These magnitudes are equivalent to increases of 0.7 and 2.9 miles in distance to school, respectively. The large marginal cost estimate reflects the fact that most charter applicants submit only one application, which implies that additional applications must be costly even after the fixed cost has been paid. Marginal costs are significantly smaller for non-white students and larger for students with limited English proficiency status and those eligible for subsidized lunch. Estimated fixed costs are larger for black and hispanic students, but these interaction estimates are imprecise.

Table 6 reports estimated distributions of of unobserved preferences. Estimates for the two-mass mixture model appear in columns (3) and (4), while columns (1) and (2) show estimates from the homogeneous charter and single normal models for comparison. The results here show important heterogeneity in unobserved tastes. Estimates of the homogeneous charter model indicate that a one standard deviation increase in θ_i is roughly equivalent to a 6.7 mile increase in distance. The corresponding estimate for the application cost η_{ij} is 0.6 miles. Adding charter heterogeneity in column (2) reduces the role for application costs and reveals substantial variation in idiosyncratic tastes. The estimated standard deviation of τ_{ij} in this model is equivalent to 3.3 miles of distance.

The two-mass mixture estimates in columns (3) and (4) show that heterogeneity in preferences for charter schools is well-described by two unobserved types of students. The first type, which includes 42 percent of the population, has a high average taste for charter schools: the mean charter utility for this group is $1.64 - 1.10 = 0.54$, implying that these students prefer charter schools to traditional public schools. As shown in column (4), the average charter taste is very negative for the second type. Estimated within-type variances of θ_i are small, suggesting that two discrete types are sufficient to characterize much of the variation in preferences for the charter sector as a whole. The standard deviation of τ_{ij} remains large, implying significant preference variation within the charter sector as well.

6.2 Achievement Estimates

Estimates of equation (16) for eighth grade math and reading scores appear in Table 7. The control functions are posterior mean unobservables from the two-mass mixture model. Results based on the other two preference models and for other grades are similar; these estimates appear in Appendix Tables A6 and A7.[17] The main effects in columns (2) and (4) of Table 7 imply that charter attendance raises eighth grade math and reading scores by 0.71σ and 0.52σ on average. Non-white students, poor students, and students with lower past achievement lag behind other students in traditional public schools, and receive larger benefits from charter school attendance. In this sense, charter schools tend to reduce 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 (Abdulkadiroğlu et al., 2011; Angrist et al., 2012). Charter effects are similar for boys and girls.

Estimates of the selection parameters reveal that stronger unobserved preferences for charters are associated with *smaller* achievement benefits from charter attendance. The control function coefficients in columns (1) and (3) show that students with stronger preferences for charters do better in traditional public schools, implying that higher-ability students tend to select into the charter sector. Similar to the pattern for observed characteristics, column (2) shows that charter attendance produces smaller math gains for these students: a one unit increase in θ_i (equivalent to roughly 0.7 standard deviations) reduces the charter math effect by 0.1σ,

[17] Test scores are an ordinal measure of performance, and patterns of test score effects may be sensitive to the scaling used for test scores (Nielsen, 2015). Appendix Table A8 reports estimates for eighth grade math scores measured in percentile units, log percentile units, and changes in percentile rank between fourth and eighth grade. The key results are similar for each of these transformations, indicating that the findings are robust to standard changes in scaling. Fewer than 0.3 percent of Boston students and 0.5 percent of charter applicants earned the maximum score in each subject, which suggests the results are not due to "ceiling effects."

and this estimate is statistically significant. The corresponding estimate for reading is also negative but not statistically significant. The estimated coefficients on the idiosyncratic taste τ_{ij} are small and insignificant in both subjects, which suggests that students do not systematically choose between charters on the basis of school-specific match effects in academic achievement.[18]

7 Absolute and Comparative Advantage in Charter School Choice

7.1 Tests of Cross-Equation Restrictions

The estimated model can be used to test cross-equation restrictions implied by the theory of comparative advantage. As shown in equations (10) and (11), test score maximization implies that achievement gains should be larger for students with stronger preferences for charter attendance. Moreover, differences in utility and test score effects should be proportional. The estimates in Tables 5 and 7 appear inconsistent with this restriction: preferences for charters are weaker for disadvantaged groups, but test score effects are larger for these groups. Table 8 reports ratios of charter preference coefficients to achievement gain coefficients for observed and unobserved student characteristics. Many of these ratios are negative, and a Wald test rejects the hypothesis that the ratios are equal and weakly positive ($p < 0.01$).[19]

Equation (10) also has implications for heterogeneity across charter schools. Specifically, average utilities should be larger for charters that generate larger test score gains. Figure 3 plots school-specific average treatment effect estimates against school-specific mean utilities. In contrast to the prediction of equation (10), this relationship is downward-sloping in both math and reading, implying that less-popular charter schools tend to produce larger test score impacts. The hypothesis that these parameters lie on a line with weakly positive slope is rejected in both math and reading ($p < 0.01$).

7.2 Selection and Charter School Effects

These test results imply that students do not sort into charter schools to maximize test scores. To further explore the pattern of selection into the charter sector, I next consider a summary measure of the relationship between achievement impacts and preferences for charter schools. Define the preference index

$$\mathcal{P}_{\text{i}} \equiv -\left(X_i'\beta + \theta_i\right).$$

$\mathcal{P}_i$ may be viewed as student i's overall utility cost from entering the charter sector as a function of observed and unobserved characteristics, ignoring distance and application costs. Let $F_{\mathcal{P}}(\cdot)$ represent the cumulative

[18]A natural alternative specification in the two-mass mixture model is to allow potential outcomes to depend on an indicator for type rather than a linear term in θ_i. Appendix Table A9 reports estimates from a model using the posterior type probability as the control function. Results from this model show a similar pattern: the unobserved type with stronger tastes for charter schools performs better in public school and gains less from charter attendance.

[19]I implement this test using methods described by Kodde and Palm (1986) for Wald tests of hypotheses combining equality and inequality constraints.

distribution function of this index, and let $U_i = F_{\mathcal{P}}(\mathcal{P}_i)$ denote student i's percentile. The relationship between preferences and potential public school outcomes is summarized by the function

$$m_0(u) = E\left[Y_{i0}|U_i = u\right].$$

The corresponding average charter outcome is

$$m_1(u) = \sum_{j=1}^{J} s_j E\left[Y_{ij}|U_i = u\right],$$

where $s_j = Pr\left[S_i = j|S_i > 0\right]$ is the enrollment share for school j among charter students. The average achievement benefit generated by charter attendance for students at cost percentile u is then

$$\Delta(u) = m_1(u) - m_0(u).$$

The function $\Delta(u)$ describes the relationship between charter preferences and the causal effects of charter attendance. This function is closely related to the Marginal Treatment Effect (MTE) concept developed by Heckman and Vytlacil (1999, 2005, 2007b). MTEs measure treatment effects at each percentile of the unobserved cost of participating in a treatment. Here $\Delta(u)$ measures variation in treatment effects as a function of both observed and unobserved components of costs. I later separately explore the roles of observables and unobservables.

Figure 4 characterizes patterns of absolute and comparative advantage in charter school choice. I focus on results for eighth grade math scores. Panel A plots the marginal mean potential outcome functions $m_1(u)$ and $m_0(u)$, while Panel B plots the charter effect $\Delta(u)$. These functions are computed via local linear regressions fit to data simulated from the two-mass mixture model. The dotted vertical line shows the mean preference for charter students, and the dashed/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 outcomes and causal effects for typical charter and non-charter students.

The $m_0(u)$ function in Figure 4 is downward sloping, which indicates that students with stronger charter preferences have an absolute advantage in the traditional public sector. The slope of $m_1(u)$ is less steep than the slope of $m_0(u)$, so the effect of charter attendance $\Delta(u)$ in panel B rises sharply as charter costs increase. An increasing $\Delta(u)$ implies that 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 $E\left[\Delta(U_i)|S_i > 0\right]$, is slightly over 0.5σ. The effect of treatment on the non-treated (TNT), defined as $E\left[\Delta(U_i)|S_i = 0\right]$, is over 0.7σ, which represents a 40 percent increase over the TOT. This implies that expanding charter schooling to new populations that are not currently served would generate larger effects than the current charter system.

Figure 5 separates this pattern into components due to observed and unobserved student characteristics. Let $F_{-X\beta}(\cdot)$ and $F_{-\theta}(\cdot)$ denote CDFs of the observed charter cost $-X_i'\beta$ and the unobserved cost $-\theta_i$, respectively. Panels A and B plot average treatment effects as functions of observed and unobserved cost percentiles, $U_i^{obs} = F_{-X\beta}(-X_i'\beta)$ and $U_i^{unobs} = F_{-\theta}(-\theta_i)$. For comparison, this figure also plots treatment effect estimates from the model with homogeneous charter schools. The unobserved component of treatment effects from this two-sector model is exactly the MTE function of Heckman and Vytlacil (2005). Results here show that the positive association between charter effects and utility costs is driven both by observed characteristics (since disadvantaged students and those with low past scores have weaker tastes for charters and larger test score gains) and unobserved characteristics (since high-θ_i students have stronger tastes for charters and smaller gains).

One possible explanation for these results is that parents who invest more in human capital on other margins may also be more motivated to enroll their children in effective 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 figures 4 and 5.

7.3 Alternative Approaches to Extrapolation

To highlight the value of the selection model estimated here, it is worth comparing causal parameters derived from the model to atheoretical predictions generated by lottery estimates of the type reported in the previous literature. Lottery-based estimates in Abdulkadiroğlu et al. (2011) show larger impacts for poorer and lower-achieving students. This section compares the results of reduced-form extrapolation based on these and other covariates to the insights gleaned from the structural selection model.

Atheoretical covariate-based approaches to extrapolation reweight experimental or quasi-experimental treatment effect estimates to match the distribution of observed characteristics in a new population (see, e.g., Hotz et al., 2005 and Angrist and Fernandez-Val, 2013). This approach can be operationalized through estimation of a set of 2SLS models for lottery applicants, with second stage

$$Y_i = \beta C_i + \sum_a \gamma_a 1\{A_i = a\} + \epsilon_i \tag{17}$$

and first stage

$$C_i = \pi Z_i^{max} + \sum_a \lambda_a 1\{A_i = a\} + \eta_i,$$

where $C_i = 1\{S_i > 0\}$ is a charter attendance indicator, and $Z_i^{max} = \max_j Z_{ij}$ is an indicator equal to one if student i receives an offer from any charter school. Let $G_i \in \{1...\bar{G}\}$ denote an exclusive and exhaustive set of covariate-based groups. Simple predictions of the TNT and TOT are $\sum_g \beta_g Pr[G_i = g|C_i = 0]$ and $\sum_g \beta_g Pr[G_i = g|C_i = 1]$, where β_g is the coefficient from estimation of (17) 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.

Table 9 compares covariate-based and model-based predictions of several treatment effect parameters for eighth grade math scores. Column (1) replicates the pooled 2SLS math estimate from Table 3, which equals

0.55σ. As shown in Appendix C, this estimate produces a particular weighted average of lottery-specific LATEs that is not generally interpretable as an effect for any subpopulation of economic interest. Column (5) shows that a model-based prediction of this parameter, which is constructed by applying the 2SLS weights to data simulated from the model, is similar to the 2SLS estimate.

Columns (2), (3) and (4) of Table 9 reveal that reweighting 2SLS estimates based on subsidized lunch status, terciles of fourth grade math score, or interactions of these covariates with race and special education status tends to raise the implied ATE and TNT relative to the LATE and 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 results based on the structural model. The predicted magnitudes generated by the covariate-based and model-based approaches are very different, however. Covariate-based estimates suggest small gaps between the TNT and TOT (between 0.03σ and 0.05σ), while the structural approach indicates a large gap (0.22σ). This is driven by the link between unobserved preferences and treatment gains uncovered by the selection 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 of effects 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.

8 Counterfactual Simulations

8.1 Description of Counterfactuals

I next explore the policy implications of self-selection into charter schools by simulating the impacts of changes to the Boston charter landscape. I report on three sets of counterfactual simulations. The first, a "baseline" charter school expansion, uses the estimates in Tables 5, 6 and 7 to predict the effects of expanding the charter sector to 20 schools.[20] The second "reduced cost" expansion modifies preferences to eliminate marginal application costs. This counterfactual approximates the effects of providing information and eliminating logistical barriers. The final "altered preference" simulation increases the utility 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 charter schools.

To focus attention on demand-side behavior I make several simplifying assumptions about the supply side of the charter school market. The supply side is defined by a set of charter schools, with each school characterized by a location, an admission probability π_j, an average utility α_j, and a mean achievement parameter μ_j. I choose locations for the first six expansion schools using the addresses of new campuses that opened after the

[20]The number of possible application portfolios grows exponentially with the number of charter schools. I manage the number of choices in the counterfactual simulations by limiting the choice set to portfolios with two or fewer schools.

application data used here were collected. Additional schools are placed at the center of randomly selected zip codes with no charter schools.

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 capacity constraints. Capacities for new schools are set equal to the mean capacity for existing schools. Appendix D describes the methods used to compute counterfactual admission probabilities.

The average test score and utility parameters for new schools are set equal to the means of α_j and μ_j from the estimated model. There are several reasons this assumption may fail to hold in practice. If peer characteristics play a role in charter demand or effectiveness, the current values of α_j and μ_j will partly reflect peer attributes that may change as the composition of the sector evolves. For example, positive peer effects may be diluted in expansions that draw in less positively selected students, dampening charter effectiveness.[21]

Along similar lines, it may be difficult for new charters to replicate the production technology used by existing campuses if teachers, principals, or other inputs are supplied inelastically (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). As a result of these issues predictions for counterfactuals that are farther out of sample should be viewed as more speculative.

8.2 Charter Expansion Effects

Figure 6 summarizes the counterfactual simulations. All simulations are based on the two-mass mixture model of charter preferences. The outcomes of interest are school choices, charter oversubscription, and charter school treatment effects. In each panel, a dotted vertical line indicates the number of charter schools used to estimate the model, and a dashed/dotted line indicates the size of Boston's subsequent 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 Average Marginal Treatment Effect (AMTE) parameter discussed by Heckman et al. (2016, forthcoming). For a student i receiving at least one charter offer, let

$$j^*(i) = \arg \max_{j \in \mathcal{O}(Z_i), j \neq 0} V_{ij}$$

[21] Existing evidence suggests that peer effects are not the main source of charter school impacts. Angrist et al. (2013) show that variation in impacts across charter lotteries is unrelated to changes in peer quality resulting from lottery admission. I replicate this finding in Appendix Figure A2: test score gains for Boston charter middle school applicants are not larger in lotteries that generate larger changes in peers' past achievement.

denote the preferred charter school among those offering seats. Define

$$\Delta^*(t) = E\left[Y_{ij^*(i)} - Y_{i0} \mid |V_{ij^*(i)} - V_{i0}| \leq t, \mathcal{O}(Z_i) \neq \{0\}\right]. \quad (18)$$

For small t, this parameter describes causal effects for students who are on the margin of deciding whether to remain in traditional public schools, and would be induced to enter the charter sector by a small increase in the attractiveness of charter schools. Figure 6 reports $\Delta^*(t)$ in each counterfactual, setting t equal to one tenth of the standard deviation of $|V_{ij^*(i)} - V_{i0}|$ in the current charter system.

Results for the baseline simulation imply that demand for charter schools in Boston may be exhausted as the system expands. Panel B shows that charter expansion is predicted to reduce oversubscription: admission probabilities rise with the number of schools, and the share of seats left empty also increases when the number of schools moves beyond 15.[22] 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, 25 percent attend one, and 10 percent of charter seats are empty. This pattern is driven by the large application costs and negative average utilities reported in Table 5. Panels C and D shows that average and marginal treatment effects increase with the size of the charter sector, a consequence of the selection pattern documented in Section 7: 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 rising marginal treatment effects and weak demand indicates that many high-benefit students choose to remain in traditional public schools even when charter seats are widely available.

Counterfactuals that alter the pattern of self-selection into charter schools increase the effectiveness of charter school expansion. The reduced cost simulation eliminates marginal application costs, increasing overall charter demand by construction. Treatment effects are 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 cost of doing so is lower, 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 marginal application 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. TOTs in this counterfactual are larger than corresponding effect for the reduced cost counterfactual. MTEs are even larger, a

[22] These simulation results roughly match the growth of charter middle school enrollment that has occurred since the data used here were collected. The 2010 charter expansion reform resulted in six new Commonwealth charter schools. The seven sample schools and six new schools enrolled 18 percent of Boston sixth graders in 2012-2013. The corresponding prediction in Figure 6 is 19 percent.

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 6 suggest that reforms aimed at changing self-selection into charter schools have the potential to boost achievement much more than reforms that merely add more charter seats.

9 Conclusion

This paper develops and estimates a generalized Roy model of charter school applications, attendance decisions, and academic achievement to analyze patterns of absolute and comparative advantage in charter school choice. The estimates reveal that tastes for charter schools among Boston students are negatively associated with 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 in academic achievement. 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.

This pattern of self-selection may reflect a greater willingness of motivated parents to both seek out effective schools and invest more in human capital on other dimensions. It may also reflect a lack of knowledge about school quality among disadvantaged families. These possibilities are 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 (Hastings et al., 2009; Brand and Xie, 2010; Hoxby and Avery, 2012; Butler et al., 2013; Dillon and Smith, 2017).

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 most 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 high-benefit 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 results raise the further question of whether parents who forgo large potential achievement gains are 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. The mechanisms through which parents form preferences over schools are an important topic for future work.

Figure 1: Sequence of events

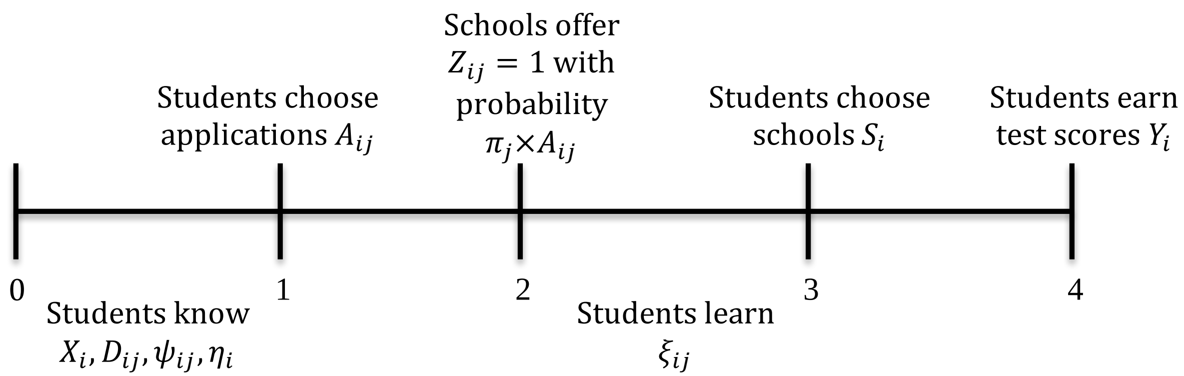
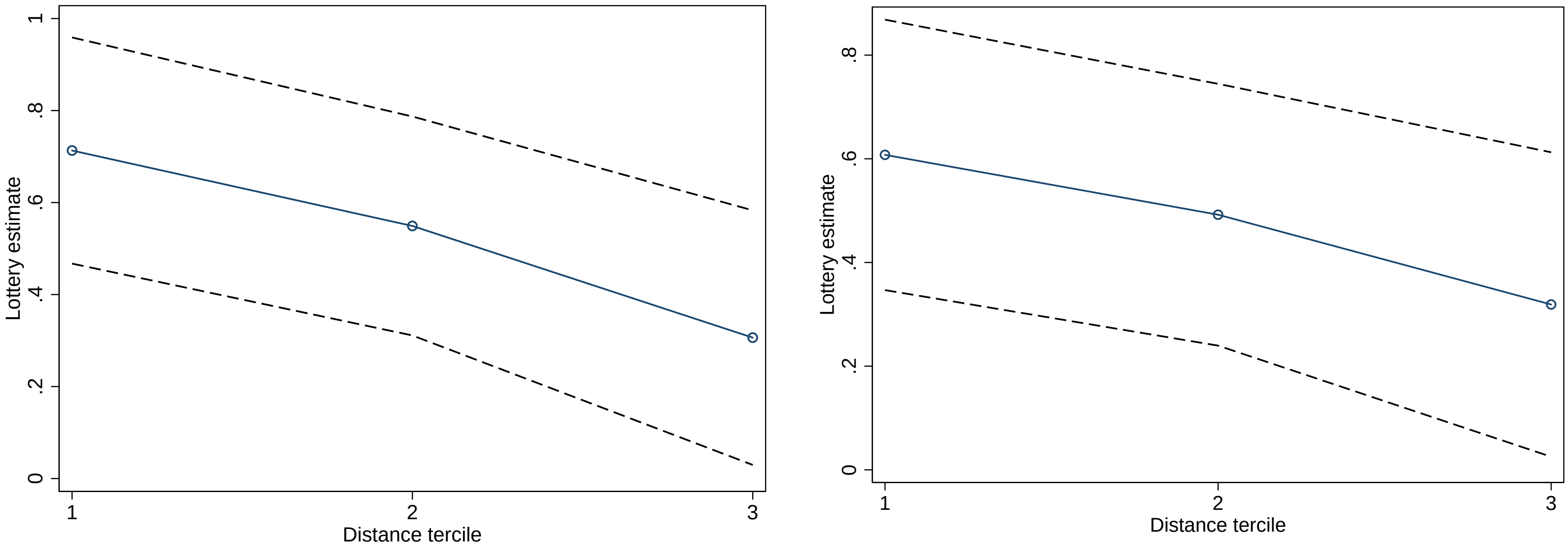


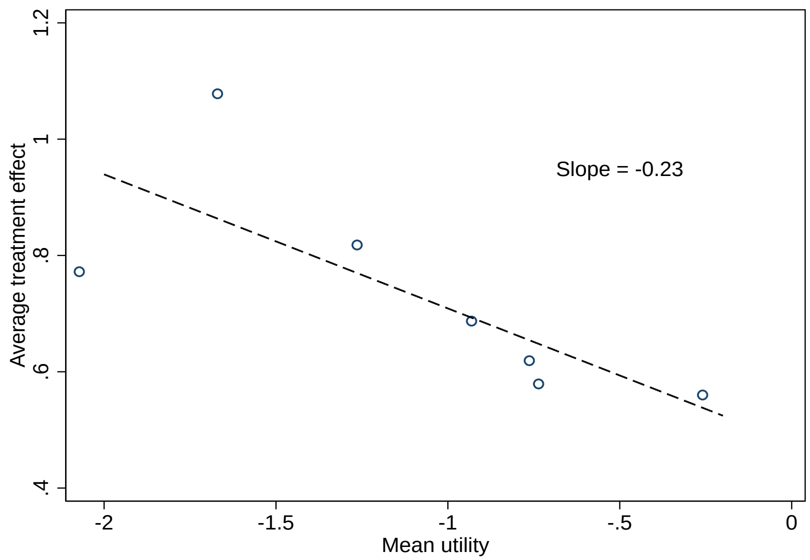
Figure 2: Relationship between distance to charter schools and lottery estimates



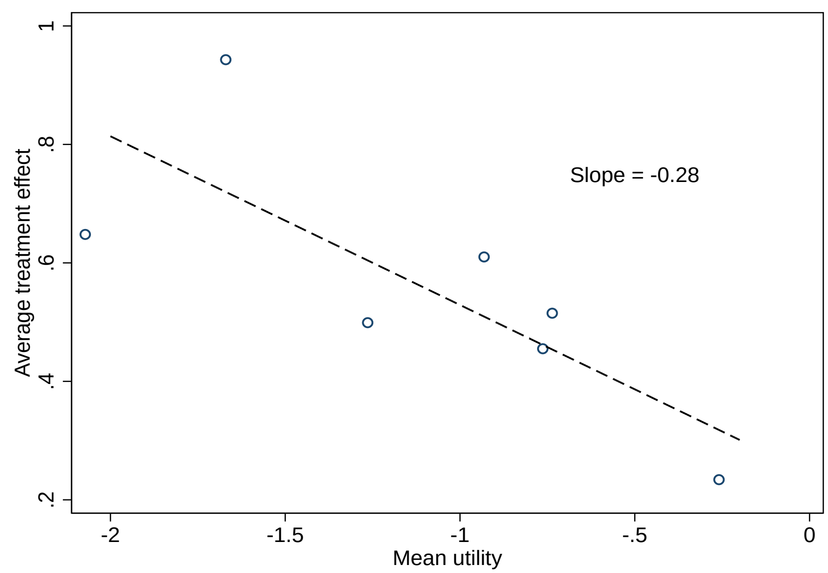
A. Math *B. Reading*

Notes: This figure displays relationships between lottery-based instrumental variables estimates of charter school effects on eighth grade test scores and distance that applicants travel to apply. Panel A shows results for math scores, and panel B displays results for reading. Estimates come from a two-stage least squares model that interacts charter school attendance with indicators for terciles of the differential distance between the closest charter school and closest traditional public school. The instruments are interactions of a lottery offer indicator with differential distance terciles, and both stages control for lottery portfolio indicators and tercile main effects. Dashed lines indicate 95 percent confidence intervals.

Figure 3: Relationship between school mean utilities and average treatment effects



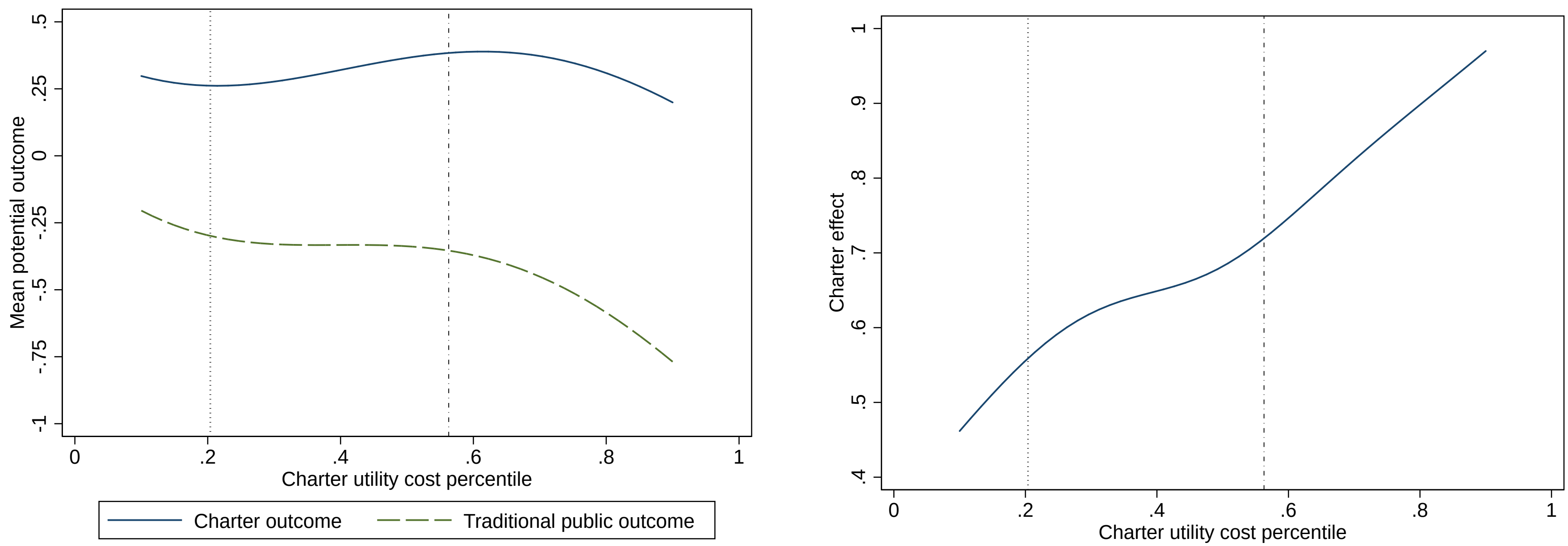
A. Math



B. Reading

Notes: This figure displays estimates of average utilities and average treatment effects for Boston charter middle schools. Estimates come from the two-mass mixture model in column (3) of Table 4. Dashed lines are least squares regression lines weighted by the inverse variance of the estimated average treatment effects.

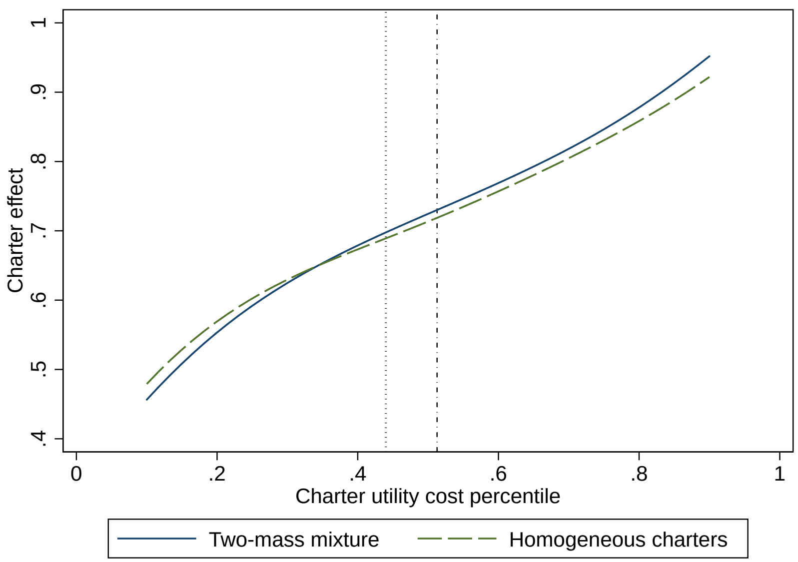
Figure 4: Absolute and comparative advantage in charter school choice



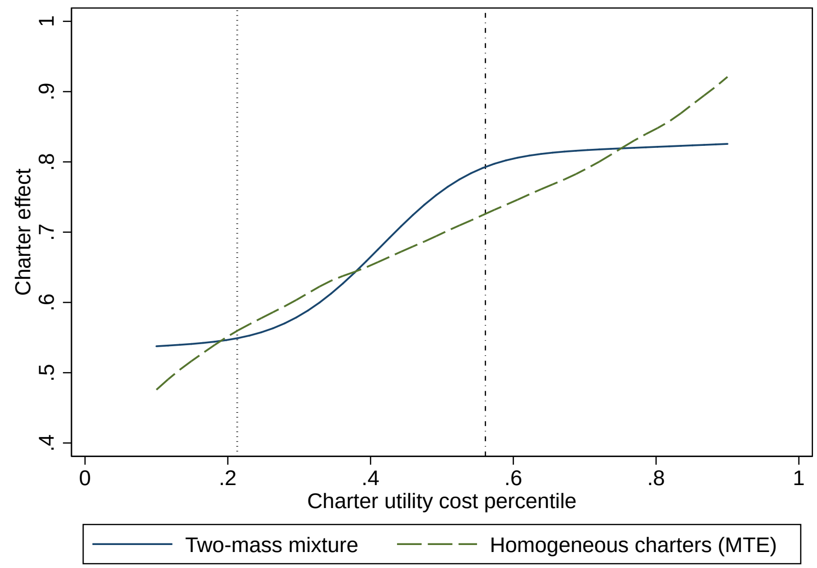
A. Selection on levels *B. Selection on gains*

Notes: This figure displays relationships between preferences for charter school attendance and outcome levels and gains in eighth grade math. Panel A plots conditional expectations of potential outcome levels in charter and traditional public schools as functions of percentiles of a charter utility cost index that combines observed and unobserved student characteristics. Panel B plots conditional expectations of charter school causal effects as functions of the utility cost index. Conditional expectations are estimated via local linear regressions in a data set of 1,000,000 individuals simulated from the two-mass mixture model. Covariates and spatial locations in these simulations are obtained by sampling with replacement from the observed data. The dotted and dashed/dotted vertical lines in each panel indicate mean preferences for students enrolled in charter and traditional public schools.

Figure 5: Selection on observables and unobservables



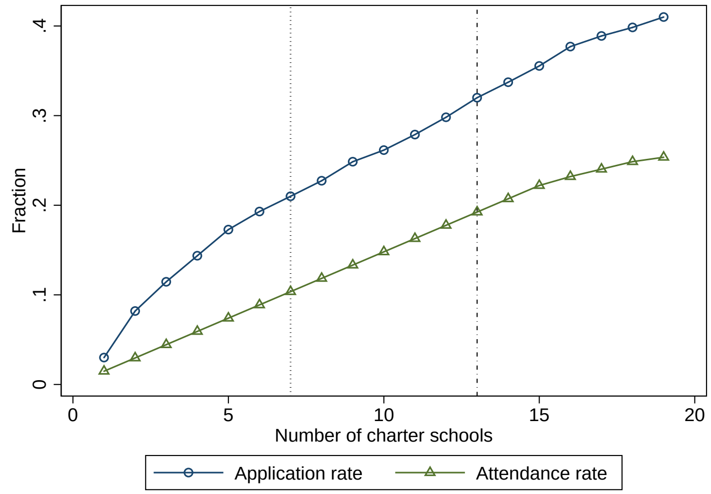
A. Selection on observables



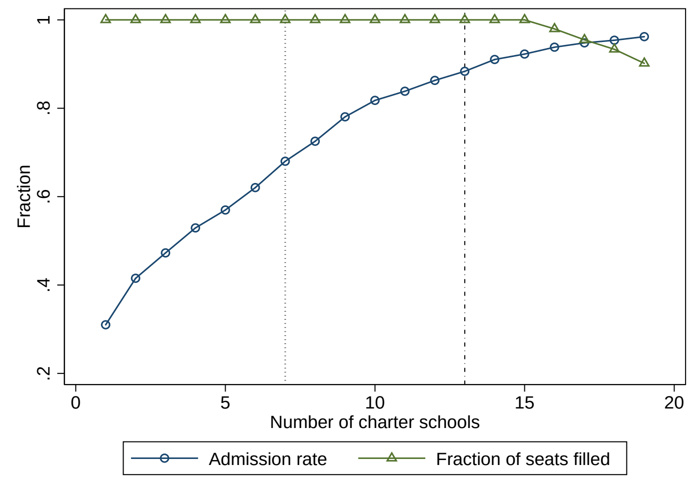
B. Selection on unobservables

Notes: This figure displays relationships between charter school preferences and causal effects on eighth grade math scores, separately for observed and unobserved components of preferences. Solid lines show estimates from the two-mass mixture model with heterogeneous charter schools, and dashed lines show estimates from a model in which schools are assumed to be homogeneous. Panel A shows relationships between the observed component of the utility cost of charter attendance expressed in percentile units and average charter school effects. Panel B shows relationships between percentiles of the unobserved cost and charter effects. Conditional expectations are estimated via local linear regressions in data sets of 1,000,000 individuals simulated from each model. Covariates and spatial locations in these simulations are obtained by sampling with replacement from the observed data. The dotted and dashed/dotted vertical lines in each panel indicate mean preferences for students enrolled in charter and traditional public schools in the heterogeneous school model.

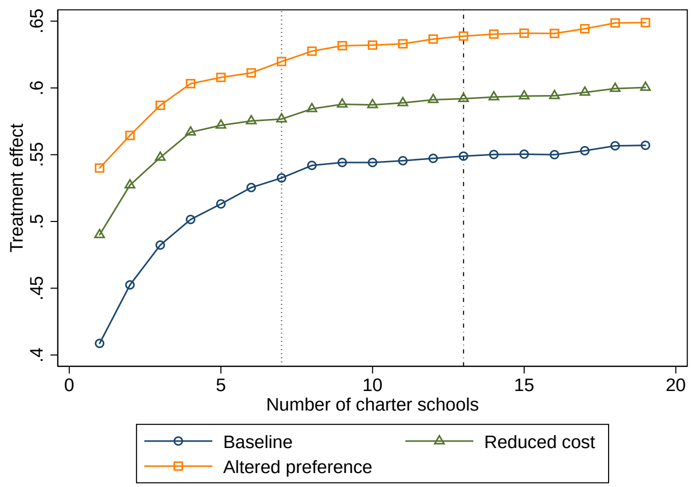
Figure 6: Counterfactual simulations



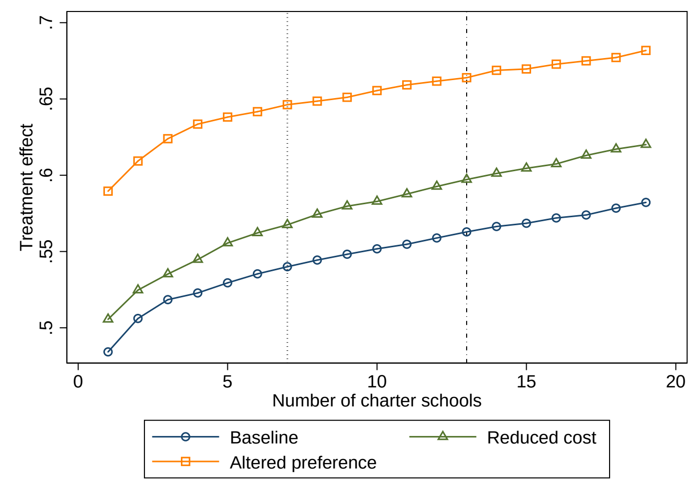
A. Applications and attendance, baseline simulation



B. Oversubscription, baseline simulation



C. Effects of treatment on the treated



D. Average marginal treatment effects (AMTE)

Notes: This figure displays simulated effects of charter school expansion. The dotted vertical line in each panel corresponds to the number of charter schools in the sample, while the dashed/dotted line corresponds to Boston's subsequent expansion. Locations for new charters are chosen at random in zip codes without a charter school. The baseline simulation is based on the two-mass mixture model of charter preferences. The reduced cost simulation sets marginal application costs 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. Average marginal treatment effects are average effects of charter attendance for students who are approximately indifferent between attending and not attending charter schools in each counterfactual. Results are based on 1,000,000 simulations of the each model.

Table 1: Descriptive statistics for Boston middle school students

	All Boston students		Charter applicants	
	Mean	Standard deviation	Mean	Standard deviation
	(1)	(2)	(3)	(4)
	A. Charter school applications and attendance			
Applied to charter school	0.175	0.380	1.000	0.000
Applied to more than one charter	0.046	0.210	0.265	0.442
Received charter offer	0.125	0.331	0.718	0.450
Attended charter school	0.112	0.316	0.600	0.490
	B. Student characteristics			
Female	0.492	0.500	0.490	0.500
Black	0.460	0.498	0.518	0.500
Hispanic	0.398	0.490	0.317	0.465
Subsidized lunch	0.821	0.383	0.723	0.448
Special education	0.226	0.418	0.170	0.376
Limited English proficiency	0.212	0.409	0.136	0.343
Fourth grade math score	-0.520	1.070	-0.314	0.990
Fourth grade reading score	-0.636	1.137	-0.413	1.036
	C. Nearby schools			
Miles to closest charter school	2.105	1.168	1.859	1.087
Miles to closest district school	0.512	0.339	0.580	0.372
Value-added of closest district school	0.032	0.154	0.022	0.167
N	9,156		1,601	

Notes: This table shows descriptive statistics for students attending fourth grade at traditional public schools in Boston between 2006 and 2009. Column (1) shows means for the full sample, while column (2) shows means for charter applicants. The sample excludes students without eighth grade test scores. Fourth grade test scores are normalized to have mean zero and standard deviation one in the population of all Massachusetts students. District school value-added is measured as the average residual from a regression of sixth grade math scores on sex, race, subsidized lunch, special education, limited English proficiency, and fourth grade math and reading scores in the sample of students enrolled in traditional public schools. The value-added calculation is jackknifed to remove the influence of a student's own score.

Table 2: Relationship between distance to charter middle schools and fourth grade test scores

	Math scores		Reading scores	
	Distance to closest charter	Differential distance	Distance to closest charter	Differential distance
Controls	(1)	(2)	(3)	(4)
None	0.038	0.012	0.048	0.022
	(0.010)	(0.009)	(0.010)	(0.009)
Baseline characteristics	0.010	-0.004	0.015	0.001
	(0.009)	(0.008)	(0.008)	(0.008)
N	9,156			

Notes: This table reports coefficients from regressions of fourth grade math and reading scores on measures of distance to charter middle schools. Columns (1) and (3) show estimates from regressions of test scores on distance to the closest charter middle school measured in miles. Columns (2) and (4) display estimates from regressions of test scores on distance to the closest charter middle school minus distance to the closest traditional public middle school. The first row controls for no other covariates. The second row adds controls for sex, race, subsidized lunch, special education, limited English proficiency, and value-added of the closest traditional public middle school.

Table 3: Two-stage least squares estimates of charter school effects

		Second stage	
	First stage	Math scores	Reading scores
Instrument	(1)	(2)	(3)
Lottery offer	0.641	0.553	0.492
	(0.025)	(0.087)	(0.092)
N		1,601	
Differential distance	-0.026	0.453	0.380
	(0.002)	(0.212)	(0.217)
N		9,156	

Notes: This table reports two-stage least squares estimates of the effects of charter school attendance on eighth grade test scores. The endogenous variable is an indicator equal to one if a student attended a charter school at any time prior to the test. The first row instruments for charter attendance using a lottery offer indicator, and the second row instruments for charter attendance using distance to the closest charter school minus distance to the closest district school. Column (1) reports first stage impacts of the instruments on charter school attendance, and columns (2) and (3) report second stage effects on math and reading scores. The lottery sample is restricted to charter school applicants, while the distance sample includes all Boston students. The lottery models control for lottery portfolio indicators. The distance models control for sex, race, subsidized lunch, special education, limited English proficiency, the value-added of the closest traditional public school, and fourth grade math and reading scores.

Table 4: Charter school preference models

	Homogeneous charter schools	Heterogeneous charter schools: One normal distribution	Heterogeneous charter schools: Two-mass mixture of normals
	(1)	(2)	(3)
Number of parameters	43	51	53
Log-likelihood	-12,917.0	-12,062.3	-11,850.2
Likelihood ratio tests: χ^2 statistic (d.f.)	-	1,709.4 (8)	424.2 (2)
P-value		0.000	0.000

Notes: This table reports maximized log-likelihood values and numbers of parameters for three charter school preference models. Column (1) shows results from a model in which students view charter schools as homogenous. Columns (2) and (3) report results from models in which mean utilities vary across charter schools and students have idiosyncratic preferences for particular charters. Columns (1) and (2) use a single normal distribution to model θ_i, the unobserved taste common to all charter schools. Column (3) uses a two-mass mixture of normal distributions. Likelihood ratio test statistics in columns (2) and (3) come from tests of each model against the model in the previous column. The sample size for all models is $N = 9{,}156$.

Table 5: Charter school preference parameter estimates

	Charter school utility	Disutility of distance	Application costs	
			Log fixed cost	Log marginal cost
	(1)	(2)	(3)	(4)
Constant/main effect	-1.099	0.182	-2.098	-0.664
	(0.095)	(0.016)	(0.187)	(0.016)
Female	-0.046	-0.018	-0.006	0.027
	(0.097)	(0.011)	(0.130)	(0.026)
Black	-0.465	-0.156	1.286	-0.241
	(0.152)	(0.018)	(1.035)	(0.047)
Hispanic	-0.376	-0.128	1.713	-0.232
	(0.164)	(0.019)	(1.041)	(0.051)
Subsidized lunch	-0.298	-0.008	0.379	0.091
	(0.124)	(0.014)	(0.210)	(0.032)
Special education	-0.228	-0.025	0.098	0.025
	(0.137)	(0.015)	(0.162)	(0.039)
Limited English proficiency	-0.118	0.024	0.038	0.100
	(0.148)	(0.014)	(0.182)	(0.040)
Value-added of closest district school	-1.156	-0.177	0.429	-0.075
	(0.306)	(0.035)	(0.392)	(0.075)
Fourth grade math score	0.138	0.007	-0.028	0.009
	(0.070)	(0.008)	(0.092)	(0.019)
Fourth grade reading score	0.161	0.008	-0.067	0.022
	(0.073)	(0.008)	(0.097)	(0.019)
Distance squared	-	0.001	-	-
		(0.001)		

Notes: This table reports maximum simulated likelihood estimates of the parameters of student preferences for charter schools. Estimates come from the two-mass mixture model in column (3) of Table 4. Covariates are de-meaned in the estimation sample so that main effects are effects at the mean. Column (1) reports estimates of the utility function for charter attendance relative to traditional public school. The constant in this column is the average of school-specific utility intercepts. Column (2) reports estimates of the disutility of distance to school. The constant in this column is the main effect of differential distance between a charter school and the closest traditional public school. The subsequent rows show coefficients on interactions between differential distance and observed characteristics. The bottom row shows the effect of the difference in squared distances. Column (3) reports estimates of the charter school application fixed cost function, and column (4) reports estimates of the marginal cost function.

Table 6: Distributions of unobserved preferences for charter schools

		Homogeneous charter schools	Heterogeneous charter schools		
			One normal distribution	Two-mass mixture of normals	
				Type one	Type two
		(1)	(2)	(3)	(4)
Charter preference, θ_i	Mean	0.000	0.000	1.641	-1.174
		-	-	(0.105)	(0.078)
	Standard deviation	1.188	0.863	0.140	0.093
		(0.018)	(0.018)	(0.040)	(0.061)
	Type probabilitity	1.000	1.000	0.417	0.583
		-	-	(0.020)	(0.020)
Idiosyncractic preference, τ_{ij}	Standard deviation	0.000	0.598	0.936	
		-	(0.008)	(0.016)	
Application cost, η_{ij}	Standard deviation	0.113	0.038	0.024	
		(0.000)	(0.007)	(0.010)	

Notes: This table reports maximum simulated likelihood estimates of the distributions of unobserved charter school preferences. See the notes to Table 4 for a description of the preference models.

Table 7: Selection-corrected estimates of charter school effects on eighth grade test scores

	Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)
Constant/main effect	-0.390	0.705	-0.508	0.522
	(0.015)	(0.092)	(0.016)	(0.096)
Female	-0.024	0.060	0.184	-0.019
	(0.015)	(0.046)	(0.016)	(0.048)
Black	-0.193	0.250	-0.087	0.199
	(0.025)	(0.073)	(0.026)	(0.077)
Hispanic	-0.100	0.260	-0.041	0.243
	(0.025)	(0.078)	(0.027)	(0.081)
Subsidized lunch	-0.128	0.192	-0.126	0.149
	(0.022)	(0.056)	(0.023)	(0.059)
Special education	-0.370	0.097	-0.397	0.134
	(0.020)	(0.065)	(0.021)	(0.068)
Limited English proficiency	0.075	-0.091	0.044	-0.074
	(0.020)	(0.069)	(0.021)	(0.072)
Value-added of closest district school	0.136	0.003	0.113	-0.041
	(0.049)	(0.138)	(0.051)	(0.145)
Fourth grade math score	0.476	-0.122	0.165	-0.043
	(0.011)	(0.033)	(0.011)	(0.035)
Fourth grade reading score	0.066	-0.019	0.366	-0.078
	(0.011)	(0.034)	(0.011)	(0.036)
Charter school preference, θ_i	0.058	-0.096	0.046	-0.039
	(0.016)	(0.047)	(0.017)	(0.049)
Idiosyncratic preference, τ_{ij}	-	-0.017	-	0.010
		(0.052)		(0.055)
P-values: No selection on unobservables	0.001		0.051	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade math and reading test scores. Estimates come from regression of test scores on indicators for attendance at traditional public and charter schools, covariates and their interactions with charter attendance, and control functions correcting for selection on unobservables. The control functions are posterior means from the two-mass mixture model in column (3) of Table 4. Columns (1) and (3) display public school coefficients, while columns (2) and (4) display interactions with charter attendance. Main effects of charter attendance are enrollment-weighted averages of effects for the seven schools. P-values are from tests of the hypothesis that the control function coefficients equal zero. Standard errors are adjusted for estimation of the control functions.

Table 8: Test of restrictions implied by test score maximization

		Math scores		Reading scores	
	Preference coefficient	Test score gain coefficient	Ratio	Test score gain coefficient	Ratio
	(1)	(2)	(3)	(4)	(5)
Female	-0.046	0.060	-1.313	-0.019	0.407
Black	-0.465	0.250	-0.538	0.199	-0.429
Hispanic	-0.376	0.260	-0.691	0.243	-0.646
Subsidized lunch	-0.298	0.192	-0.644	0.149	-0.499
Special education	-0.228	0.097	-0.426	0.134	-0.588
Limited English proficiency	-0.118	-0.091	0.773	-0.074	0.626
Value-added of closest district school	-1.156	0.003	-0.003	-0.041	0.036
Fourth grade math score	0.138	-0.122	-0.883	-0.043	-0.315
Fourth grade reading score	0.161	-0.019	-0.117	-0.078	-0.481
Charter school preference, θ_i	1.000	-0.096	-0.096	-0.039	-0.039
Idiosyncratic preference, τ_{ij}	1.000	-0.017	-0.017	0.010	0.010
P-values: Test score maximization			0.000		0.000

Notes: This table reports tests of restrictions implied by test score maximization based on coefficients for observed characteristics and unobserved tastes. Estimates come from the two-mass mixture model in column (3) of Table 4. Column (1) reports the coefficient on each variable in the charter school utility function, and columns (2) and (4) report the additional test score gain resulting from charter attendance for students with each characteristic. Columns (3) and (5) report ratios of impacts on test score gains to impacts on preferences. P-values come from Wald tests of the hypothesis that all ratios in a column are equal and weakly positive.

Table 9: Comparison of reduced form and structural approaches to extrapolation

	Lottery IV	Covariate-based prediction			Model-based
	estimate	Subsidized lunch	4th grade score	Interacted covs.	prediction
Parameter	(1)	(2)	(3)	(4)	(5)
LATE	0.553	0.553	0.553	0.553	0.525
TOT		0.562	0.587	0.552	0.508
ATE		0.588	0.626	0.596	0.705
TNT		0.591	0.632	0.602	0.730

Notes: This table compares charter school treatment effects on eighth grade math scores obtained by covariate-based reweighting of lottery IV estimates vs. prediction from the structural selection model. Columns (1) through (4) are based on 2SLS models estimated in the lottery sample. Models in columns (2) through (4) interact charter school attendance with observed covariates, instrumenting with interactions of the lottery offer and covariates and controlling for covariate main effects and application portfolio indicators. Column (2) uses subsidized lunch status, column (3) uses terciles of baseline test score, and column (4) uses interactions of subsidized lunch, race, baseline score tercile and special education status. Column (5) reports predicted effects based on 1,000,000 simulations of the two-mass mixture model. The LATE in column (5) is a model-based prediction using the implicit weights underlying the IV estimate in column (1), as described in Appendix C. The TOT in column (5) is a predicted average effect for charter students, the TNT is a predicted effect for non-charter students, and the ATE is a predicted effect for the full population.

References

Abdulkadiroğlu, A., J. D. Angrist, S. Dynarski, T. J. Kane, and P. A. Pathak (2011): "Accountability and flexibility in public schools: evidence from Boston's charters and pilots," *Quarterly Journal of Economics*, 126(2), 699–748.

Abdulkadiroğlu, A., J. D. Angrist, P. D. Hull, and P. A. Pathak (2016): "Charters without lotteries: testing takeovers in New Orleans and Boston," *American Economic Review*, 16, 1878–1920.

Ajayi, K. F. (2013): "School choice and educational mobility: lessons from secondary school applications in Ghana," IED working paper no. 259.

Angrist, J. D. (1998): "Estimating the labor market impact of voluntary military service using Social Security data on military applicants," *Econometrica*, 66, 249–288.

Angrist, J. D., S. R. Cohodes, S. M. Dynarski, P. A. Pathak, and C. R. Walters (2016): "Stand and deliver: Effects of Boston's charter high schools on college preparation, entry and choice," *Journal of Labor Economics*, 34, 275–318.

Angrist, J. D., S. M. Dynarski, T. J. Kane, P. A. Pathak, and C. R. Walters (2012): "Who benefits from KIPP?" *Journal of Policy Analysis and Management*, 31, 837–860.

Angrist, J. D. and I. Fernandez-Val (2013): "ExtrapoLATE-ing: External validity and overidentification in the LATE framework," in *Advances in Economics and Econometrics*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, Cambridge University Press, vol. III of *Econometric Society Monographs, Econometrics*, chap. 11, 401–434.

Angrist, J. D. and G. W. Imbens (1995): "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *Journal of the American Statistical Association*, 90, 431–442.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996): "Identification of causal effects using instrumental variables," *Journal of the American Statistical Association*, 91, 444–455.

Angrist, J. D., P. A. Pathak., and C. R. Walters (2013): "Explaining charter school effectiveness," *American Economic Journal: Applied Economics*, 5, 1–27.

Arcidiacono, P. (2005): "Affirmative action in higher education: how do admission and financial aid rules affect future earnings?" *Econometrica*, 73, 1477–1524.

Arcidiacono, P., E. Aucejo, and V. J. Hotz (2016): "University differences in the graduation of minorities in STEM fields: evidence from California," *American Economic Review*, 106, 525–562.

Barseghyan, L., D. Clark, and S. Coate (2014): "Public school choice: an economic analysis," NBER working paper no. 20701.

Bergman, P. and I. McFarlin (2016): "An experimental analysis of cream skimming in charter schools," Mimeo, Columbia University.

Bloom, H. S. (1984): "Accounting for no-shows in experimental evaluation designs," *Evaluation Review*, 8, 225–246.

Booker, K., T. R. Sass, B. Gill, and R. Zimmer (2011): "The effects of charter high schools on educational attainment," *Journal of Labor Economics*, 29, 377–415.

Boston Municipal Research Bureau (2008): "The status of charter schools in Boston," http://www.bmrb.org /content/upload/Charter608.pdf.

Brand, J. E. and Y. Xie (2010): "Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education," *American Sociological Review*, 75, 273–302.

Butler, J., D. A. Carr, E. F. Toma, and R. Zimmer (2013): "Choice in a world of new school types," *Journal of Policy Analysis and Management*, 32, 785–806.

Cameron, S. V. and J. J. Heckman (1998): "Life cycle schooling and dynamic selection bias: models and evidence for five cohorts of American males," *Journal of Political Economy*, 106, 262–333.

Card, D. (1995): "Using geographic variation in college proximity to estimate the return to schooling," in *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*, ed. by L. Christofides, E. Grant, and R. Swidinsky, Toronto: University of Toronto Press.

Center for Research on Education Outcomes, CREDO (2013): "National charter school study," Mimeo, Stanford University.

Chade, H., G. Lewis, and L. Smith (2014): "Student portfolios and the college admissions problem," *Review of Economic Studies*, 81, 971–1002.

Chade, H. and L. Smith (2006): "Simultaneous search," *Econometrica*, 74, 1293–1307.

Chandra, A. and D. O. Staiger (2007): "Productivity spillovers in health care: evidence from the treatment of heart attacks," *Journal of Political Economy*, 115, 103–140.

Commonwealth of Massachusetts (2010): "An act relative to the achievement gap: turning around low-performing schools and promoting innovation for all," Http://www.mass.gov/Eoedu/docs/legislation_policy/20100125_ed_law_fact_sheet.pdf.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2016): "Who benefits from universal childcare? Estimating marginal returns to early childcare attendance," Working paper.

Curto, V. and R. G. Fryer (2014): "The potential of urban boarding schools for the poor: Evidence from SEED," *Journal of Labor Economics*, 32, 65–93.

Deming, D. J., J. S. Hastings, T. J. Kane, and D. O. Staiger (2014): "School choice, school quality, and postsecondary attainment," *American Economic Review*, 104, 991–1013.

Dillon, E. W. and J. A. Smith (2017): "Determinants of the match between student ability and college quality," *Journal of Labor Economics*, 35, 45–66.

Dobbie, W. and R. G. Fryer (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, 158–187.

——— (2013): "Getting beneath the veil of effective schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 5, 28–60.

——— (forthcoming): "The medium-term impacts of high-achieving charter schools," *Journal of Political Economy*.

Epple, D., R. Romano, and R. Zimmer (2015): "Charter schools: a survey of research on their characteristics and effectiveness," NBER working paper no. 21256.

Ferreyra, M. M. and G. Kosenok (2012): "Charter school entry and school choice: the case of Washington, DC," Mimeo, Carnegie Mellon University.

Fryer, R. G. (2014): "Injecting charter school best practices into traditional public schools: evidence from field experiments," *Quarterly Journal of Economics*, 129, 1355–1407.

Geweke, J., G. Gowrisankaran, and R. J. Town (2003): "Bayesian inference for hospital quality in a selection model," *Econometrica*, 71, 1215–1238.

Gleason, P., M. Clark, C. C. Tuttle, and E. Dwoyer (2010): "The evaluation of charter school impacts: final report," Mathematica report.

Hastings, J. S., T. J. Kane, and D. O. Staiger (2009): "Heterogeneous preferences and the efficacy of public school choice," NBER working papers no. 12145 and 11805.

Heckman, J. J. (1979): "Sample selection bias as a specification error," *Econometrica*, 47, 153–161.

——— (1990): "Varieties of selection bias," *American Economic Review: Papers & Proceedings*, 80, 313–318.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016): "Dynamic treatment effects," *Journal of Econometrics*, 191, 276–292.

——— (forthcoming): "Returns to education: the causal effects of education on earnings, health and smoking," *Journal of Political Economy*.

Heckman, J. J. and S. Navarro (2007): "Dynamic discrete choice and dynamic treatment effects," *Journal of Econometrics*, 136, 341–396.

Heckman, J. J. and B. Singer (1984): "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica*, 52, 271–320.

Heckman, J. J., J. L. Tobias, and E. Vytlacil (2001): "Four parameters of interest in the evaluation of social programs," *Southern Economic Journal*, 68, 210–223.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006): "Understanding instrumental variables estimates in models with essential heterogeneity," *Review of Economics and Statistics*, 88, 389–432.

Heckman, J. J. and E. Vytlacil (1999): "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences*, 96, 4730–4734.

——— (2005): "Structural equations, treatment effects, and econometric policy evaluation," *Econometrica*, 73, 669–738.

——— (2007a): "Econometric evaluation of social programs, part I: causal models, structural models and econometric policy evaluation," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6.

——— (2007b): "Econometric evaluation of social programs, part II: using the marginal treatment effect to organize alternative economic estimators to evaluate social programs and to forecast their effects in new environments," in *Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: Elsevier, vol. 6.

Hotz, V. J., G. W. Imbens, and J. H. Mortimer (2005): "Predicting the efficacy of future training programs using past experiences at other locations," *Journal of Econometrics*, 125, 241–270.

Howell, J. S. (2010): "Assessing the impact of eliminating affirmative action in higher education," *Journal of Labor Economics*, 28, 113–166.

Hoxby, C. M. (1998): "The economics of school reform," NBER reporter.

——— (2003): "School choice and school productivity: could school choice be a tide that lifts all boats?" in *The Economics of School Choice*, ed. by C. M. Hoxby, Chicago, IL: University of Chicago Press.

Hoxby, C. M. and C. Avery (2012): "The missing "one-offs": the hidden supply of high-achieving, low-income students," NBER working paper no. 18586.

Hoxby, C. M. and S. Murarka (2009): "Charter schools in New York City: who enrolls and how they affect their students' achievement," NBER working paper no. 14852.

Hoxby, C. M. and J. E. Rockoff (2004): "The impact of charter schools on student achievement," Harvard Institute of Economic Research working paper.

Imbens, G. W. and J. D. Angrist (1994): "Identification and estimation of local average treatment effects," *Econometrica*, 62, 467–475.

Imberman, S. A. (2011): "The effect of charter schools on achievement and behavior of public school students," *Journal of Public Economics*, 95, 850–863.

Kirkebøen, L. J., E. Leuven, and M. Mogstad (2016): "Field of study, earnings, and self-selection," *Quarterly Journal of Economics*, 131, 1057–1111.

Kline, P. and C. R. Walters (2016): "Evaluating public programs with close substitutes: the case of Head Start," *Quarterly Journal of Economics*, 131, 1795–1848.

Kodde, D. A. and F. C. Palm (1986): "Wald criteria for jointly testing equality and inequality restrictions," *Econometrica*, 54, 1243–1248.

Ladd, H. F. (2002): *Market-based Reforms in Urban Education*, Washington, DC: Economic Policy Institute.

Massachusetts Department of Elementary and Secondary Education (2011): "Understanding charter school reimbursements," http://www.doe.mass.edu/charter/finance/tuition/Reimbursements.html.

Matzkin, R. L. (1992): "Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models," *Econometrica*, 60, 239–270.

——— (2003): "Nonparametric estimation of nonadditive random functions," *Econometrica*, 71, 1339–1375.

Mehta, N. (forthcoming): "Competition in public school districts: charter school entry, student sorting, and school input determination," *International Economic Review*.

Mulligan, C. B. and Y. Rubinstein (2008): "Selection, investment, and women's relative wages over time," *Quarterly Journal of Economics*, 123, 1061–1110.

Murphy, K. M. and R. H. Topel (1985): "Estimation and inference in two-step econometric models," *Journal of Business and Economic Statistics*, 3, 370–379.

Neal, D. (1997): "The effects of Catholic secondary schooling on educational achievement," *Journal of Labor Economics*, 15, 98–123.

——— (2004): "The measured black-white wage gap among women is too small," *Journal of Political Economy*, 112, S1–S28.

Nielsen, E. R. (2015): "The income-achievement gap and adult outcome inequality," Federal Reserve Finance and Economics Discussion Series no. 2015-041.

Pathak, P. A. and T. Sönmez (2008): "Leveling the playing field: sincere and sophisticated players in the Boston mechanism," *American Economic Review*, 98, 1636–1652.

Rothstein, J. (2006): "Good principals or good peers? Parental valuation of school characteristics, Tiebout equilibrium, and the incentive effects of competition among jurisdictions," *American Economic Review*, 96, 1333–1350.

Roy, A. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3, 135–146.

Scharfenberg, D. (2016): "Mass. voters reject ballot question on charter schools," *Boston Globe*, November 8th.

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

Vytlacil, E. (2002): "Independence, Monotonicity, and Latent Index Models: An Equivalence Result," *Econometrica*, 70, 331–341.

Wald, A. (1940): "The fitting of straight lines if both variables are subject to error," *The Annals of Mathematical Statistics*, 11, 284–300.

Willis, R. J. and S. Rosen (1979): "Education and self-selection," *Journal of Political Economy*, 87, S7–S36.

Wilson, S. F. (2008): "Success at scale in charter schooling," American Enterprise Institute Future of American Education Project, working paper no. 2008-02.

Appendix A: Semi-Parametric Identification

This appendix considers semi-parametric identification of a model with one charter school and no heterogeneity in unobserved application costs. I first derive students' optimal application and attendance rules in this version of the model. I then then show that these rules imply a representation equivalent to the single-spell discrete duration model analyzed by Heckman and Navarro (2007). Finally, I apply Theorem 2 in Heckman and Navarro (2007) to give conditions under which the model is semi-parametrically identified.

Application and Attendance Choices

In a model with one charter school, the expected utility associated with receiving a charter school offer is given by

$$w(1|X_i, D_i, \psi_{i1}) = E\left[\max\left\{v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \xi_{i1}, 0\right\}|X_i, D_i, \psi_{i1}\right],$$

while the expected utility of receiving no offer equals zero. The probability of an offer is π_1, and the cost of applying is $c(X_i, \eta_i)$. The optimal application rule is therefore

$$A_{i1} = 1\left\{\pi w(1|X_i, D_i, \psi_i) - c(X_i, \eta_i) > 0\right\}$$

$$= 1\left\{\pi_1 h\left(v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1}\right) > c(X_i, \eta_i)\right\},$$

where

$$h(v) \equiv E\left[\max\left\{v + \xi_{i1}, 0\right\}\right].$$

We can rewrite the function $h(v)$ as

$$h(v) = F_{-\xi_1}(v) \times [v - K_{-\xi_1}(v)],$$

where $F_{-\xi_1}(v)$ is the CDF of $-\xi_{i1}$ evaluated at v and $K_{-\xi_1}(v)$ is the conditional expectation of $-\xi_{i1}$ truncated from above at v. When ξ_{i1} has full support on the real line, it is straightforward to show that this function satisfies $h(v) > 0\ \forall v$, $\lim_{t\to-\infty} h(v) = 0$, and $h'(v) = F_{-\xi_1}(v) > 0\ \forall v$.

Assuming $c(X_i, \eta_i)$ is always positive, the application rule can then be rewritten

$$A_{i1} = 1\left\{v_1(X_i, D_{i1}, D_{i0}) - h^{-1}\left(c(X_i, \eta_i)/\pi_1\right) + \psi_{i1} > 0\right\}.$$

Due to the non-linearity of the $h^{-1}(\cdot)$ function the left-hand side of this inequality is not additively separable in the observables (X_i, D_{i1}, D_{i0}) and the unobservables (ψ_{i1}, η_i). This is true even when the cost function itself is separable as $c(x, \eta) = c_1(x) + c_2(\eta)$.

To obtain an additively separable representation I consider a special case where there is no cost heterogeneity, so $c(x, \eta) = c(x)$. In this case the application rule can be rewritten as a separable threshold crossing model:

$$A_i = 1\left\{\Omega(X_i, D_{i1}, D_{i0}) < \psi_{i1}\right\},$$

with $\Omega(x, d_1, d_0) = h^{-1}(c(x)/\pi_1) - v_1(x, d_1, d_0)$.

A student attends the charter school if she applies, receives a lottery offer, and the final utility for the charter school exceeds the final utility for public school. Then we can write

$$S_i = A_{i1} \times Z_{i1} \times 1\{v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \xi_{i1} > 0\}.$$

$$= A_{i1} \times Z_{i1} \times 1\{-v_1(X_i, D_{i1}, D_{i0}) < \zeta_{i1}\},$$

where $\zeta_{i1} = \psi_{i1} + \xi_{i1}$.

Identification

To analyze identification of the model, it is useful to show it is a special case of the discrete duration model analyzed by Heckman and Navarro (2007). Define

$$T_i = 1 + A_{i1} + Z_{i1} + S_i.$$

T_i may be viewed as the number of periods that student i participates in the charter enrollment process. If she chooses not to apply, then $T_i = 1$. If she applies but loses the lottery, then $T_i = 2$. If she applies, wins the lottery, but turns down the charter offer, then $T_i = 3$. If she applies, wins the lottery, and accepts the offer, then $T_i = 4$, which is the maximum duration.

Let B_{it} denote an indicator equal to one if student i decides to stop at step t, observed only when $T_i \geq t$. Each step in this process obeys a separable threshold-crossing model. We have $B_{i1} = 1\{\Omega(X_i, D_{i1}, D_{i0}) \geq \psi_{i1}\}$. At $t = 2$, the student exits when she loses the lottery, which occurs with probability $1 - \pi_1$ at this stage. Hence we can write $B_{i2} = 1\{1 - \pi_1 \geq \nu_{i1}\}$, where we have normalized $\nu_{i1} \sim U(0, 1)$ and ν_{i1} is independent of all other variables since the lottery is random. At $t = 3$, we have $B_{i3} = 1\{-v_1(X_i, D_{i1}, D_{i0}) \geq \zeta_{i1}\}$. We define $B_{i0} = 0$ and $B_{i4} = 1 - B_{i3}$. Finally, potential outcomes for $T_i \in \{1, 2, 3\}$ equal Y_{i0}, while the potential outcome for $T_i = 4$ is Y_{i1}.

This argument shows that the one-charter choice model is a special case of the model in Heckman and Navarro (2007). Their Theorem 2 gives identification of the joint distribution of latent utilities and each potential outcome in this model. The following is a restatement of this theorem, slightly adapted to the problem at hand.

Theorem A1: *Suppose charter application, lottery offer, and attendance rules are given by B_{i1}, B_{i2} and B_{i3} as defined above, and that $\Omega(X_i, D_{i1}, D_{i0})$ and $v_1(X_i, D_{i1}, D_{i0})$ are elements of the Matzkin (1992) class of functions (see Appendix A of Heckman and Navarro, 2007 for a definition of this class). Write potential outcomes as $Y_{ij} = \mu_j(X_i) + \epsilon_{ij}$ with $E[\epsilon_{ij}|X_i] = 0$ for $j \in \{0, 1\}$. Suppose a random sample of data on $(X_i, D_{i1}, D_{i0}, A_{i1}, Z_{i1}, S_i, Y_i)$ is available. Assume:*

1. *$(\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1})$ are continuous mean-zero random variables with finite variances and supports with upper limits $(\bar{\epsilon}_1, \bar{\epsilon}_0, \bar{\psi}_1, \bar{\zeta}_1)$ and lower limits $(\underline{\epsilon}_1, \underline{\epsilon}_0, \underline{\psi}_1, \underline{\zeta}_1)$. These conditions also hold for each component of each subvector.*

2. $(\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1}) \perp\!\!\!\perp (X_i, D_{i1}, D_{i0})$.

3. $Supp(\mu_j(X_i), \Omega(X_i, D_{i1}, D_{i0}), v_1(X_i, D_{i1}, D_{i0})) = Supp(\mu_j(X_i)) \times Supp(\Omega(X_i, D_{i1}, D_{i0})) \times Supp(v_1(X_i, D_{i1}, D_{i0}))$.

4. $Supp\,(\Omega(X_i, D_{i1}, D_{i0}), -v_1(X_i, D_{i1}, D_{i0})) \supseteq Supp\,(\psi_{i1}, \zeta_{i1})$.

5. $\nu_{i1} \perp\!\!\!\perp (\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1}, X_i, D_{i1}, D_{i0})$.

Then we can identify $\mu_1(x)$, $\mu_0(x)$, $\Omega(x, d_1, d_0)$, $v_1(x, d_1, d_0)$, and the joint distribution functions $F_{\psi_1\zeta_1\epsilon_1}(\psi_1, \zeta_1, \epsilon_1)$ and $F_{\psi_1\zeta_1\epsilon_0}(\psi_1, \zeta_1, \epsilon_0)$ up to scale if the Matzkin class is specified up to scale, and exactly if a specific normalization is used.

This theorem follows exactly from the argument for Theorem 2 in Heckman and Navarro (2007). The only slight twist is that there is not full support for the lottery offer "choice" index at $t = 2$, but this is irrelevant since by condition 5 ν_{i1} is independent of all other variables in the model. The remaining primitives of the model are then identified by using the joint distribution of (ψ_{i1}, ζ_{i1}) to recover the marginal distribution of $\xi_{i1} = \zeta_{i1} - \psi_{i1}$. The probability π_1 is identified by the offer rate among lottery applicants, and the function $h(\cdot)$ is determined by the distribution of ξ_{i1}. We can then recover $c(x) = \pi_1 h\left(-v_1(x, d_1, d_0) - \Omega(x, d_1, d_0)\right)$ for any (d_1, d_0).

A final observation is that though the assumptions of Theorem A1 are sufficient for identification, they are stronger than necessary for this special case: the model is overidentified. To see this, note that Heckman and Navarro (2007) establish identification of separate potential outcome distributions corresponding to each node in the duration model. Since charter applications and lottery offers are assumed to have no direct effect on outcomes, however, the potential outcomes corresponding to $T_i = 1$, $T_i = 2$ and $T_i = 3$ are known to be the same in this case. With full support of $\Omega(X_i, D_{i1}, D_{i0})$ and $v_1(X_i, D_{i1}, D_{i0})$, there are multiple ways to identify $\mu_0(x)$: we could look at individuals in a limit set with zero probability of applying to charter schools, or we could look at individuals in a limit set with an application probability equal to one and a conditional attendance probability equal to one who are lotteried back into traditional public schools. In principle one could use this fact to weaken the support conditions without sacrificing identification.

Appendix B: Model Fit

The model estimated here fits the data well. This can be seen in Table A4 and Figure A1, which compare observed and model-predicted patterns of heterogeneity across choices, outcomes, and schools. Panel A of Figure A1 splits the sample into deciles 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 near zero 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 ($p = 0.04$), 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 deciles 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. Most points lie close to the 45-degree line and an F-test does not reject the hypothesis that the model fits all moments. Predicted scores exhibit substantial dispersion and there is significant overlap between predictions for charter and non-charter students.

Finally, Table A4 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 application rates and underpredicts the offer takeup rate. 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 reasonably accurate description of heterogeneity along many dimensions.

Appendix C: 2SLS Weights

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

$$Y_i = \beta C_i + \sum_a \gamma_a 1\{A_i = a\} + \epsilon_i,$$

$$C_i = \pi Z_i^{max} + \sum_a \lambda_a 1\{A_i = a\} + \eta_i,$$

where Y_i is a test score, C_i is a charter attendance dummy, Z_i^{max} is a lottery offer dummy, and the sample is assumed to be restricted to lottery applicants. The reduced form corresponding to this system is

$$Y_i = \rho Z_i{}^{max} + \sum \tau_a 1\{A_i = a\} + u_i.$$

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, we have

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

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

where ρ_a and π_a denote coefficients from regressions of Y_i and C_i on Z_i^{max} within lottery portfolio a, and

$$w_a = Pr\left[A_i = a\right] \times Var(Z_i^{max}|A_i = a).$$

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:

$$\beta = \frac{\rho}{\pi}$$

$$= \frac{\sum_a w_a \rho_a}{\sum_a w_a \pi_a}$$

$$= \frac{\sum_a (w_a \pi_a)(\rho_a/\pi_a)}{\sum_a w_a \pi_a}$$

$$= \sum_a \left(\frac{\omega_a}{\sum_{a'} \omega_{a'}}\right) IV_a,$$

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

$$\omega_a = w_a \pi_a$$

$$= Pr\left[A_i = a\right] \times Var\left(Z_i^{max}|A_i = a\right) \times \left(Pr\left[C_i = 1|Z_i^{max} = 1, A_i = a\right] - Pr\left[C_i = 1|Z_i^{max} = 0, A_i = a\right]\right).$$

This argument shows that 2SLS estimation with application portfolio fixed effects generates a weighted average of portfolio-specific IV coefficients with weights proportional to the product of sample size, the variance of the offer, and the first stage shift in charter attendance resulting from an 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 first stage generates weights proportional to $w_a \pi_a^2$ rather than $w_a \pi_a$.

Appendix D: 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 in which students make utility-maximizing choices as described in Section 3, and schools set admission probabilities to maximize enrollment subject to capacity constraints.

The time of the game follows Figure 1. 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 "probability takers" in the first stage, in the sense that they do not expect schools to react to their application choices when setting admission probabilities. 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 $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^*)$, where q_a is defined as in Section 5 and now explicitly depends on the vector of admission probabilities students expect to face in each lottery.

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

3. After receiving the offer vector z, student i chooses school j with probability $p(j|z, X_i, D_i, \theta_i, \tau_i)$, as defined in Section 5.

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_j denote the capacity of school j, which is the maximum share of students that can attend school j. 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 $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e)$. In addition, suppose that schools other than j admit students with probability π_{-j}. If school j admits students with probability π_j in the second stage, its enrollment is given by

$$e_j(\pi_j, \pi_{-j}, \pi^e) = E\left[\sum_{a\in\{0,1\}^J}\sum_{z\in\{0,1\}^J} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f(z|a; \pi_j, \pi_{-j}) p(j|z, X_i, D_i, \theta_i, \tau_i)\right],$$

where $f(z|a; \pi_j, \pi_{-j})$ is the probability mass function for offers, now explicitly written as a function of admission probabilities. School j choose π_j to solve

$$\max_{\pi_j\in[0,1]} e_j(\pi_j, \pi_{-j}, \pi^e) \quad s.t. \quad e_j(\pi_j, \pi_{-j}, \pi^e) \leq \Lambda_j. \tag{19}$$

The best response function $\pi_j^{BR}(\pi_{-j}, \pi^e)$ is the solution to problem (19). The optimal admission probability sets school j's enrollment equal to its capacity if possible. The following equation implicitly defines π_j^{BR} at interior solutions:

$$E\left[\sum_a\sum_z q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f(z|a; \pi_j^{BR}, \pi_{-j}) p(j|z, X_i, D_i, \theta_i, \tau_i)\right] = \Lambda_j.$$

Noting that $p(j|z, x, d, \theta, \tau) = 0$ when $z_j = 0$ and setting the probability mass function for the offer at school j to $a_j\pi_j$, this equation can be rewritten

$$E\left[\sum_{a:a_j=1}\sum_{z:z_j=1} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f_{-j}(z_{-j}|a_{-j}; \pi_{-j}) \pi_j^{BR} p(j|X_i, D_i, \theta_i, \tau_i)\right] = \Lambda_j,$$

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

$$\pi_j^{BR} = \frac{\Lambda_j}{E\left[\sum_{a:a_j=1}\sum_{z:z_j=1} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f_{-j}(z_{-j}|a_{-j}; \pi_{-j}) p(j|X_i, D_i, \theta_i, \tau_i)\right]}$$

$$\equiv \Gamma_j(\pi_{-j}, \pi^e).$$

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

$$\pi_j^{BR}(\pi_{-j}, \pi^e) = \min\{\Gamma_j(\pi_{-j}, \pi^e), 1\}.$$

Existence of Equilibrium

Let $\pi^{BR} : [0,1]^J \to [0,1]^J$ denote the following vector-valued function:

$$\pi^{BR}(\pi) \equiv \left(\pi_1^{BR}(\pi_{-1}, \pi), \ldots, \pi_J^{BR}(\pi_{-J}, \pi)\right).$$

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 D1: *There exists a* $\pi^* \in [0,1]^J$ *such that* $\pi^{BR}(\pi^*) = \pi^*$.

Proof: Note that $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi)$ is continuous in π and strictly positive, $p(j|X_i, D_i, \theta_i, \tau_i)$ is strictly positive when $z_j = 1$, and $f_{-j}(z_{-j}|a_{-j}; \pi_{-j})$ is continuous in π_{-j} and sums to one for each a_{-j}, so the denominator of Γ_j is always non-zero and continuous in π. π_j^{BR} is therefore a composition of continuous functions, and is continuous. Then π^{BR} is a continuous function that maps the compact, convex set $[0,1]^J$ to itself. Brouwer's Fixed Point Theorem immediately applies and π^{BR} has at least one fixed point in $[0,1]^J$.

Uniqueness of Equilibrium

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

$$\ell_j(\pi) \equiv \pi_j - \min\{\Gamma_j(\pi_{-j}, \pi), 1\}$$

and let $\ell(\pi) = (\ell_1(\pi), \ldots, \ell_J(\pi))$. A vector π^* supporting an equilibrium satisfies $\ell(\pi^*) = 0$. A sufficient condition for a unique zero of this function and therefore a unique equilibrium is that the Jacobian of $\ell(\pi)$ is a positive dominant diagonal matrix. This requires the following two conditions to hold at every value of $\pi \in [0,1]^J$:

$$\text{1a. } \frac{\partial \ell_j}{\partial \pi_j} > 0 \quad \forall j.$$

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

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_j for a sufficiently small positive π_j. When $\pi_j > 0$, we can write Γ_j as

$$\Gamma_j(\pi_{-j}, \pi) = \frac{\Lambda_j \pi_j}{e_j(\pi_j, \pi_{-j}, \pi)}$$

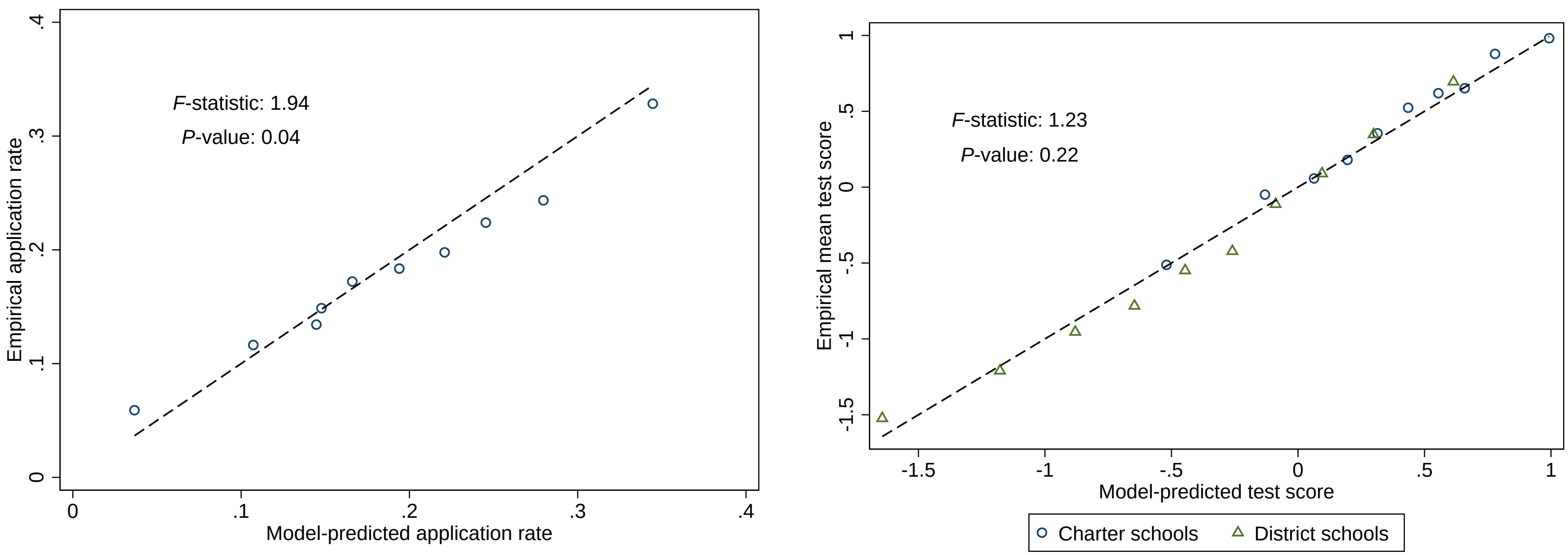
It follows that conditions 1a and 2a are equivalent to the following conditions on the model's enrollment elasticities:

$$\text{1b. } \frac{\partial \log e_j}{\partial \log \pi_j} > \left(\frac{\Lambda_j - e_j}{\Lambda_j} \right) \quad \forall j$$

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

Condition 1b is more likely to hold throughout the parameter space when demand for charter schools is strong, so that $e_j(\pi_j, \pi_{-j}, \pi) > \Lambda_k$ 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 j 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 $\pi^{BR}(\pi)$. Experimenting with starting values never produced more than one equilibrium in any counterfactual.

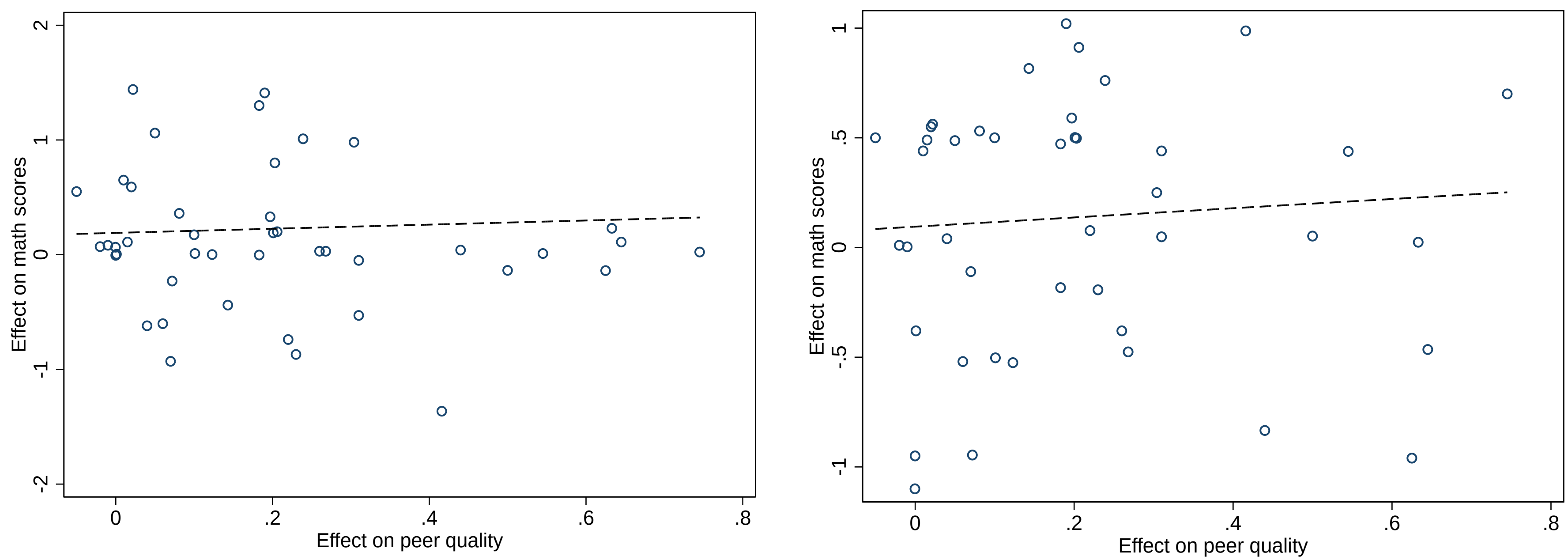
Figure A1: Model fit



A. Application probability *B. Eighth grade math scores*

Notes: This table compares charter application rates and mean test scores to predictions from the two-mass mixture model. Panel A splits the sample into deciles 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 eighth grade math scores conditional on each student's observed school choice. The sample is split into deciles of this predicted score separately for charter and traditional public schools. Circles plot mean observed scores against mean model predictions for charter schools, while triangles plot corresponding observed and predicted means for traditional public schools. Predictions are averages over 1,000,000 simulations of the two-mass mixture model, with covariates and spatial locations drawn with replacement from the empirical joint distribution. Dashed lines show the 45-degree line. F-statistics and p-values are from tests of the hypothesis that these lines fit all points perfectly up to sampling error, treating the model predictions as fixed.

Figure A2: Relationship between charter lottery effects on test scores and peer quality



A. Math *B. Reading*

Notes: This figure plots coefficients from regressions of sixth grade test scores on charter 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 fourth grade test score of the students with whom he or she attends sixth grade. The lines come from OLS regressions of test score effects on peer quality effects, weighting by lottery sample size. The slopes are 0.19 (s.e. = 0.42) for math and 0.21 (s.e. = 0.31) for reading.

Table A1: Boston charter middle schools

	Grade coverage (1)	Years open (2)	Records available (3)	Oversubscribed cohorts (4)
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	-

Notes: This table lists charter middle schools in Boston, Massachusetts. Schools are included if serve traditional student populations, accept students in fifth or sixth grade, and operated for cohorts attending fourth grade between 2006 and 2009. Column (2) lists the opening and (where relevant) closing year for each school. Column (3) indicates whether applicant records were available for cohorts attending fourth grade between 2006 and 2009, and column (4) lists the cohorts for which lotteries were held during this period.

Table A2: Covariate balance

	Differential
	(1)
Female	-0.017
	(0.035)
Black	-0.006
	(0.033)
Hispanic	0.029
	(0.031)
Subsidized lunch	-0.003
	(0.031)
Special education	-0.001
	(0.027)
Limited English proficiency	0.001
	(0.022)
Value-added of public schools in zip code	-0.005
	(0.012)
Fourth grade math score	-0.060
	(0.069)
Fourth grade reading score	0.026
	(0.073)
Miles to closest charter school	-0.045
	(0.076)
Miles to closest district school	0.013
	(0.021)
Joint *p* -value	0.846
N	1601

Notes: This table reports coefficients from regressions of pre-lottery characteristics on a charter lottery offer dummy, controlling for lottery portfolio indicators. 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 observed in eighth grade for students attending fourth grade in Boston between 2006 and 2009. Column (1) shows the follow-up rate for the full sample as well as the difference in followup rates between students with above-median and below-median predicted eighth grade math scores. Predicted scores are fitted values from regressions of eighth 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.

Table A4: Model fit

Charter applications and attendance			Eighth grade math scores				
				Lottery winners		Lottery losers	
	Model	Data		Model	Data	Model	Data
	(1)	(2)		(3)	(4)	(5)	(6)
Apply to charter	0.213	0.175	Charter applicants	0.181	0.236	-0.191	-0.151
Apply to more than one	0.068	0.046	Non-applicants	-	-	-0.459	-0.453
Offer takeup rate	0.774	0.836					
Apply to:			Applicants to:				
Charter 1	0.033	0.031	Charter 1	0.209	0.331	-0.055	0.050
Charter 2	0.046	0.036	Charter 2	0.193	0.405	-0.049	0.110
Charter 3	0.057	0.048	Charter 3	0.106	0.131	-0.051	0.013
Charter 4	0.052	0.039	Charter 4	0.234	0.401	-0.198	-0.207
Charter 5	0.062	0.051	Charter 5	0.117	0.017	-0.072	-0.145
Charter 6	0.019	0.017	Charter 6	0.234	0.241	0.009	-0.073
Charter 7	0.015	0.016	Charter 7	0.471	0.443	0.011	0.088
			χ^2 statistic (d.f.):	40.2 (27)			
			P -value:	0.049			

Notes: This table compares choices and outcomes to predictions from the two-mass mixture selection model. Model predictions are averages over 1,000,000 simulations of the model, with covariates and spatial locations drawn with replacement from the empirical joint distribution. The χ^2 statistic and *p* -value come from a Wald test of the hypothesis that the model fits all observed moments up to sampling error, treating the model predictions as fixed.

Table A5: Estimates of school-specific parameters

	Admisson probability (1)	Mean utility (2)	Math effect (3)	Reading effect (4)
Charter school 1	0.516	-0.736	0.579	0.515
	(0.064)	(0.096)	(0.104)	(0.108)
Charter school 2	0.390	-0.931	0.687	0.610
	(0.057)	(0.103)	(0.105)	(0.110)
Charter school 3	0.653	-0.763	0.619	0.455
	(0.039)	(0.092)	(0.101)	(0.106)
Charter school 4	0.706	-1.264	0.818	0.499
	(0.051)	(0.099)	(0.104)	(0.109)
Charter school 5	0.394	-0.259	0.560	0.234
	(0.074)	(0.095)	(0.099)	(0.103)
Charter school 6	0.824	-2.072	0.772	0.649
	(0.055)	(0.121)	(0.128)	(0.134)
Charter school 7	0.875	-1.670	1.078	0.942
	(0.039)	(0.120)	(0.120)	(0.126)
P-values: no heterogeneity	0.000	0.000	0.000	0.000

Notes: This table reports estimates of school-specific lottery admission probabilities, mean utilities, and test score effects. Estimates come from the two-mass mixture model in column (3) of Table 4. Column (1) shows each charter school's admission probability averaged across applicant cohorts. Column (2) displays mean utility estimates from the two-mass mixture model. Column (3) shows estimates of average causal effects on eighth grade math scores. *P*-values come from Wald tests of the hypothesis that all parameters in a column are equal.

Table A6: Selection-corrected estimates of charter school effects on sixth and seventh grade test scores

	Sixth grade				Seventh grade			
	Math scores		Reading scores		Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant/main effect	-0.492	0.650	-0.542	0.244	-0.429	0.627	-0.513	0.4307
	(0.014)	(0.085)	(0.015)	(0.090)	(0.014)	(0.088)	(0.015)	(0.089)
Female	-0.001	0.006	0.157	-0.059	0.006	0.130	0.217	-0.031
	(0.014)	(0.043)	(0.015)	(0.045)	(0.015)	(0.044)	(0.015)	(0.045)
Black	-0.208	0.163	-0.155	0.150	-0.207	0.237	-0.098	0.161
	(0.023)	(0.068)	(0.024)	(0.073)	(0.024)	(0.071)	(0.024)	(0.072)
Hispanic	-0.104	0.236	-0.094	0.128	-0.102	0.201	-0.030	0.192
	(0.024)	(0.072)	(0.025)	(0.077)	(0.025)	(0.075)	(0.025)	(0.076)
Subsidized lunch	-0.146	0.163	-0.143	0.061	-0.146	0.174	-0.132	0.170
	(0.020)	(0.053)	(0.021)	(0.056)	(0.021)	(0.054)	(0.021)	(0.055)
Special education	-0.342	-0.015	-0.326	0.006	-0.353	-0.023	-0.399	0.006
	(0.018)	(0.059)	(0.019)	(0.063)	(0.019)	(0.061)	(0.019)	(0.062)
Limited English proficiency	0.041	-0.129	-0.046	-0.005	0.081	-0.073	-0.017	-0.028
	(0.019)	(0.065)	(0.020)	(0.069)	(0.020)	(0.066)	(0.020)	(0.067)
Value-added of closest district school	0.096	-0.015	0.116	-0.115	0.099	-0.062	-0.018	0.108
	(0.046)	(0.130)	(0.049)	(0.138)	(0.047)	(0.133)	(0.048)	(0.135)
Fourth grade math score	0.569	-0.163	0.175	-0.083	0.487	-0.103	0.170	-0.064
	(0.010)	(0.031)	(0.010)	(0.033)	(0.010)	(0.032)	(0.010)	(0.032)
Fourth grade reading score	0.095	0.036	0.460	0.021	0.093	-0.039	0.377	-0.054
	(0.010)	(0.032)	(0.010)	(0.034)	(0.010)	(0.033)	(0.010)	(0.033)
Charter school preference, θ_i	0.055	-0.091	0.023	-0.011	0.046	-0.071	0.039	-0.057
	(0.015)	(0.044)	(0.016)	(0.047)	(0.016)	(0.046)	(0.016)	(0.046)
Idiosyncratic preference, τ_{ij}	-	-0.025	-	0.061	-	0.034	-	0.072
		(0.049)		(0.052)		(0.050)		(0.051)
P-value: No selection on unobservables	0.000		0.438		0.021		0.137	
N	10,122				9,731			

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on sixth and seventh grade test scores. Each pair of columns shows results from a regression of test scores on indicators for attendance at traditional public and charter schools, covariates and their interactions with charter attendance, and control functions correcting for selection on unobservables. The control functions are posterior means from the two-mass mixture model in column (3) of Table 4. Columns (1)-(4) show estimates for sixth grade, while columns (5)-(8) show estimates for seventh grade. P-values are from tests of the hypothesis that the control function coefficients equal zero. Standard errors are adjusted for estimation of the control functions.

Table A7: Selection-corrected estimates of charter school effects on eighth grade test scores for alternative preference models

	Homogeneous charter schools				Heterogeneous charter schools, single normal distribution			
	Math scores		Reading scores		Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant/main effect	-0.433	0.674	-0.510	0.4436	-0.436	0.716	-0.510	0.5807
	(0.008)	(0.077)	(0.008)	(0.081)	(0.008)	(0.067)	(0.008)	(0.070)
Female	-0.025	0.058	0.184	-0.019	-0.025	0.058	0.184	-0.018
	(0.015)	(0.046)	(0.016)	(0.048)	(0.015)	(0.046)	(0.016)	(0.048)
Black	-0.196	0.255	-0.087	0.200	-0.197	0.253	-0.087	0.197
	(0.025)	(0.073)	(0.026)	(0.077)	(0.025)	(0.073)	(0.026)	(0.077)
Hispanic	-0.102	0.257	-0.041	0.242	-0.103	0.257	-0.041	0.245
	(0.025)	(0.077)	(0.027)	(0.081)	(0.026)	(0.077)	(0.027)	(0.081)
Subsidized lunch	-0.129	0.180	-0.127	0.139	-0.127	0.185	-0.127	0.159
	(0.022)	(0.056)	(0.023)	(0.058)	(0.022)	(0.056)	(0.023)	(0.058)
Special education	-0.372	0.095	-0.397	0.140	-0.370	0.093	-0.397	0.136
	(0.020)	(0.065)	(0.021)	(0.068)	(0.020)	(0.065)	(0.021)	(0.068)
Limited English proficiency	0.077	-0.100	0.044	-0.080	0.077	-0.095	0.044	-0.067
	(0.020)	(0.069)	(0.021)	(0.072)	(0.020)	(0.069)	(0.021)	(0.072)
Value-added of closest district school	0.136	-0.009	0.112	-0.045	0.140	-0.007	0.113	-0.033
	(0.049)	(0.138)	(0.051)	(0.144)	(0.049)	(0.138)	(0.051)	(0.144)
Fourth grade math score	0.477	-0.120	0.165	-0.043	0.476	-0.121	0.165	-0.045
	(0.011)	(0.033)	(0.011)	(0.035)	(0.011)	(0.033)	(0.011)	(0.035)
Fourth grade reading score	0.065	-0.014	0.366	-0.075	0.065	-0.015	0.366	-0.080
	(0.011)	(0.034)	(0.011)	(0.035)	(0.011)	(0.034)	(0.011)	(0.035)
Charter school preference, θ_i	0.079	-0.099	0.036	-0.068	0.067	-0.084	0.060	-0.068
	(0.026)	(0.055)	(0.028)	(0.058)	(0.029)	(0.045)	(0.031)	(0.050)
Idiosyncratic preference, τ_{ij}	-	-	-	-	-	0.022	-	0.001
						(0.116)		(0.121)
P-values: No selection on unobservables	0.007		0.366		0.032		0.169	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade test scores using control functions based on the preference models in columns (1) and (2) of Table 4. Each pair of columns shows results from a regression of test scores on indicators for attendance at traditional public and charter schools, covariates and their interactions with charter attendance, and control functions correcting for selection on unobservables. Control functions in columns (1)-(4) are posterior means from the model in column (1) of Table 4. Control functions in columns (5)-(8) are posterior means from the model in column (2) of Table 4. Columns (1), (3), (5) and (7) display public school coefficients, while columns (2), (4), (6), and (8) display interactions with charter attendance. Main effects in columns (6) and (8) are enrollment-weighted averages of effects for the seven schools. P-values are from tests of the hypothesis that the control function coefficients equal zero. Standard errors are adjusted for estimation of the control functions.

Table A8: Selection-corrected estimates of charter school effects on eighth grade math scores for alternative test score transformations

	Percentile rank		Change in percentile rank		Log percentile rank	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)
Constant/main effect	0.489	0.211	-0.005	0.209	-1.027	0.5726
	(0.005)	(0.028)	(0.005)	(0.027)	(0.017)	(0.103)
Female	-0.007	0.019	-0.007	0.017	-0.017	0.045
	(0.005)	(0.014)	(0.005)	(0.014)	(0.017)	(0.051)
Black	-0.059	0.075	-0.051	0.071	-0.111	0.139
	(0.007)	(0.022)	(0.007)	(0.022)	(0.028)	(0.082)
Hispanic	-0.032	0.077	-0.026	0.076	-0.018	0.100
	(0.008)	(0.023)	(0.008)	(0.023)	(0.028)	(0.087)
Subsidized lunch	-0.039	0.057	-0.031	0.049	-0.076	0.126
	(0.007)	(0.017)	(0.006)	(0.017)	(0.024)	(0.063)
Special education	-0.106	0.023	-0.111	0.016	-0.471	0.245
	(0.006)	(0.020)	(0.006)	(0.020)	(0.022)	(0.073)
Limited English proficiency	0.022	-0.028	0.022	-0.027	0.070	-0.047
	(0.006)	(0.021)	(0.006)	(0.021)	(0.023)	(0.077)
Value-added of closest district school	0.041	0.001	0.044	-0.008	0.078	0.055
	(0.015)	(0.042)	(0.015)	(0.041)	(0.055)	(0.155)
Fourth grade math score	0.143	-0.034	-0.125	-0.051	0.437	-0.237
	(0.003)	(0.010)	(0.003)	(0.010)	(0.012)	(0.037)
Fourth grade reading score	0.019	-0.005	0.022	-0.005	0.081	-0.047
	(0.003)	(0.010)	(0.003)	(0.010)	(0.012)	(0.038)
Charter school preference, θ_i	0.018	-0.028	0.017	-0.027	0.052	-0.089
	(0.005)	(0.014)	(0.005)	(0.014)	(0.018)	(0.053)
Idiosyncratic preference, τ_{ij}	-	-0.005	-	-0.003	-	0.008
		(0.016)		(0.016)		(0.059)
P-value: No selection on unobservables	0.001		0.001		0.001	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade math scores for several transformations of test scores. The control functions are posterior means from the two-mass mixture model in column (3) of Table 4. Columns (1) and (2) report results for test scores measured as percentile ranks, columns (3) and (4) show results using the change in percentile rank between fourth and eighth grade, and columns (5) and (6) display results using the log of the percentile rank. Standard errors are adjusted for estimation of the control functions.

Table A9: Selection-corrected estimates for eighth grade test scores with alternative control functions

	Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)
Constant/main effect	-0.458	0.838	-0.512	0.5328
	(0.010)	(0.135)	(0.010)	(0.141)
Female	-0.024	0.061	0.184	-0.018
	(0.015)	(0.046)	(0.016)	(0.048)
Black	-0.193	0.250	-0.087	0.199
	(0.025)	(0.073)	(0.026)	(0.077)
Hispanic	-0.100	0.260	-0.041	0.243
	(0.025)	(0.078)	(0.027)	(0.081)
Subsidized lunch	-0.128	0.194	-0.126	0.151
	(0.022)	(0.056)	(0.023)	(0.059)
Special education	-0.370	0.098	-0.397	0.135
	(0.020)	(0.065)	(0.021)	(0.068)
Limited English proficiency	0.075	-0.090	0.044	-0.073
	(0.020)	(0.069)	(0.021)	(0.072)
Value-added of closest district school	0.136	0.005	0.113	-0.039
	(0.049)	(0.138)	(0.051)	(0.145)
Fourth grade math score	0.476	-0.122	0.165	-0.044
	(0.011)	(0.033)	(0.011)	(0.035)
Fourth grade reading score	0.066	-0.019	0.366	-0.078
	(0.011)	(0.034)	(0.011)	(0.036)
Type one (high θ_i)	0.164	-0.292	0.145	-0.114
	(0.046)	(0.134)	(0.048)	(0.140)
Idiosyncratic preference, τ_{ij}	-	-0.019	-	0.008
		(0.052)		(0.055)
P-value: No selection on unobservables	0.001		0.044	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade test scores using the posterior type probability from the two-mass mixture model as a control function. Standard errors are adjusted for estimation of the control functions.