A Structural Model of Charter School Choice and Academic Achievement

Christopher Walters*

November 6th, 2012

Abstract

Lottery-based instrumental variables estimates show that Boston's charter schools substantially increase test scores and close racial achievement gaps among their applicants. A key policy question is whether charter expansion is likely to produce similar effects on a larger scale. This paper uses a structural model of school choice and academic achievement to predict the effects of charter expansion for the citywide achievement distribution in Boston. Estimates of the model suggest that charter applicants are negatively selected on achievement gains: low-income students and students with low prior achievement gain the most from charter attendance, but are unlikely to apply to charter schools. This form of selection implies that lottery-based estimates understate gains for broader groups of students, and that charter schools will produce substantial gains for marginal applicants drawn in by expansion. Simulations suggest that realistic expansions are likely to reduce the gap in math scores between Boston and the rest of Massachusetts by up to 8 percent, and reduce racial achievement gaps by roughly 5 percent. Nevertheless, the estimates also imply that perceived application costs are high and that most students prefer traditional public schools to charter schools, so large expansions may leave many charter seats empty. These results suggest that in the absence of significant behavioral or institutional changes, the potential gains from charter expansion may be limited as much by demand as by supply.

*MIT Economics (e-mail: crwalt@mit.edu). I am grateful to Joshua Angrist, Parag Pathak, and David Autor for their guidance and support. I also thank Aviva Aron-Dine, David Chan, Sarah Cohodes, Susan Dynarski, Christopher Palmer, Stephen Ryan, Xiao Yu Wang, Tyler Williams, and participants at the MIT Labor and Industrial Organization field lunches for their useful suggestions. Special thanks go to Carrie Conaway and the Massachusetts Department of Elementary and Secondary Education for suggestions, assistance and data. Danielle Wedde provided valuable assistance by collecting lottery data and survey responses. 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.

1 Introduction

Differences in test scores between racial and socioeconomic groups are a pervasive feature of the American educational landscape. In 2008, 13-year-old black students scored more than 0.8 standard deviations (σ) below their white counterparts on the National Assessment of Educational Progress Long Term Trend (NAEP-LTT) math test. Similar achievement gaps appear in all subjects and at all grade levels (Fryer and Levitt 2004, 2006; Vanneman et al., 2009). Moreover, after several decades of convergence, the relative scores of black students stagnated in the late 1980s (Neal 1996). Achievement gaps between high- and low-income students have also grown in recent years (Reardon 2011). Numerous policy interventions with the potential to affect achievement gaps have been proposed, but few produce gains of the magnitude necessary to substantially reduce these gaps (Fryer 2010).[1] As a result, some analysts have argued that it is either impossible or inordinately expensive to significantly reduce achievement gaps using educational policies alone, especially policies that target adolescents (Rothstein 2004; Heckman 2011).

These pessimistic accounts notwithstanding, a growing body of evidence suggests that some charter schools serving poor, minority populations in urban areas boost achievement sharply. Charter schools are publicly funded, non-selective schools that typically have more freedom than traditional public schools to set curricula and make staffing decisions. Studies based on entrance lotteries show that attendance at charter schools in Boston and New York's Harlem Children's Zone raises achievement by 0.25σ per year or more (Abdulkadiroglu et al. 2011; Dobbie and Fryer 2011a). Angrist et al. (2011, 2012), Dobbie and Fryer (2011b), Gleason et al. (2010), Hoxby and Murarka (2009), and Hoxby and Rockoff (2004) also report positive effects for urban charters. [2]

These findings suggest that urban charter schools may have the potential to reduce achievement gaps. Reflecting this hope, the Massachusetts legislature recently relaxed the state's charter school cap with the explicit goal of reducing racial and socioeconomic disparities in academic performance (Commonwealth of Massachusetts 2010). Twenty Massachusetts charter schools are approved to open by Fall 2014, including 11 in Boston (Massachusetts Department of Elementary and Secondary Education 2012b). State officials have announced plans to approve

[1] A few interventions have larger effects for blacks than for whites. Krueger and Whitmore (2001) use data from the Tennessee STAR experiment to show that a 50% reduction in class size over three years raises the scores of blacks and whites by 0.26σ and 0.13σ, respectively. Similarly, Howell and Peterson (2002) estimate that providing vouchers for private school boosts scores by 0.2σ for black students and has little benefit for whites. Barnett (1992) and Ferguson (1998) report that the Perry Preschool Project, an intensive early childhood intervention for at-risk black students, increased achievement in adolescence by roughly 0.4σ.

[2] Estimates for charter schools outside urban areas are mixed. Gleason et al. (2010) find that non-urban charters are no more effective than traditional public schools, while Angrist et al. (2011) find negative effects for non-urban charter middle schools in Massachusetts.

additional Boston charter schools in the near future (Vaznis 2012).

Despite the substantial effects reported in lottery-based studies, however, some analysts argue that charter expansion is unlikely to raise achievement on a larger scale. Charter applicants are a small, self-selected subset of the student population; in Boston, only 17 percent of students apply to a charter, and applicants are less disadvantaged and higher-achieving than non-applicants.[3] Students with the largest potential benefits may be most likely to sign up for charter lotteries. In support of this view, Ravitch (2010, pp. 144-145) argues that "charter schools enroll the most motivated students in poor communities, those whose parents push them to do better...as more charter schools open, the dilemma of educating all students will grow sharper." Similarly, Rothstein (2004, p. 82) writes of the Knowledge is Power Program (KIPP), a high-performing urban charter operator: "[T]hese exemplary schools...select from the top of the ability distribution those lower-class children with innate intelligence, well-motivated parents, or their own personal drives, and give these children educations they can use to succeed in life." In this view, lottery-based estimates that identify effects for charter applicants overstate the effects of charter attendance for broader groups of students, and charter expansion is unlikely to boost citywide achievement.

On the other hand, lottery-based estimates could also understate the potential effects of charter attendance for non-applicants. The parents of low-achieving students may be unlikely to investigate alternatives to traditional public school, despite evidence that urban charters are especially effective for such students (Angrist et al. 2012). More generally, since applying to charter schools requires parental action, the parents of charter applicants may be more motivated to invest in their children's human capital than other parents. The inputs provided by charter schools may either complement or substitute for these parental investments. The efficacy of charter expansion is therefore theoretically unclear.

The aim of this paper is to predict the consequences of charter school expansion in Boston, with an emphasis on achievement gaps. To this end, I develop and estimate a structural model that links students' charter application decisions to their potential achievement gains in a parametric selection framework. The model is similar to the stochastic portfolio choice problem outlined by Chade and Smith (2006): students submit charter applications to maximize expected utility, taking account of admission probabilities and non-monetary application costs. To identify the model's parameters, I combine instruments from entrance lotteries with a second set of instruments based on proximity to charter schools. The approach taken here is similar in spirit to other recent studies that use economic theory to extrapolate from experimental or

[3]The demographics of charter applicants are discussed in detail in Section 2.3.

quasi-experimental causal estimates (see, e.g., Todd and Wolpin 2010, Card and Hyslop 2006, Attanasio et al. 2011, Duflo et al. 2012, and Hastings et al. 2009).[4] The model accommodates heterogeneity in student preferences and achievement gains on both observed and unobserved dimensions. In addition, it allows for heterogeneous effects across charter schools, and permits charter admission probabilities to adjust endogenously to changes in the set of available schools.

Estimates of the model imply that charter applicants are *negatively* selected on gains from charter attendance. Specifically, higher-achieving, less-disadvantaged students have the strongest preferences for charter schools, but charters are most effective for poor students and those with low previous achievement. The structural estimates also imply that charter applicants are negatively selected on unobserved dimensions of achievement gains. Surprisingly, these findings imply that lottery-based estimates *understate* the achievement effects of charter schools for broader sets of students. As a result, charter expansion has the potential to substantially raise achievement for marginal applicants: simulations show that Boston's proposed expansion, which raises the share of middle schoolers attending charters from 9 percent to 15 percent, is expected to reduce the gap in math scores between Boston and the rest of Massachusetts by 5 to 8 percent, and reduce citywide racial achievement gaps by roughly 5 percent.

At the same time, estimates of the model also show that the effects of further expansions may be limited by weak demand. Students act as if application costs are high, and most prefer to attend traditional public schools despite the achievement gains they would receive from charters. As a result, large expansions of the charter system are likely to leave many charter seats empty without substantially increasing enrollment. In the long run, demand for charters may rise if parents become more informed or measures are taken to boost application rates. Taken together, however, the results reported here suggest that at present, parents in Boston are either unaware of, or unresponsive to, the achievement gains produced by charter schools: The overall level of demand for charters is low, and the students with the largest potential gains are relatively unlikely to apply.

These results contribute to a nascent literature assessing the possible consequences of charter school expansion. Fryer (2011) examines the effects of introducting practices from successful charter schools in nine low-performing traditional public schools in Houston, Texas. The results show clear gains in math, suggesting that these practices are effective outside the charter context. Other related studies include Mehta (2011) and Ferreyra and Kosenok (2011), who develop

[4] Angrist and Fernandez-Val (2011) and Hotz et al. (2005) describe approaches to extrapolation that emphasize variation in treatment effects as a function of observed covariates. Heckman and Vytlacil (2001, 2005), Carneiro et al. (2010), and Heckman (2010) discuss nonparametric estimation of marginal treatment effects (MTE), effects at particular values of the unobserved propensity to receive treatment. The approach taken here includes elements of both of these approaches.

equilibrium models of charter school entry and use them to conduct counterfactual analyses. Strategic entry models are less appropriate for the Boston setting because entry is effectively determined by Massachusetts' charter school cap legislation. The model estimated here focuses on the demand for charter schools, with an emphasis on the relationship between preferences and the achievement effects of charter attendance.

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 benchmarks the effects of charter schools in the sample of lottery applicants. Section 4 outlines the structural model, and Section 5 discusses identification and estimation of the model. Section 6 reports the structural estimates. Section 7 uses these estimates to simulate the effects of charter expansion. Section 8 concludes.

2 Setting and Data

2.1 Context: Charter Schools in Boston

Non-profit organizations, teachers, or other groups wishing to operate charter schools in Massachusetts submit applications to the state's Board of Education. If authorized, charter schools are granted freedom to organize instruction around a philosophy or curricular theme, as well as budgetary autonomy. Charter employees are also typically exempt from local collective bargaining agreements, giving charters more discretion over staffing than traditional public schools.[5] The Board of Education reviews each charter school's academic and organizational performance at five year intervals, and decides whether charters should be renewed or revoked. Enrollment at Massachusetts charter schools is open to all students who live in the local school district; if a charter school receives more applications than it has seats, it must accept students by random lottery. 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).[6]

The Boston Public Schools (BPS) district is the largest school district in Massachusetts, and it also enrolls an unusually large share of charter students. In the 2010-2011 school year, 14

[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 middle schools operated in Boston during my data window, though three were scheduled to open in 2010 or later.

[6] In the first year after an increase in charter enrollment, the state fully reimburses the local district for the additional charter tuition payments. Subsequent reimbursement rates are 60 percent in the second year, 40 percent in the third year, and zero thereafter.

charter schools operated in Boston, accounting for 9 percent of BPS enrollment. The analysis here focuses on middle schools, defined as schools that accept students in fifth or sixth grade; 12 percent of Boston middle schoolers attended charter schools in 2010-2011. Columns (1) through (3) of Table 1 list names, grade structures and years of operation for the ten Boston charter middle schools that opened between 1997 and 2010. These schools are marked in black on the map in Figure 1.[7]

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

These encouraging findings make Boston an appealing setting for studying the effects of charter school expansion. The effects of expansion in Boston are also relevant to an ongoing policy debate. In recent years, the growth of charters in Massachusetts has been slowed by the state's charter cap, a law that limits expenditures on charter schools to 9 percent of the host district total.[8] The Board of Education stopped accepting proposals for new Boston charters in 2008 when charter expenditure hit the cap (Boston Municipal Research Bureau 2008).

In 2010, the Massachusetts legislature relaxed the charter cap for low-performing school districts. Specifically, for districts with test scores in the lowest decile, the limit on charter expenditures is to rise incrementally from 9 percent in 2010 to 18 percent in 2017 (Commonwealth of Massachusetts 2010). The new law gives priority to "proven providers" who have previously held leadership positions at schools demonstrating academic success for similar student populations (Massachusetts Department of Elementary and Secondary Education 2012a). The law also requires schools to specify recruitment plans aimed at attracting applicants who are demographically similar to the local population, though all students are free to apply and admissions will continue to be determined by lottery (Massachusetts Department of Elementary and Secondary Education 2012c).

Through 2011, the Board of Education received 51 charter applications under the new law. Of these, 32 were selected as finalists, and 20 charters were granted, eleven to schools in Boston

[7] Only nine locations are marked in Figure 1 because Boston Preparatory Charter School opened in the building previously occupied by Frederick Douglass Charter School after the latter closed.

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

(Massachusetts Department of Elementary and Secondary Education 2012). Table 2 lists the Boston charter middle schools scheduled to open in the 2011-2012 and 2012-2013 school years, as well as existing schools operated by the same providers. Two opened in 2011-2012, while four opened in 2012-2013. Five new schools are linked to existing charters in Boston; the sixth, KIPP Academy Boston, is part of the Knowledge is Power Program, the nation's largest charter management organization.[9] The locations of the newly approved schools are marked in red in Figure 1. The state Board of Education has announced that it will consider proposals for additional Boston charter schools opening in 2013 and later (Vaznis 2012).

2.2 Data Sources and Sample Construction

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

The DESE database covers all Massachusetts public school students from the 2001-2002 school year through the 2010-2011 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 ELA achievement tests. I begin by selecting all white, black, and Hispanic students in the database who attended a traditional BPS school in 4th grade between 2006 and 2009. I also require students to have non-missing 4th grade demographics and test scores, as well as school attendance information and test scores in 6th, 7th, or 8th grade. I use only the earliest test taken by a given student in a particular subject and grade.

Next, I merge the student address database to the DESE administrative file using a crosswalk between BPS and state student identifiers. The address database includes a record for every year that a student attended a traditional BPS school between 1998 and 2011. I drop students in the state database without BPS address data. This restriction eliminates less than 1 percent of Boston 4th graders. The address information is used to measure proximity to each Boston charter school. I measure proximity using road distance in miles.[10]

[9]KIPP operates two charter schools in Lynn, a poor suburb of Boston. In a lottery-based evaluation of one of these schools, Angrist et al. (2012) estimate effects similar to those of Boston's charter middle schools.

[10]Road distance is obtained using the STATA *traveltime* command, which queries Google Maps to obtain travel distance between any two locations. Versions of the model using other measures of proximity, such as direct "as the crow flies" distance or driving time, produced very similar results.

I then match the student data to records from lotteries held at seven charter middle schools in Boston.[11] I focus on middle schools because applicant records were more consistently available in middle school than in elementary or high school, and I restrict attention to applicant cohorts attending 4th grade in 2006 and later because records for earlier cohorts were missing for several schools. Column (4) of Table 1 summarizes the availability of lottery records for the ten charter middle schools that operated between 1997 and 2010.[12] Of the three schools without available records, two closed prior to the 2009-2010 school year; the third declined to provide records. In the analysis below, I treat these schools as equivalent to traditional public school. I matched the available lottery records to the administrative data by name, grade, year, and (where available) date of birth. This process produced unique matches for 92 percent of lottery applicants.

After matching the lottery files to the student data, I constructed two subsamples for statistical analysis. The first is used to estimate causal effects for lottery applicants. This sample excludes students that did not apply to charter schools, as well as applicants who were not randomized. The latter group includes siblings guaranteed admission, and late applicants, who are typically placed on a waiting list. The lottery sample includes 1,822 applicants to charter middle schools. A second sample is used to estimate the structural model. In addition to randomized applicants, this sample includes students who did not apply to charter schools and applicants who were not randomized. The structural sample includes 10,986 students who attended BPS schools in 4th grade between 2006 and 2009.

2.3 Descriptive Statistics

Applicants to Boston charter schools differ from the general population of Boston students. Specifically, charter applicants tend to have higher socioeconomic status, and to enter middle school with higher prior achievement than non-applicants. This can be seen in Table 3, which reports summary statistics for the structural sample in column (1) and the applicant sample in column (2). Seventeen percent of students applied to at least one charter lottery, and ten percent attended a charter school. Compared to the general student population, charter applicants are less likely to be Hispanic, to be eligible for subsidized lunch, to have special education status, or to be classified as limited English proficient. Charter applicants are more likely to be black, and also live closer to charter schools on average (1.64 miles from the closest charter school, compared to 1.84 miles for the full population).

[11]The lotteries used here are an expanded version of the middle school sample used by Abdulkadiroglu et al. (2011), including two additional schools.

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

The last two rows of Table 3 display information about 4th grade MCAS scores. I standardize MCAS scores to have mean zero and standard deviation one within each grade-year in Massachusetts. Boston 4th graders lag behind the state average by 0.54σ and 0.66σ in math and ELA. Students who apply to charter schools have substantially higher scores than the general Boston population: applicants' 4th grade scores exceed the Boston average by more than 0.2σ in both subjects. Taken together, these summary statistics indicate that charter applicants are higher-achieving, less economically disadvantaged, and less likely to have academic problems than students who do not apply to charter schools.

3 Effects on Lottery Applicants

3.1 Lottery Estimates

To motivate my analysis of the effects of charter expansion, I begin by benchmarking the effects of charter attendance among the selected subset of students who apply to charter schools. I interpret these effects in the Local Average Treatment Effect (LATE) notation described by Imbens and Angrist (1994), which provides a formal framework for analyzing heterogeneity in causal effects across individuals. Let $Y_i(1)$ be applicant i's potential test score if she attends a charter school, and let $Y_i(0)$ be her test score if she attends a public school. S_i indicates charter attendance (the "treatment"), and Z_i is a lottery offer dummy. Let $S_i(1)$ and $S_i(0)$ denote potential treatment status as a function of Z_i. The LATE framework is based on the following assumptions for the lottery applicant sample:

A1 *Independence and Exclusion*: $(Y_i(1), Y_i(0), S_i(1), S_i(0))$ is independent of Z_i.

A2 *First Stage*: $0 < Pr[Z_i = 1] < 1$ and $Pr[S_i(1) = 1] > Pr[S_i(0) = 1]$.

A3 *Monotonicity*: $S_i(1) \geq S_i(0)\ \forall i$.

The Independence and Exclusion assumption is motivated by the observation that offers are randomly assigned among applicants, and are unlikely to affect test scores through any channel but charter attendance. The First Stage assumption requires that winning the lottery makes applicants more likely to attend charter school on average. Monotonicity requires that winning the lottery does not discourage any applicant from attending charter school.

Under assumptions A1-A3, applicants can be partitioned into three groups: never takers, who never attend charters ($S_i(1) = S_i(0) = 0$), always takers, who attend regardless of the offer ($S_i(1) = S_i(0) = 1$), and compliers, who are induced to attend by receiving offers ($S_i(1) >$

$S_i(0)$). Imbens and Angrist (1994) show that conventional instrumental variables (IV) methods consistently estimate LATE, the average treatment effect for compliers. We have

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[S_i|Z_i = 1] - E[Y_i|S_i = 0]} = E[Y_i(1) - Y_i(0)|S_i(1) > S_i(0)] \tag{1}$$

The Wald (1940) IV estimator is the empirical analogue of the left-hand side of equation (1).

I estimate LATE using a two-stage least squares (2SLS) procedure that combines observations from multiple lotteries. Specifically, the estimating equation for the lottery analysis is

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

where Y_i is a test score for applicant i, S_i is a dummy variable indicating charter school attendance, and ψ_ℓ is a set of fixed effects capturing all combinations of charter lotteries entered by students in the data. I code a student as attending charter school if she attends a charter at any time after the lottery and prior to the test. The first stage equation is

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

The instrument Z_i is one for students who receive any charter offer before the start of the school year following the lottery.[13] The 2SLS estimate of β can be interpreted as a weighted average of within-lottery LATEs.[14] To use all available test score information, the sample stacks test scores in grades six through eight. Standard errors are clustered at the student level.

Consistent with the results reported by Abdulkadiroglu et al. (2011), the 2SLS estimates show that Boston's charter schools have dramatic effects on student achievement for lottery applicants. As shown in column (1) of Table 4, receipt of a lottery offer increases the probability of charter attendance by 0.65. The second-stage estimates, reported in columns (2) and (3), imply that attending a charter school increases math scores by 0.50σ and boosts ELA scores by 0.31σ. These effects are precisely estimated.

These pooled results mask substantial heterogeneity in the benefits of charter school attendance across racial groups. The second row of Table 4 shows that the effects of charter schools are relatively modest for white students: the math estimate for whites is a statistically insignificant 0.14σ, and the ELA estimate is negative and insignificant. In contrast, the third

[13] With random assignment of Z_i, pre-lottery characteristics should be balanced across winners and losers. Appendix Table A1 examines balance for observable student characteristics. There are few significant differences between lottery winners and losers, and joint tests of balance never reject at conventional significance levels. Even with random assignment, the validity of the instrument can be compromised by non-random attrition. Appendix Table A2 shows that the followup rate for the lottery sample is 84 percent, and followup rates for lottery winners and losers are very similar. Column (1) of Appendix Table A2 shows that followup rates are similar in the lottery and structural samples.

[14] The 2SLS estimate is a weighted average of within-lottery Wald estimates, with weights proportional to the variance of the first-stage fitted values (Angrist and Imbens 1995).

and fourth rows reveal large, significant effects for black and Hispanic students in both subjects. Charter attendance boosts scores for black students by 0.62σ in math and 0.38σ in ELA. The corresponding effects for Hispanics are 0.57σ and 0.53σ. The last row of Table 4 reports p-values from Wald tests of the equality of charter effects across races. The null hypothesis of equal effects is rejected at conventional significance levels ($p < 0.1$) for both subjects. These results show that Boston's charter schools raise test scores for non-white students much more than for whites.

3.2 Effects on Score Distributions by Race

As a final piece of motivation for the structural analysis to follow, I next use the lottery sample to ask whether charter schools close racial achievement gaps among applicant compliers. To estimate effects on black and white score distributions, I modify the methods described by Abadie (2002, 2003). Abadie notes that in addition to LATE, the marginal distributions of $Y_i(1)$ and $Y_i(0)$ are separately identified for compliers in instrumental variables settings. Intuitively, the distribution of Y_i for students with $S_i = Z_i = 0$ is a mixture of the distributions of $Y_i(0)$ for compliers and never takers. The distribution of $Y_i(0)$ for never takers is directly observable among students with $Z_i = 1$ and $S_i = 0$. The distribution of $Y_i(0)$ for compliers can therefore be recovered by a deconvolution procedure that uses these two observed distributions. A similar argument shows that the distribution of $Y_i(1)$ for compliers can be recovered using the distribution of Y_i for students with $S_i = Z_i = 1$ together with the distribution for students with $S_i = 1$ and $Z_i = 0$. Abadie provides simple methods for estimating CDFs of potential outcome distributions for compliers, and outlines bootstrap procedures for testing hypotheses about these distributions. I extend these methods to estimate potential outcome densities separately by race, and test for black-white equality among applicant compliers who are randomly assigned to charter schools or public schools.

The estimating equations for the distributional analysis are of the form

$$K_h\left(y - Y_i\right) \cdot S_i = \kappa_{\ell y} + \gamma(y) \cdot S_i + \eta_{iy}, \tag{4}$$

where S_i is treated as an endogenous regressor and instrumented with lottery offers. Here $K_h(t) = \frac{1}{h}K\left(\frac{t}{h}\right)$, $K(t)$ is a kernel function, and h is a bandwidth. Let $f_s^c(y)$ be the density of $Y_i(s)$ for lottery compliers. Appendix A shows that the probability limit of the 2SLS estimate of $\gamma(y)$ is $f_1^c(y)$, the density function for treated compliers. $f_0^c(y)$ can be estimated by replacing S_i with $(1 - S_i)$ in equation (4). I use a Kolmogorov-Smirnov (KS) statistic to test distributional equality for blacks and whites.[15]

[15]The KS statistic is proportional to the maximum difference in complier CDFs between racial groups. I

Boston's charter schools cause math score distributions for black and white applicant compliers to converge. This can be seen in Figure 2, which plots the estimated complier densities of $Y_i(0)$ and $Y_i(1)$ for math, separately by race and year since application. The densities are estimated using a triangle kernel with bandwidth 1σ.[16] Black vertical lines indicate the pooled mean of $Y_i(0)$ in each figure, while red lines mark the mean of $Y_i(1)$ in plots for treated students. At baseline (prior to treatment), distributions for treated and non-treated compliers are similar, and black students lag behind whites throughout the distribution. The KS test rejects baseline distributional equality at conventional significance levels ($p = 0.01$ for the untreated, $p = 0.09$ for the treated). In post-treatment years, the black distribution moves towards the white distribution for treated compliers, and in 7th and 8th grade the null hypothesis of distributional equality cannot be rejected for treated students ($p \geq 0.54$).

In contrast with the results for treated compliers, no convergence occurs for untreated compliers. The left-hand panels of Figure 2 show that black compliers who attend public schools lag behind their white counterparts in every year, with little relative change in the distributions after baseline. The null hypothesis of distributional equality between untreated black and white compliers is always rejected at the 10-percent level or lower. These results suggest that Boston's charter schools close otherwise persistent achievement gaps between black and white compliers in math. As shown in Figure 3, black-white convergence is less pronounced for ELA than for math, though large shifts in mean ELA scores are evident in the plots for 7th and 8th grade.

4 Modeling Charter School Attendance

4.1 Setup

The lottery estimates in Section 3 show that Boston's charter schools have dramatic effects on average test scores and racial achievement gaps for applicant compliers. At the same time, effects for non-applicants may differ systematically from those for applicants, so these results need not provide an accurate guide to the likely consequences of charter expansion. To extrapolate from the lottery-based estimates, I use a structural model of charter application, attendance, and achievement. As in Chade and Smith (2006) and Ajayi (2011), the charter

estimate these CDFs by replacing $K_h\left(y - Y_i\right)$ with $1\{Y_i \leq y\}$ in equation (4). Inference is based on a stratified bootstrap procedure. For each of 200 bootstrap replications, I draw observations with replacement within lotteries to obtain a new sample with the same lottery-specific sample sizes as the original sample. I then randomly assign observations in each lottery to racial groups in the same proportions as in the original sample, and recalculate the KS statistic. The results provide the sampling distribution of the KS statistic under the null hypothesis of distributional equality for black and white compliers.

[16] Imbens and Rubin (1997) point out that instrumental variables estimates of potential outcome densities are not guaranteed to be positive. I follow their suggestion and set the estimated densities to zero in a small number of cases where the 2SLS estimate is negative.

school application decision is modeled as a random utility portfolio choice problem: students choose a set of applications to maximize expected utility, taking into account admission probabilities and application costs. The model also allows for heterogeneous effects across charter schools.

Figure 4 explains the sequence of events described by the model. First, students decide whether to apply to each of K charter schools, indexed by $k \in \{1...K\}$. The dummy variable $A_{ik} \in \{0, 1\}$ indicates that student i applies to school k. Second, charter schools randomize offers to applicants. The dummy variable $Z_{ik} \in \{0, 1\}$ indicates an offer for student i at school k, and π_k denotes the admission probability for applicants to school k. In the third stage, students choose schools denoted S_i, where $S_i = 0$ indicates public school attendance. Any student can attend public school, but student i can attend charter school k only if $Z_{ik} = 1$. Finally, students take achievement tests in grades six through eight. Y_{ijg} denotes student i's score in grade g on math and ELA tests, indexed by $j \in \{m, e\}$.

4.2 Student Choice Problem

4.2.1 Preferences

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

$$U_{ik} = \gamma_k^0 + X_i'\gamma^x + \gamma^d \cdot D_{ik} + \theta_i + v_{ik} - \gamma^a \cdot |A_i| \tag{5}$$

where X_i is a vector of characteristics for student i including sex, race, subsidized lunch status, special education status, limited English proficiency, and 4th grade math and ELA scores. D_{ik} measures distance to school k. The utility of public school attendance is

$$U_{i0} = v_{i0} - \gamma^a \cdot |A_i|.$$

The quantity $\gamma^a \cdot |A_i|$ represents the utility cost of applying to $|A_i|$ charter schools.[17] Application costs include the disutility of filling out application forms and the opportunity cost of time spent attending lotteries. These costs may also capture frictions associated with learning about charter schools. Charter schools are not included in the standard BPS school choice system, and are not typically listed in informational resources provided to parents by the district.[18] Parents who wish to learn about charter schools must undertake additional efforts, such

[17]Variables without k subscripts refer to vectors, so that $A_i \equiv (A_{i1}...A_{iK})'$ and so on.

[18]For example, the "What Are My Schools?" tool located at https://externalweb.mybps.org provides a list of the BPS schools to which children are eligible to apply, but does not list charter schools.

as visiting individual school websites or attending information nights. Applicants pay these costs whether or not they attend a charter.

The variables θ_i and v_{ik} represent unobserved heterogeneity in tastes. θ_i, which characterizes student i's preference for charter schools relative to traditional public school, is the key unobservable governing selection into the charter sector. This variable includes any latent factors that influence students to opt out of traditional public school in favor of charter schools, such as the perceived achievement gain from attending charter schools, proximity to the relevant public school, and parental motivation.[19] In the language of the random-coefficients logit model (see, e.g., Hausman and Wise 1978, Berry et al. 1995, and Nevo 2000), θ_i is the random coefficient on a charter school indicator. The presence of θ_i implies that charter schools are closer substitutes for each other than for traditional public schools. I assume that θ_i follows a normal distribution with mean zero and variance σ_θ^2.

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

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

Students know τ_{ik} and θ_i before applying to charter schools, and learn ξ_{ik} after applying. The post-application preference shock explains why some applicants decline charter school offers. To generate multinomial logit choice probabilities, τ_{ik} and ξ_{ik} are assumed to follow independent extreme value type I distributions with variances σ_τ^2 and $\pi^2/6$.[20]

4.2.2 School Lotteries

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

$$f(Z_i|A_i;\pi) = \prod_k \left[A_{ik} \cdot (\pi_k Z_{ik} + (1-\pi_k)(1-Z_{ik})) + (1-A_{ik}) \cdot (1-Z_{ik})\right]. \quad (6)$$

Initially, admission probabilities are treated as parameters to be estimated. However, admission rates are likely to change as the system of charter schools expands. In the simulations to follow, the π_k adjust so that schools fill their seats in equilibrium. Section 7.2 and Appendix D discuss the determination of endogenous admission probabilities.

[19] Proximity to public school is treated as an unobservable because Boston has a citywide choice plan, so students have a large number of traditional public schools to choose from.

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

4.2.3 Application and Attendance Decisions

I derive students' optimal application and attendance rules by backward induction. A student is faced with a unique attendance decision after each possible combination of charter school offers, because the set of offers in hand determines the available school choices. Consider the decision facing a student at stage 3 in Figure 4. At this point, the student knows her charter offers, application costs are sunk, and there is no uncertainty about preferences. Student i can attend public school or any charter school that offers a seat. Her choice set is

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

Define

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

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

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

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

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

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

$$\begin{aligned} W_i(Z_i, \theta_i, \tau_i) &\equiv E[\max_{k \in \mathcal{C}(Z_i)} \tilde{U}_{ik}(\theta_i, \tau_{i0}) + \xi_{ik} | X_i, D_i, Z_i, \theta_i, \tau_i] \\ &= \nu + \log\left(\sum_{k \in \mathcal{C}(Z_i)} \exp\left(\tilde{U}_{ik}(\theta_i, \tau_{ik})\right)\right) \end{aligned}$$

where ν is Euler's constant.

Students choose charter applications to maximize expected utility, anticipating offer probabilities and their own attendance choices. Consider the application decision facing a student at stage 1 in Figure 4. The student knows θ_i and τ_i, but does not know ξ_i, and her choice of A_i induces a lottery over values of Z_i at a cost of $\gamma^a \cdot |A_i|$. The expected utility from choosing $A_i = a$ is given by

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

The optimal application rule is therefore

$$\begin{aligned} A_i &= \arg\max_{a \in \{0,1\}^K} V_i(a, \theta_i, \tau_i) \\ &\equiv A^*(X_i, D_i, \theta_i, \tau_i). \end{aligned}$$

4.3 Achievement

Students are tested after application and attendance decisions have been made. *Potential* achievement for student i at charter school k in subject j and grade g is given by

$$Y_{ijg}(k) = \alpha^0_{jk} + \alpha^0_{jgc} + X_i'\alpha^x_{jgc} + \alpha^\theta_{jgc} \cdot \theta_i + \epsilon_{ijgk} \tag{7}$$

while potential traditional public school achievement is

$$Y_{ijg}(0) = \alpha^0_{jgp} + X_i'\alpha^x_{jgp} + \alpha^\theta_{jgp} \cdot \theta_i + \epsilon_{ijg0}. \tag{8}$$

The subscripts c and p in these equations refer to charter school and public school. The causal effect of attending charter school k relative to traditional public school for student i in subject j and grade g is $Y_{ijg}(k) - Y_{ijg}(0)$. Observed scores for student i are the potential scores associated with her optimal school choice: $Y_{ijg} = Y_{ijg}(S_i)$.

The unobserved determinants of academic achievement may be correlated both over time and across subjects. To capture this possibility, I allow ϵ_{ijgk} to follow the first-order autoregressive process

$$\epsilon_{ijgk} = \rho_{jk} \cdot \epsilon_{ij(g-1)k} + \zeta_{ijgk}, \tag{9}$$

where the ζ_{ijgk} are serially independent and the vector $(\zeta_{imgk}, \zeta_{iegk})'$ has a bivariate normal distribution with covariance matrix Σ_k. I assume that ρ_{jk} and Σ_k are the same across charter schools, though they may differ between charter schools and traditional public schools.

4.4 Comments on Modeling Choices

Equations (5) through (9) provide a complete description of charter demand and potential academic achievement. This section provides intuition for some of the key modeling choices implicit in these equations.

First, the model emphasizes differences between charter and traditional public schools, while limiting differences between charter schools. Heterogeneity in preferences and achievement across students with different observed characteristics is governed by the vectors γ^x and $\alpha^x_{jg\ell}$ for $\ell \in \{c, p\}$. This specification allows observed characteristics to affect the choice of charter schools relative to traditional public schools, and to interact differently with achievement in charter and public schools, but requires that these characteristics affect preferences and achievement the same way at every charter. Similarly, equation (7) implies that the relationship between the unobserved taste θ_i and student achievement is the same at every charter school. Heterogeneity in preferences and achievement across charter schools is captured by the school-specific intercepts γ^0_k and α^0_{jk}. These restrictions allow me to limit the number of parameters to be estimated while also parsimoniously summarizing heterogeneity in preferences and achievement gains across both students and schools. Moreover, this emphasis on differences between charters and traditional public schools mirrors the approach to identification described in Section 5.3, which emphasizes selection into the charter sector rather than across charter schools.

A second notable feature of the model is that potential achievement does not enter directly in students' utility functions. Instead, achievement and preferences are linked through the charter taste θ_i, which appears in equations (7) and (8). Appendix C formally shows that this specification nests a standard model of Roy (1951) selection in which students seek to maximize achievement and have private information about their potential scores in charter and public schools. The model described here is more flexible than this Roy model, in the sense that it allows students' preferences to depend on unobserved factors besides achievement. For example, students may be more likely to choose charter schools if they expect to receive large achievement gains, but they may also be more likely to choose charters if they have more motivated parents, and parental motivation may be positively or negatively correlated with achievement gains. Equations (7) and (8) admit either possibility by allowing for a flexible relationship between the unobserved charter taste θ_i and potential achievement. The next section formally discusses my strategy for identifying this relationship and outlines my estimation procedure.

5 Identification and Estimation

5.1 Exclusion Restriction

The central challenge in extrapolating from lottery-based estimates of charter effects is that lottery applicants are self-selected. This leads potential achievement distributions for applicants to differ from the corresponding population distributions. The model described here accounts for this self-selection while placing structure on the selection process for the purposes of identification. Specifically, identification of the parameters of equations (7) and (8) is based on the following exclusion restriction:

$$E[\zeta_{ijgk}|X_i, \theta_i, Z_i, D_i, \tau_i, \xi_i] = 0. \tag{10}$$

Equation (10) three identifying assumptions. First, the lottery offer vector Z_i is excluded from equations (7) and (8). The offer choice A_i is a deterministic function of X_i, D_i, θ_i, and τ_i, and lottery offers are randomly assigned conditional on A_i. The exclusion of Z_i is therefore equivalent to assuming that offers have no direct effect on student achievement, a standard assumption in the charter lottery literature. Second, assumption (10) requires that distance to charter schools is unrelated to the idiosyncratic component of potential achievement, ζ_{ijgk}.

Finally, the school-specific preference shocks τ_i and ξ_i are also taken to be unrelated to potential achievement. I next discuss the latter two assumptions in detail and provide suggestive evidence in support of them.

5.2 Exclusion of Distance

If distance to charter schools is unrelated to ζ_{ijgk} and affects students' charter attendance decisions, it is a valid instrument for charter school attendance. The use of this instrument parallels the use of proximity-based instruments in research on higher education (see, e.g., Card 1993). The exclusion of distance seems plausible in the model described here, since X_i includes detailed student characteristics and baseline achievement. The exclusion restriction requires that distance to charter schools is effectively random conditional on these variables. Together, the lottery and distance instruments identify the taste coefficients $\alpha^{\theta}_{jg\ell}$ in equations (7) and (8). For example, if lottery applicants who are willing to travel long distances to attend charters benefit more than students who live nearby, this suggests that unobserved tastes for charters are positively related to achievement gains from charter attendance. Appendix B demonstrates how the combination of lotteries and distance identifies the taste coefficients in a simplified model with one charter school.

Table 5 explores the validity of the distance instrument and compares IV estimates based on lotteries and distance. Columns (1) and (2) report coefficients from ordinary least squares (OLS) regressions of 4th grade test scores on distance to the closest charter middle school, measured in miles. The estimates in the first row show that students who live further from charter middle schools have significantly higher 4th grade test scores, suggesting that charter schools tend to locate in low-achieving areas of Boston. The second row shows that adding controls for observed characteristics shrinks these imbalances considerably and renders the coefficients statistically insignificant. This suggests that observable student characteristics capture the relationship between location and academic achievement, and lends plausibility to the use of distance as an instrument in models that control for these characteristics.

Columns (3) through (5) of Table 5 compare 2SLS estimates using lottery offers and distance as instruments for charter attendance. Models using the lottery instrument control for lottery fixed effects and limit the sample to applicants, while models using the distance instrument control for demographics and 4th grade test scores, and include the full sample. Column (3) shows that both instruments have strong, statistically significant first stage effects on charter attendance: winning a lottery increases the probability of charter attendance among applicants by 0.65, while a one-mile increase in distance decreases the probability of charter attendance by

2.2 percentage points. Columns (4) and (5) show that the two instruments produce similar estimates of the effects of charter attendance, though the distance estimates are much less precise. The lottery instrument produces estimates of 0.50σ and 0.31σ for math and ELA tests, while the distance instrument generates estimates of 0.54σ and 0.19σ. The structural estimates reported below efficiently combine information from both sources of variation in charter attendance.[21]

5.3 Exclusion of School-specific Preferences

The exclusion of school-specific tastes from equations (7) and (8) implies that selection on unobservables has a "single-index" form: the relationship between potential achievement and unobserved preferences is driven only by the average charter school taste θ_i. This allows selection on unobservables to be parsimoniously characterized by the set of coefficients $\alpha^{\theta}_{jg\ell}$. The single-index restriction requires that students view charter schools as a homogeneous treatment, implemented at multiple sites throughout Boston. Students may know about cross-site heterogeneity in average effects (captured by α^{0}_{jk}) and about their own suitability for the charter treatment (captured by θ_i), but the model requires that students do not make choices based on their own idiosyncratic treatment gains across sites.

The exclusion of school-specific preferences is more plausible if the charter treatment is in fact homogeneous across schools. Table 6 lists responses to a survey on school practices for the seven charter middle schools included in the sample.[22] For comparison, column (8) reports average responses for other charter middle schools in Massachusetts. The survey results show that practices are highly uniform across Boston's charter middle schools, and differ markedly from schools elsewhere in the state. Boston middle schools have more instructional time than other charter schools; five of seven have longer school years than the non-Boston average, and six of seven have longer school days. The seven Boston middle schools all strongly identify with the No Excuses educational approach, choosing at least 4 on a 5-point scale measuring adherence to No Excuses, whereas other charter schools are less likely to identify with No Excuses. Boston's schools uniformly emphasize traditional reading and math skills, discipline and comportment, and measurable results, while other charters place less emphasize on these ideas. With a few exceptions, Boston middle schools ask parents and students to sign commitment contracts, require students to wear uniforms, and utilize formal merit/demerit systems to reward and punish student behavior. In the classroom, cold-calling and drills for math and reading are

[21] The lottery and distance instruments need not produce similar estimates even if both instruments are valid, because they identify effects for different sets of compliers. The close correspondence between the lottery and distance estimates in Table 5 suggests that average effects for lottery and distance compliers are similar, though I cannot reject relatively large differences due to the imprecision of the distance estimates.

[22] Schools are randomly ordered to avoid divulging information about individual schools.

commonly used by Boston's charters, and less likely to be used by charters elsewhere. The pattern in Table 6 shows that educational practices are similar across charter middle schools in Boston, lending support to the assumption that students are unlikely to make choices based on private information about school-specific achievement gains.

To further motivate the exclusion of τ_i and ξ_i from equations (7) and (8), Table 7 summarizes the relationship between distance to charter schools and the choice of schools among charter applicants. In the model outlined above, the decision to choose one charter school over another is determined by the combination of distance and school-specific tastes. If application portfolio choices are dominated by distance, then there is little scope for selection on school-specific tastes.

The results in Table 7 show that the choice of school conditional on applying is determined mostly by distance. Thirty-nine percent of applicants apply to the closest school, and these students travel an average of 1.63 miles to their chosen schools. An additional twenty-two percent apply to the second closest charter, traveling an average of 1.17 miles beyond the closest school, and 12 percent choose the third closest, on average traveling 1.82 miles further than necessary. Only 14 percent of applicants apply to the fifth, sixth, or seventh closest school. These facts show that although students are free to apply to distant schools, few do so; conditional on choosing to apply to a charter, most students apply to one close by, leaving little potential for matching on school-specific achievement gains.

5.4 Estimation

I estimate the parameters of the model by maximum simulated likelihood (MSL). Let Ω denote the parameters of equations (5) through (9). The likelihood contribution of a student with endogenous variables (A_i, Z_i, S_i, Y_i) is given by

$$\mathcal{L}_i(\Omega) = \int 1\{A^*(X_i, D_i, \theta, \tau) = A_i\} \cdot f(Z_i|A_i;\pi) \cdot P_{is(i)}(Z_i, \theta, \tau) \times \phi_m\left(\tilde{Y}_i(\theta)\right) dF(\theta, \tau | X_i, D_i, \Omega) \quad (11)$$

where

$$\tilde{Y}_{ijg}(\theta) \equiv Y_{ijg} - \alpha^0_{js(i)} - \alpha^0_{jgs(i)} - X_i'\alpha^x_{jgs(i)} - \alpha^\theta_{jgs(i)} \cdot \theta$$

and $\phi_m(\cdot)$ is the multivariate normal density function of the ϵ_{ijgk} implied by equation (9). The subscript $s(i)$ denotes the school attended by student i.[23]

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

I evaluate the integral in equation (11) by simulation. The indicator function in the integrand creates practical difficulties: for some values of Ω there may be no simulations where the observed value of A_i is chosen, leading to a zero value for the likelihood (note that there are $2^7 = 128$ possible values of A_i). I therefore approximate the indicator function with a logit kernel smoother. For λ close to zero, we have

$$1\{A^*(X_i, D_i, \theta_i, \tau_i) = a\} \approx \frac{\exp\left(\frac{V_i(a,\theta_i,\tau_i)}{\lambda}\right)}{\sum_{a' \in \{0,1\}^K} \exp\left(\frac{V_i(a',\theta_i,\tau_i)}{\lambda}\right)}. \tag{12}$$

The kernel smoothing approach, suggested by McFadden (1989) and discussed in detail by Train (2003), is equivalent to estimation without smoothing in the limit as λ approaches zero, and it produces an objective function with better computational properties. In the empirical work to follow, I set $\lambda = 0.1$, the smallest value of λ for which the exponential functions in equation (12) did not evaluate to infinity during the estimation procedure.[24]

Let θ_i^r and τ_i^r be draws of θ and τ for individual i in simulation r. Define

$$\hat{\ell}_i^r(\Omega) = \frac{\exp\left(\frac{V_i(A_i,\theta_i^r,\tau_i^r)}{\lambda}\right)}{\sum_a \exp\left(\frac{V_i(a,\theta_i^r,\tau_i^r)}{\lambda}\right)} \cdot f(Z_i|A_i;\pi) \cdot P_{is(i)}(Z_i, \theta_i^r, \tau_i^r) \cdot \phi_m\left(\tilde{Y}_i(\theta_i^r)\right)$$

The simulated likelihood for observation i is

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

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

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

If R rises faster than $\sqrt{N}$, the MSL estimator is $\sqrt{N}$-consistent and has the same asymptotic distribution as the conventional maximum likelihood estimator (Train 2003).

6 Estimates of the Structural Model

6.1 Preference Parameters

The MSL estimates were produced using 200 draws of θ_i and τ_i for each student. The results were not sensitive to increasing the number of draws. To code S_i, I assigned a student to the

[24]Values of λ between 0.1 and 0.3 produced very similar results. An alternative to smoothing is to view the right-hand side of equation (12) as the true choice probability and treat λ as an additional parameter to be estimated (see, e.g., Berkovec and Stern 1991). Walker et al. (2004) note that estimating the smoothing parameter typically results in an objective function that is flat at the optimum, with poor numerical performance. Assuming the existence of tastes specific to combinations of charter applications also seems economically unattractive. For these reasons, I use the kernel smoothing approach.

charter school where she attended the most days in grades five through eight; students who spent no time in charter schools were assigned to public school.

Table 8 reports MSL estimates of the parameters governing preferences for charter schools.[25] Estimates of the vector γ^x are consistent with the demographic patterns reported in Table 2. Subsidized lunch status, special education, and limited English proficiency are associated with weak demand for charter schools, while black students and students with higher baseline math and ELA scores have stronger preferences for charters. The intercept of charter utility γ^0, which is the average of γ_k^0 across schools, is negative and significant. This implies that most students in the omitted demographic category would prefer traditional public schools to charter schools even in the absence of application and distance costs.[26] Column (3) of Table 8 shows average marginal effects of observed characteristics on the probability of applying to at least one charter school.[27] Poverty has the largest marginal effect on application behavior: Holding other variables constant, subsidized lunch status reduces the probability of submitting a charter application by 7 percentage points.

The bottom half of Table 8 reports estimates of the parameters governing preferences for distance, application costs, and heterogeneity in tastes. Increased distance significantly reduces the utility of charter school attendance. The marginal effect in column (3) shows that a one-mile increase in driving time to a charter school reduces the probability of applying to that school by roughly one percentage point.[28] The application cost γ^a is positive, large, and statistically significant. Its magnitude suggests that applying to a charter school involves a utility cost roughly equivalent to a 2.3 mile increase in driving distance. The estimates capturing unobserved heterogeneity in preferences for charter schools are statistically significant and economically important: in utility terms, a one-standard-deviation increase in θ_i is roughly equivalent to a 10-mile increase in distance to all charter schools. The corresponding effect for τ_i is about three miles. The last row of Table 8 shows that the average lottery admission probability is 0.64.[29]

[25] I calculate standard errors using the average outer product of the gradient of the simulated likelihood. Define

$$\hat{V} \equiv \frac{1}{N}\sum_i \nabla_\Omega \log \hat{\mathcal{L}}_i\left(\hat{\Omega}_{MSL}\right) \nabla_\Omega \log \hat{\mathcal{L}}_i\left(\hat{\Omega}_{MSL}\right)'$$

Then $\hat{V}^{-1}$ is a consistent estimate of the asymptotic variance of $\sqrt{N}\left(\hat{\Omega}_{MSL} - \Omega_0\right)$, where Ω_0 is the true parameter vector.

[26] Fourth grade scores are de-meaned in the estimation sample, so the omitted category is white males without subsidized lunch, limited English proficiency, or special education status with average 4th grade scores.

[27] Marginal effects for discrete variables are computed by simulating the model first with the relevant characteristic set to zero for each student and then with it set to one, and computing the average difference in application probabilities across these simulations. Marginal effects for continuous variables are average simulated numerical derivatives of the application probability.

[28] The reported marginal effect for distance is obtained by first computing the marginal effect of a one-mile increase in distance to each school on the probability of applying to that school, and then averaging these effects across schools.

[29] I allow admission probabilities to differ across schools and cohorts, and set them equal to one for students with

The MSL estimates imply that charter applicants are a highly selected subset of the student population. Define

$$\mathcal{P}_i \equiv X_i'\gamma^x + \theta_i$$

$\mathcal{P}_i$ indexes student i's preference for charter schools relative to public school as a function of observed characteristics and unobserved tastes. Panel A of Figure 5 plots kernel densities of $\mathcal{P}_i$ for applicants and all students in standard deviation units. The mean of the applicant preference distribution is more than a full standard deviation above the population mean, and the variance of preferences is also lower among applicants. This figure highlights the intuition for identification: the model uses variation in preferences among lottery applicants driven by the distance instrument to extrapolate beyond effects for applicants and predict the distribution of achievement effects for the full population.

6.2 Achievement Parameters

The structural estimates reveal important heterogeneity in the achievement effects of charter schools on both observed and unobserved dimensions. This can be seen in Table 9, which reports estimates of the parameters of the math potential outcome equations. Estimates for 6th, 7th, and 8th grade are shown in panels A, B, and C, respectively. In each panel, column (1) shows estimates for charter schools, column (3) shows estimates for public schools, and column (5) shows the difference, the causal effect of charter school attendance. Columns (2), (4), and (6) report standard errors.

The estimates in Table 9 show that charter schools have larger effects on math scores for more disadvantaged students. The constant term reported in column (5) implies that charter attendance raises math scores for students in the omitted demographic category by 0.37σ in 6th grade, 0.25σ in 7th grade, and 0.39σ in 8th grade.[30] Subsidized lunch students receive further benefits of around 0.15σ in every grade, while black and Hispanic students also experience larger gains. A one standard deviation improvement in baseline math scores decreases the effect of charter attendance by between 0.11σ and 0.16σ, and these baseline interactions are statistically significant. As shown in columns (1) and (3), blacks, Hispanics, and subsidized lunch students lag behind other students in public school, but these characteristics are not predictive of scores in charter schools conditional on the other included covariates. In this sense, the structural estimates imply that charter schools close math achievement gaps between racial and socioeconomic groups.

siblings at charter schools (siblings are guaranteed admission). The reported estimate is the average admission probability for randomized applicants across schools and cohorts.

[30]The constant reported for charter schools is an average of school-specific estimates.

The last row of each panel shows the relationship between unobserved charter preferences and math scores. The estimates in column (1) suggest that tastes for charter schools are not systematically related to achievement in charter schools. On the other hand, column (3) shows that students with stronger unobserved preferences for charters do better in public school in every grade. A one standard deviation increase in charter school tastes is associated with an increase of between 0.03σ and 0.07σ in public school math achievement. Together, these estimates imply that stronger tastes for charter schools are associated with slightly *smaller* math gains, a fact documented in column (5).

Table 10 shows that the pattern of estimates for ELA is broadly similar to the corresponding results for math.[31] Column (5) shows that charter schools have substantial effects on the ELA scores of students in the omitted demographic category (0.41σ and 0.39σ in 7th and 8th grade), with significantly larger effects for subsidized lunch students and students with low baseline math scores. As in math, race and subsidized lunch status are not predictive of ELA scores in charter schools. Furthermore, the pattern of negative selection on unobservables is more pronounced in ELA than in math: a one standard deviation increase in tastes for charter schools reduces the effect of charter attendance by 0.06σ and 0.10σ in 6th and 8th grade, though the estimate for 7th grade is not significant.

Taken together, the estimates reported in tables 9 and 10 reveal two important patterns. First, charter schools reduce differences in achievement between racial and socioeconomic groups. Race and subsidized lunch status are highly predictive of test scores in public school, but not predictive of charter school scores conditional on other characteristics. Baseline achievement is also less predictive of scores in charter schools than public schools, especially in math. Broadly speaking, charter schools raise scores the most for the students who lag furthest behind their peers, reducing achievement gaps relative to the public school counterfactual.

Second, when combined with the utility parameters in Table 8, the achievement parameters show a consistent pattern of *negative* selection on gains from charter attendance. Students with subsidized lunch status and those with low baseline math scores receive large achievement gains from charter schools, but have atypically weak preferences for charter attendance. Students with stronger unobserved tastes for charter schools also experience smaller gains, especially in reading. Black students are an exception to this pattern, as they have stronger-than-average preferences for charters and receive larger-than-average gains. As shown in panel B of Figure 5, however, the full set of preference and achievement coefficients implies a decreasing relationship between preferences for charter schools and achievement gains. To summarize the relation-

[31] Estimates of the model's covariance parameters are reported in Appendix Table A3.

ship between gains and preferences, this figure plots the causal effect of charter attendance conditional on the charter preference $\mathcal{P}_i$, averaged across K schools and 3 grades. Define

$$\beta_j(p) \equiv E\left[\frac{1}{3 \cdot K}\sum_{g=6}^{8}\sum_{k=1}^{K}\left(Y_{ijg}(k) - Y_{ijg}(0)\right)|\mathcal{P}_i = p\right]$$

Figure 5 plots the $\beta_j(p)$ functions implied by the MSL estimates.[32] For both math and ELA, $\beta_j(p)$ is downward sloping, reflecting the fact that students with stronger preferences for charter schools benefit less from charter attendance. This pattern is somewhat surprising; one might have expected the students with the largest potential benefits to be the most likely to seek out charter schools. Instead, the findings reported here suggest that disadvantaged students struggle the most in traditional public schools, but are unlikely to investigate the charter alternative.

6.3 School Effects

Table 11 reports estimates of the model's school-specific parameters, including the average utility γ_k^0, the admission probability π_k, and the test score effects $(\alpha_{jk}^0 - \alpha_{j60}^0)$ for $j \in \{m, e\}$ (the omitted grade effect in equation (7) is 6th grade, so the school-specific intercepts measure effects in grade six). The utility estimates in column (1) show that some charters are more popular than others, but all of the estimates are negative, indicating that on average students prefer traditional public school to charter attendance. The admission probabilities, which are averages across applicant cohorts, range from 0.39 to 0.87.

The achievement effects reported in Table 11 suggest that the math effects of Boston's charters are not driven by any particular school; all seven charter middle schools increase student achievement relative to traditional public school for students in the omitted demographic category. The estimated math effects range from 0.23σ to 0.64σ, and all of them are statistically significant. The ELA effects reported in column (4) vary more across schools, and only two are positive and statistically significant. Interestingly, the two most effective schools, School 6 and School 7, are relatively unpopular as measured by the average utility parameters in column (1). This mirrors the negative selection with respect to student characteristics documented in tables 8 through 10: Students with the most to gain are less likely to apply to charter schools, and students who apply to charters are less likely to choose the most effective schools.

6.4 Model Fit

The model estimated here matches observed charter application and attendance behavior well. This can be seen in Table 12, which shows empirical choice probabilities, together with

[32] There is no closed-form expression for $\beta_j(p)$. The plots are constructed from simulated data using local linear regressions with a triangle kernel and a bandwidth of 0.5 standard deviations.

the corresponding model-based predictions.[33] The model slightly under-predicts the fraction of students applying to charter schools (16.6 percent in the data compared to 14.4 percent from the model) and the fraction of students attending charters (10.0 percent compared to 9.3 percent). It also under-predicts the fraction of applicants that submit multiple applications: 26.7 percent of applicants choose more than one charter, while the model predicts a rate of 18.1 percent. The remaining rows of Table 11 show that the model generally matches the relative popularity of particular charter schools among charter applicants and attenders.

Table 13 shows that predicted math and ELA score distributions closely match the corresponding empirical distributions. For both math and ELA and in every grade, the model-predicted mean scores for charter and public school students are within 0.03σ of the corresponding empirical means. Predicted standard deviations are also very close to their empirical counterparts for all tests and grades in both charters and public schools.

7 Predicting Expansion Effects

7.1 Description of Expansions

The model estimated in this paper allows for a parsimonious description of charter demand that can be used to make out-of-sample predictions about student choices and achievement test scores in counterfactual environments. I use the model to investigate the effects of changing Boston's charter school network on the distribution of middle school test scores for the cohort of students attending 4th grade in 2009. I begin with a look at the effects of closing all of Boston's charter schools. This simulation shows how the existence of the charter sector has affected test scores. I then simulate the effects of Boston's planned expansion, which adds six new charter middle schools to the existing set of seven. This is followed by an analysis of progressively larger expansions that add schools one by one until the number of charter schools reaches 30. Finally, I simulate the effects of *forcing* all Boston middle school students to attend charter schools. While unrealistic, this last scenario allows me to compare the population average treatment effect (ATE) to other treatment parameters and put upper bounds on the possible effects of charter expansion.

7.2 Additional Assumptions

Expansions are defined by sets of charter schools, with each school characterized by a location, an average utility γ_k^0, math and ELA intercepts α_{mk}^0 and α_{ek}^0, and an admission probability

[33] The predictions are produced by simulating the model 100 times for each observation in the data set, and averaging across simulations and then observations.

π_k. Choosing these characteristics for new schools requires additional assumptions. I next describe these assumptions in detail and outline the procedures I use to acquire the parameters for each new school.

7.2.1 School Locations

To choose charter locations for Boston's planned expansion, I use the addresses of the new schools scheduled to open through 2013. Larger expansions require that I specify where additional schools are located. Charter schools in Boston usually locate in vacant buildings, such as empty churches (Roy 2010). To manage the computational complexity of the model, I do not model charter locations as strategic choices, and instead assign them randomly. Specifically, I draw addresses at random from a grid of half-mile by half-mile blocks covering Boston; if a drawn block does not already contain a charter school, I add the next school in this location. Other models of location choice might produce slightly different predictions of the effects of charter expansion. As shown in Section 7.3, however, my estimates imply that the market for charter schools will be saturated by the time 30 schools have been added: at this point, almost all students live in close proximity to a charter offering guaranteed admission. For expansions at this scale, other models of location choice are therefore likely to produce results similar to those reported below. Figure 6 shows the locations used for the counterfactual simulations.[34]

7.2.2 Utility and Test Score Parameters

I assign utility and test score effects for Boston's planned expansion using the MSL estimates for the linked schools listed in Table 2.[35] For larger expansions, I randomly assign each new school a draw of the vector $(\gamma_k^0, \alpha_{mk}^0, \alpha_{ek}^0)'$ from the estimated distribution of school effects. This approach requires two assumptions. First, I assume that new charter schools are drawn from the same distribution that generated the existing set of schools, which implies that new charter entrants can successfully replicate the education production function used by existing schools. This assumption seems plausible for Boston's planned expansion, which involves the expansion of existing charter schools to a small number of satellite campuses located elsewhere in the city. For larger expansions, replicating the production function may become increasingly difficult, as teachers, principals, or other inputs used by charter schools could become scarce. As with all out-of-sample simulation exercises, therefore, my quantitative predictions for larger expansions should be viewed as more uncertain.

[34] For the scenario in which all students are forced to attend charter schools, I randomly assign students to charters, so spatial locations are not relevant.

[35] KIPP Academy Boston is not linked to an existing Boston charter. I assign this school the mean utility and test score effects.

Second, I assume that the school-specific parameters of the utility and achievement equations remain constant as students' application and attendance decisions change in counterfactual scenarios. Importantly, this assumption implies that there are no peer effects on student achievement, so that students' own potential test scores are not affected by changes in their classmates' characteristics. I abstract from peer effects because differences in peer quality are unlikely to be the primary mechanism driving the effects of charter schools. As shown in Table 3, charter applicants' baseline test scores are about 0.25σ above the Boston average. For this difference in peer quality to explain the effects reported in Table 4, peer effects would have to be roughly 2σ and 1σ per standard deviation of baseline peer quality in math and ELA. In a summary of the peer effects literature, Sacerdote (2011) reports a wide range of estimates, almost all of which are substantially less than these magnitudes. This suggests that charter schools produce achievement gains mostly through channels other than peer effects, so ignoring peer effects is unlikely to cause important errors in predictions of charter expansion effects.[36]

7.2.3 Admission Probabilities

To predict admission probabilities in the counterfactual simulations, I assume that schools set their admission probabilities to maximize their own enrollments, accounting for optimal student behavior and admission probabilities at other schools. Specifically, I look for a Subgame Perfect Nash Equilibrium in which students correctly anticipate π_k and apply optimally, and schools optimally choose π_k given students' application decisions and other schools' admission probabilities. Appendix D formally outlines the structure of the game, shows that an equilibrium exists, and gives conditions under which the equilibrium is unique. I numerically solve for the equilibrium admission probabilities in each counterfactual simulation.

7.3 Simulation Results

Figures 7, 8, and 9 summarize the simulated effects of charter expansion on school choices, charter oversubscription, and test scores. In each panel, a vertical black line indicates the existing number of charter schools, and a red line indicates the size of Boston's planned charter expansion. Panel A of Figure 7 shows how application and attendance rates change as the number of charter schools rises, while panel B shows the effects of expansion on admission probabilities and the share of charter seats filled. Figure 8 shows effects on average math and

[36] To directly explore the relevance of peer effects in my sample, I followed the approach of Abdulkadiroglu et al. (2011) and estimated 2SLS regressions in the lottery sample that interact charter attendance with average peer baseline scores for students in the same lottery, instrumenting this interaction term with the interaction of peer quality and the lottery offer. The coefficients on the interaction terms were small and statistically insignificant for both math and ELA, suggesting that the effects of charter attendance are not larger for applicants who experience larger increases in peer quality when they attend charter schools.

ELA scores, and Figure 9 shows effects on white-black achievement gaps. Tables 14, 15, and 16 show numerical results for choice behavior, math scores, and ELA scores, respectively.

The simulations imply that charter schools have had a substantial impact on the distribution of test scores in Boston. This can be seen in the second row of Tables 15 and 16, which show the effects of closing all charter schools. Without charter schools, average middle school math scores in Boston would fall by between 5 and 11 percent relative to the state average, and ELA scores would fall by up to 8 percent. Closing Boston's charters is also predicted to increase citywide white-black achievement gaps by roughly 5 percent; the largest increase in the gap, 7 percent, occurs for 8th grade math and ELA.

The charter expansions scheduled to take place through 2013 are likely to produce further increases in average test scores and reductions in the achievement gap. Specifically, these expansions are predicted to raise charter application and attendance rates to 22 and 15 percent, raise average scores by 5 to 8 percent in math and 1 to 6 percent in ELA, and further reduce achievement gaps by 4 to 6 percent in math and 2 to 3 percent in reading. Columns (3), (6), and (9) of Tables 15 and 16 show the effect of treatment on the treated (TOT), the average effect of charter attendance for the students who attend charters in each counterfactual. The TOTs associated with Boston's planned expansion are all larger than the TOT for the existing system. This reflects the pattern of selection discussed in Section 6: at the margin, charter expansion draws in students with weaker tastes for charter schools, who receive larger achievement gains. Columns (3) and (4) of Table 14 show that the availability of additional charter seats reduces oversubscription. The average admission probability increases from 0.63 to 0.80, and only 85 percent of charter seats are filled, reflecting the fact that 6 of 13 schools are undersubscribed (see column (2) of Table A4).

Figure 7 through 9 display the effects of opening additional charter schools one by one. Opening more charters smoothly raises application and attendance rates, increases test scores, and reduces achievement gaps. However, panel B of Figure 7 shows that charters are increasingly undersubscribed as more are added, and the share of charter seats filled falls sharply as the number of schools increases. With 30 schools, 24 percent of students attend a charter, average 8th grade scores rise by 20 percent and 5 percent in math and ELA, and 8th grade achievement gaps in these subjects fall by 14 and 11 percent. However, only 60 percent of charter seats are filled.

The counterfactual simulations highlight two important insights implied by the structural estimates. First, since there is negative selection on achievement gains, the achievement effects of charter schools are increasing in the share of students who attend charters. This can be seen most clearly in the last rows of tables 15 and 16, which report the effects of a hypothetical expansion that forces all students to attend charter schools. When all students attend charters, the TOT is equal to the population average treatment effect (ATE). In both subjects and for all grades, the ATE is larger than the TOT for today's charter students. This result implies that effects for current charter applicants are *smaller* than potential effects for typical students in Boston.

Second, the simulation results imply that despite their large potential achievement effects, additional demand for charter schools in Boston is limited. In the 30-school expansion, few schools are oversubscribed, so almost all students live in close proximity to a charter offering guaranteed admission. Nevertheless, only one-third of students apply to a charter, and 40 percent of charter seats are left unfilled. This result is due to the large application cost reported in Table 8 together with the negative average utilities reported in Table 11, which are in turn driven by the fact that charter application rates in Boston are low despite reasonably high admission probabilities. Only students in the upper tail of the taste distribution are interested in attending charter schools, and these students are spread increasingly thin as the charter system expands. Since the model also predicts that charters would have substantial impacts on the test scores of the remaining non-applicants, this lack of demand implies that preferences for charter schools are dominated by factors other than achievement gains. This finding suggests that large-scale charter expansions may be ineffective without concomitant efforts to boost charter application rates.

8 Conclusion

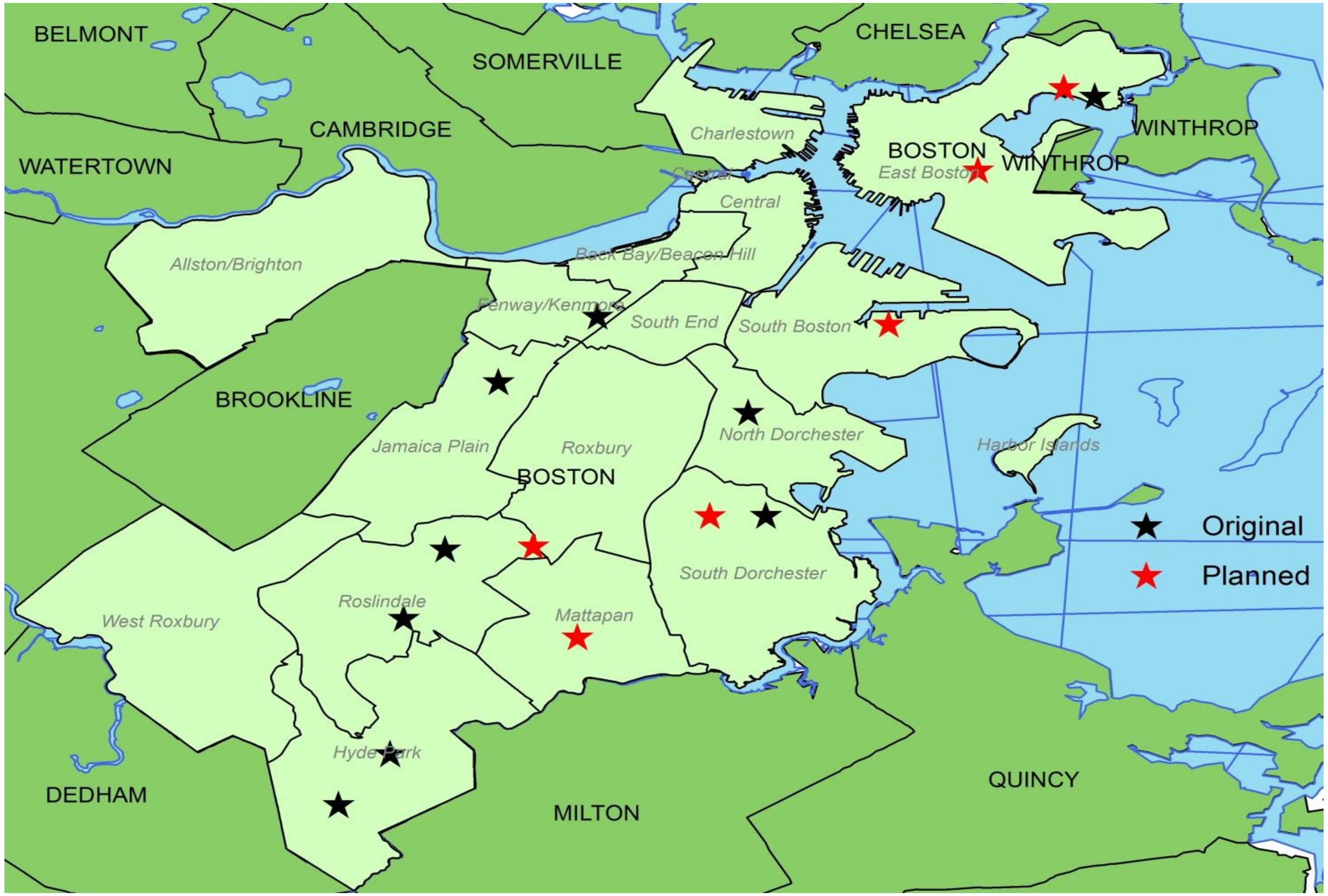
Estimates based on admission lotteries show that Boston's charter middle schools have substantial positive effects on test scores and quickly close racial achievement gaps among their applicants. At the same time, the implications of these findings for charter school expansion are unclear; applicants are a small, non-random subset of the student population, so such gains may be atypical. This paper develops a structural model of charter applications, school choice, and academic achievement that links the decision to apply to charter schools to achievement gains from charter attendance. To identify the parameters of the model, I combine two sets of instruments based on random lotteries and proximity to charter schools. I then use estimates of the model to predict out-of-sample effects for non-applicants and simulate the effects of charter expansion on the citywide test score distribution.

Estimates of the model reveal that tastes for charter schools are inversely related to achievement gains. Specifically, low-achievers and poor students gain the most from charter attendance, but are unlikely to apply to charters. Consistent with this finding, counterfactual simulations show that the average effect of charter schools is increasing in the size of the charter sector, as larger expansions draw in students with weaker preferences who receive larger gains. This pattern is surprising – the standard Roy model of selection suggests that students with larger potential gains should be more likely to apply. However, the "reverse Roy" pattern described here is consistent with results from other contexts, such as the long-term care insurance market described by Finkelstein and McGarry (2006). Finkelstein and McGarry note that the decision to purchase long-term care insurance is driven by both risk aversions and health risk; since more risk-averse people also tend to have lower health risks, they find that those who purchase more insurance are not higher risk on average. More generally, in settings where participation decisions are driven by multiple factors, selection on one dimension can lead to apparent negative selection on gains on another dimension. In the charter school context, application decisions are driven by socioeconomic status and baseline achievement, which are negatively correlated with achievement gains from charter attendance.

Despite the large effects of charter schools for marginal applicants, however, the structural estimates also imply that charter demand in Boston is limited. Most students prefer traditional public school attendance to charter schools, and act as if applying to charter schools is costly. As a result, when the charter market share reaches 24 percent, most schools are undersubscribed and 40 percent of charter seats are left empty. This finding suggests that skeptical views of large-scale charter school expansion, such as those expressed by Ravitch (2010) and Rothstein (2004), may reach the right conclusion, but for the wrong reason. Charter schools do not attract students who receive atypically large benefits, but the effects of charter expansion may nevertheless be attenuated by weak demand.

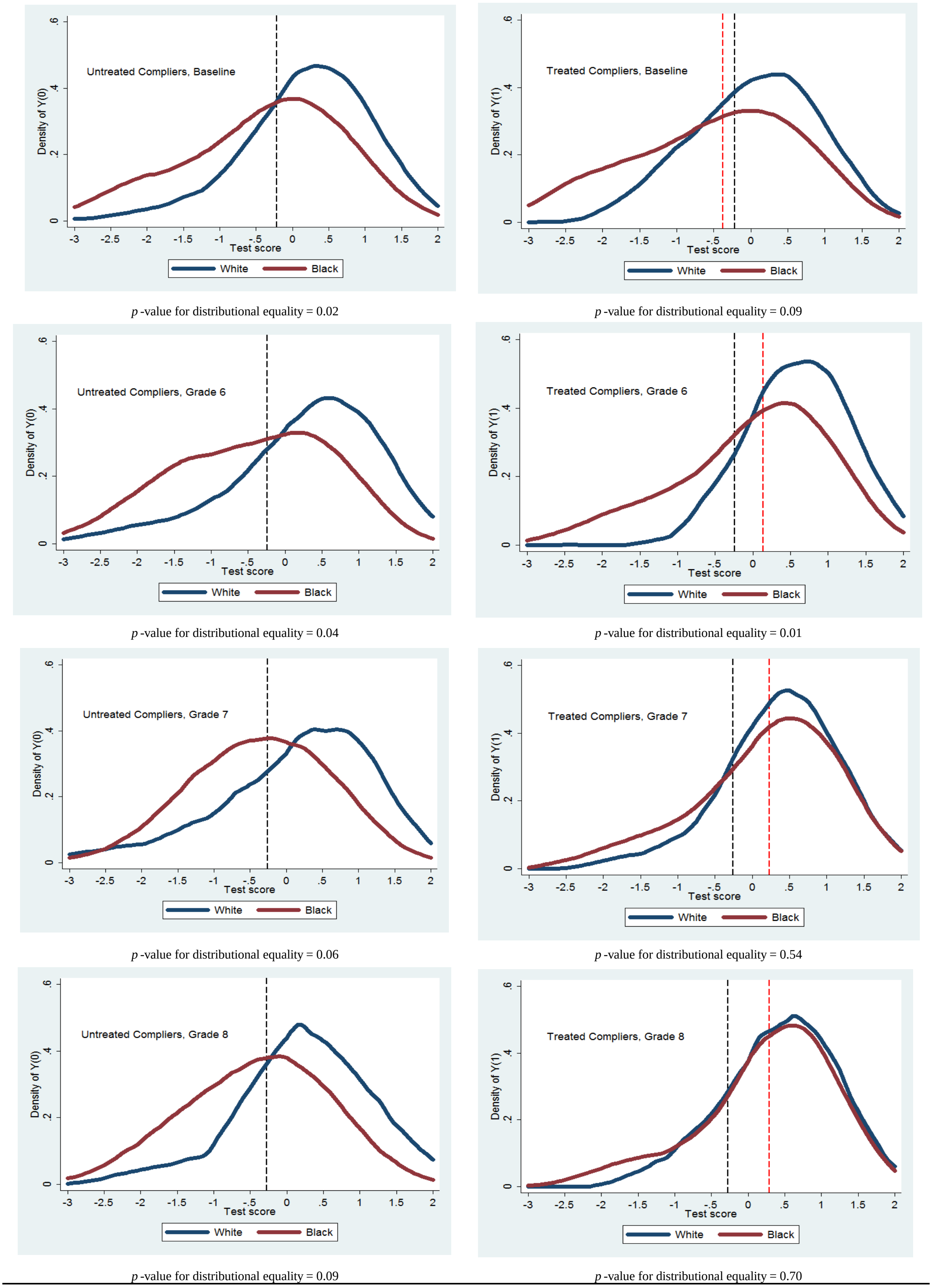
These findings raise the further question of whether parents who forgo large potential achievement gains are truly uninterested in achievement, or simply unaware of differences in effectiveness across schools. The model estimated in this paper does not distinguish between these two possibilities. In the former case, the simulation results reported here reflect the long-run demand for charter schools in Boston, and the potential achievement gains from charter expansion are ultimately limited. On the other hand, if the lack of demand for charter schools reflects a lack of parental information, the demand for charters may rise in the long run as parents become more informed, and the simulation results correspond to a short-run equilibrium. In related work, Hastings and Weinstein (2008) show that providing test score information leads parents to choose higher-performing schools, which suggests that informational frictions may play a role. Changes in recruitment practices may also change the pattern of selection into charter schools; the recent legislation authorizing charter expansion in Massachusetts requires schools to take efforts to recruit applicants who are demographically similar to students in the local district. In future work, I plan to use data from Boston's planned expansion to validate the model estimated here, and study changes in the demand for charter schools as the city's charter network expands.

Figure 1: Charter Middle Schools in Boston



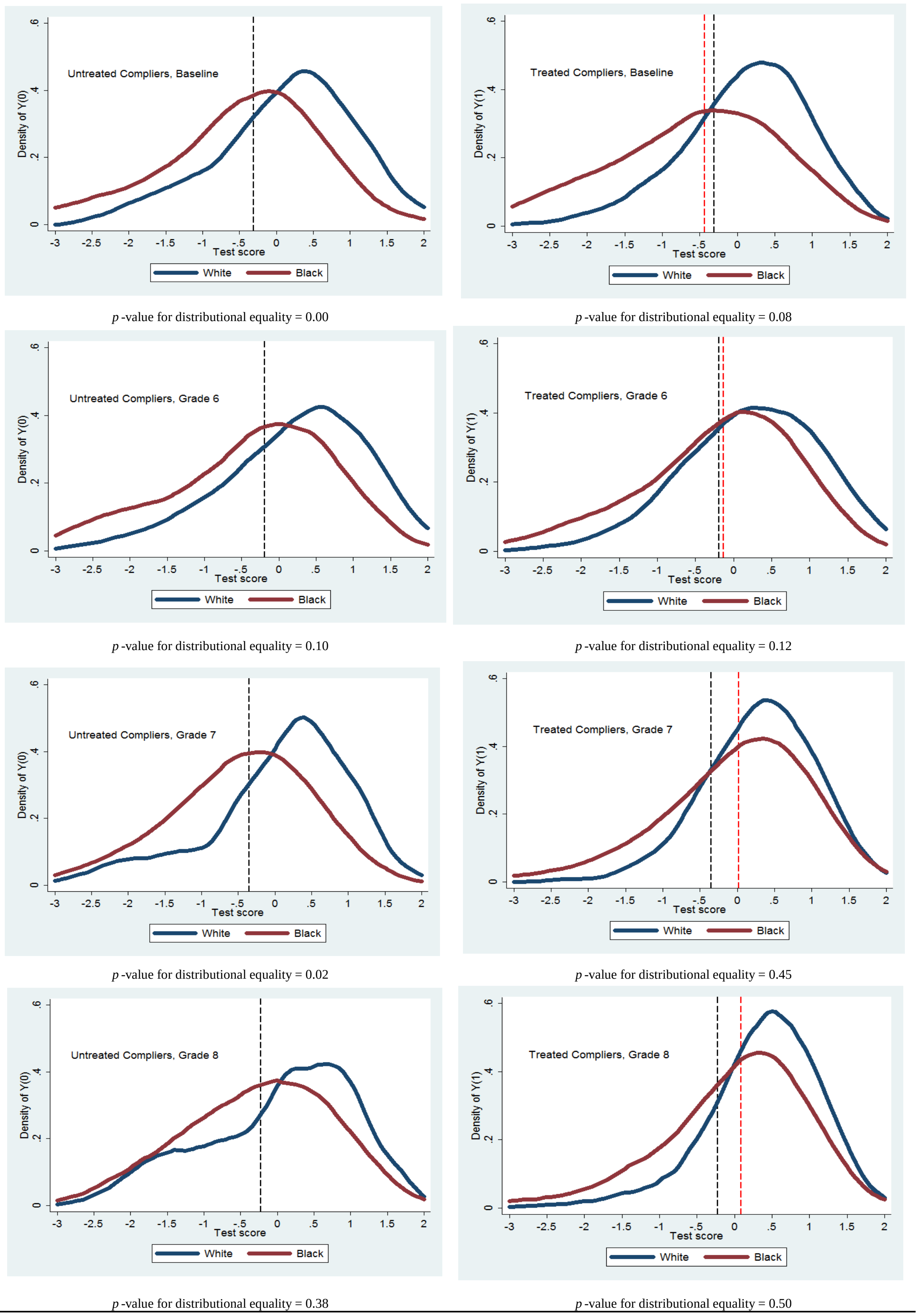
Notes: Black stars mark the locations of charter middle schools operating between 1995-1996 and 2010-2011. Red stars mark the locations of charter middle schools scheduled to open in 2011-2012 or later.

Figure 2: Math Score Distributions for Black and White Compliers



Notes: This figure plots kernel densities of math test scores for black and white lottery compliers. Black vertical lines show mean scores for untreated compliers, while red vertical lines in the treated plots show means for treated compliers. The treated densities are estimated from 2SLS regressions of a kernel smoother interacted with a charter dummy on the charter dummy and lottery fixed effects. The charter dummy is instrumented with an indicator for winning the lottery. All densities use a triangle kernel with a bandwidth of 1.25 standard deviations. Untreated densities are estimated using analogous regressions that replace the charter dummy with a non-charter dummy. *P* -values are from Kolmogorov-Smirnov tests of distributional equality.

Figure 3: ELA Score Distributions for Black and White Compliers



Notes: This figure plots kernel densities of ELA test scores for black and white lottery compliers. The treated densities are estimated from 2SLS regressions of a kernel smoother interacted with a charter dummy on the charter dummy and lottery fixed effects. The charter dummy is instrumented with an indicator for winning the lottery. All densities use a triangle kernel with a bandwidth of one standard deviation. Untreated densities are estimated using analogous regressions that replace the charter dummy with a non-charter dummy. *P* -values are from Kolmogorov-Smirnov tests of distributional equality.

Figure 4: Sequence of Events

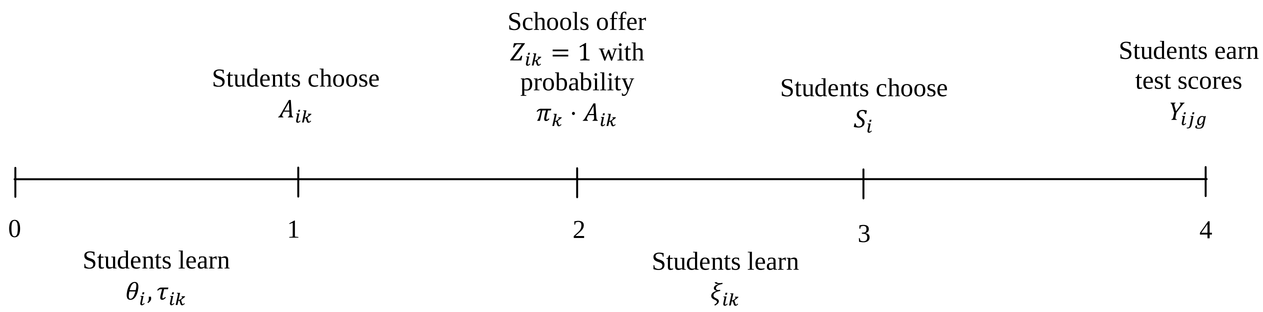
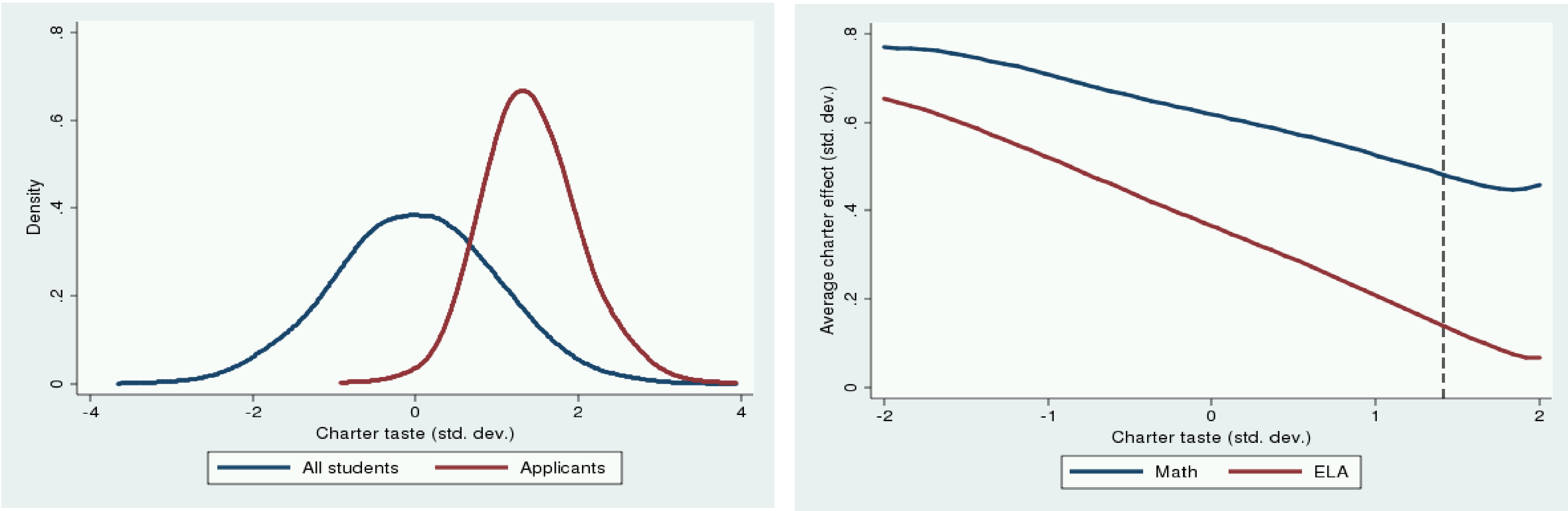


Figure 5: Charter Tastes and Achievement Effects

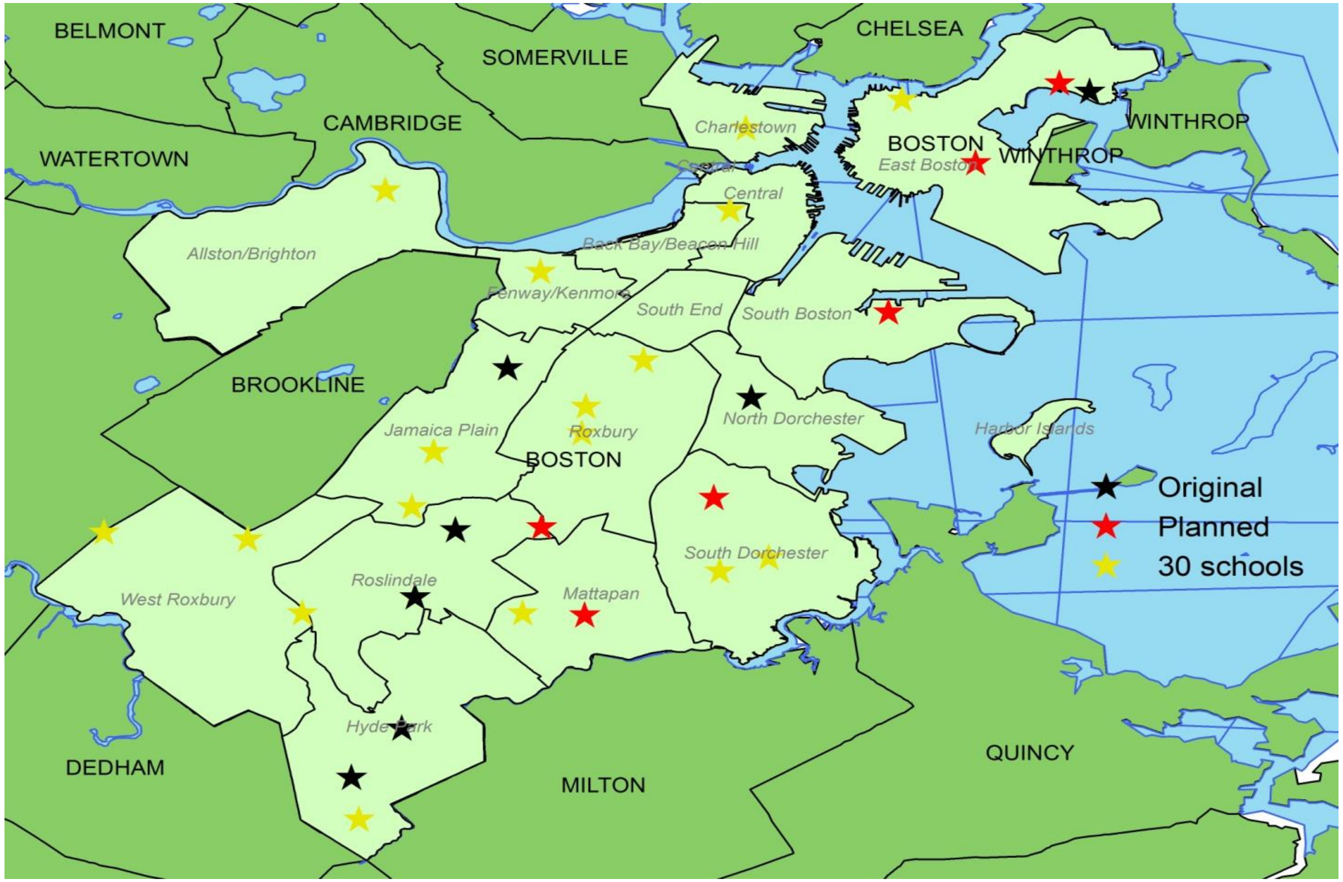


A. Charter Tastes

B. Charter Effects

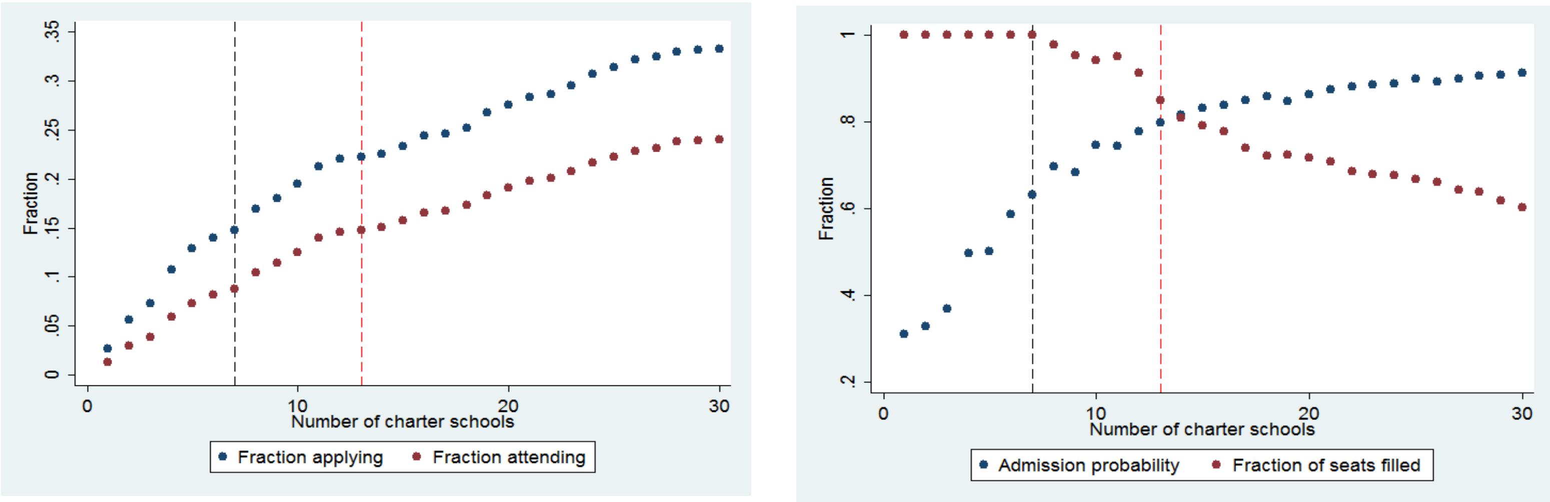
Notes: This figure describes the relationship between tastes for charter school attendance and the achievement effects of charter schools relative to traditional public school. Panel A plots the distribution of preferences for charter schools among charter applicants and all students as a function of observed and unobserved characteristics. Panel B plots the average causal effect of charter school attendance at each value of charter tastes for math and ELA. The vertical dashed lines indicates the mean preferences among applicants. The plots are produced from simulating the model using the MSL estimates and the empirical distribution of observed student characteristics. The graphs show kernel densities (panel A) or local linear regressions (panel B) estimated in the simulated data with a triangle kernel and a bandwidth of 0.5 standard deviations.

Figure 6: Charter School Expansions



Notes: Black stars mark the locations of the charter middle schools used to estimate the structural model. Red stars mark the locations of schools scheduled to open in 2011-2012 and 2012-2013. Yellow stars mark the locations of schools opened in a hypothetical expansion that raises the number of charters to 30.

Figure 7: Simulated Effects of Charter School Expansion -- Choice Behavior

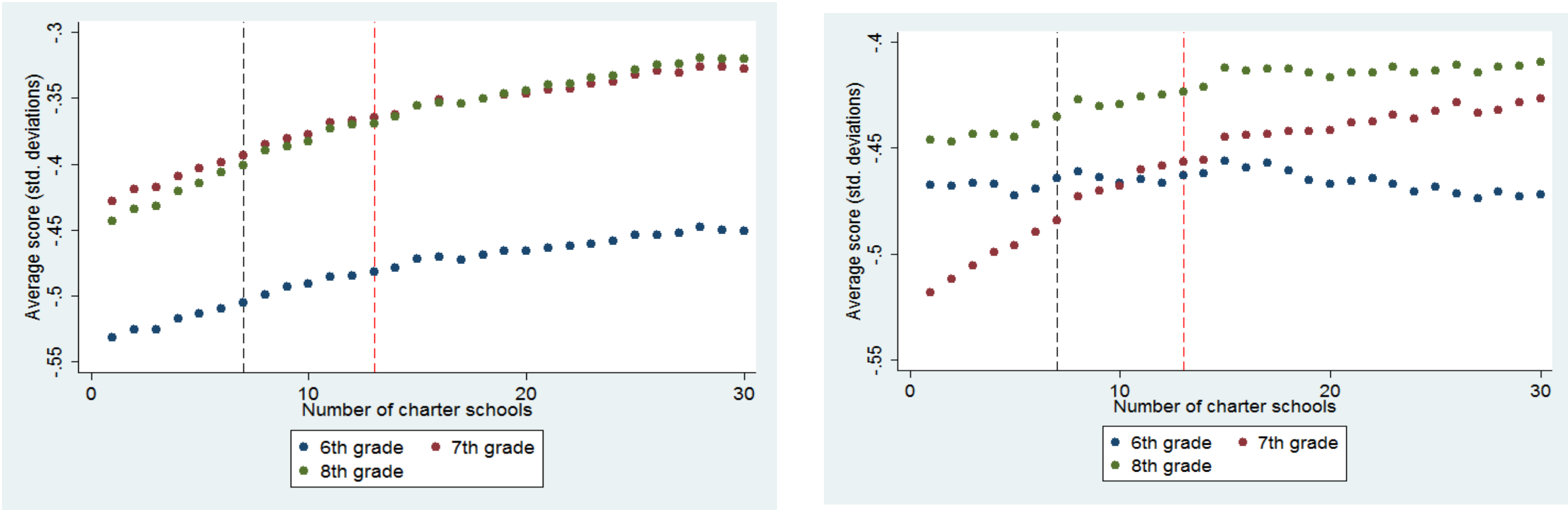


Panel A. Applications and attendance

Panel B. Oversubscription

Notes: This figure displays simulated effects of charter school expansio on application and attendance behavior. The black dashed line in each panel corresponds to the existing number of charter schools, while the red dashed line corresponds to Boston's planned expansion. Panel A shows the fraction of students who apply to and attend charter schools. Panel B shows the average admission probability for applicants and the fraction of charter seats that are filled. The figures are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the structural sample.

Figure 8: Simulated Effects of Charter School Expansion -- Average Citywide Test Scores

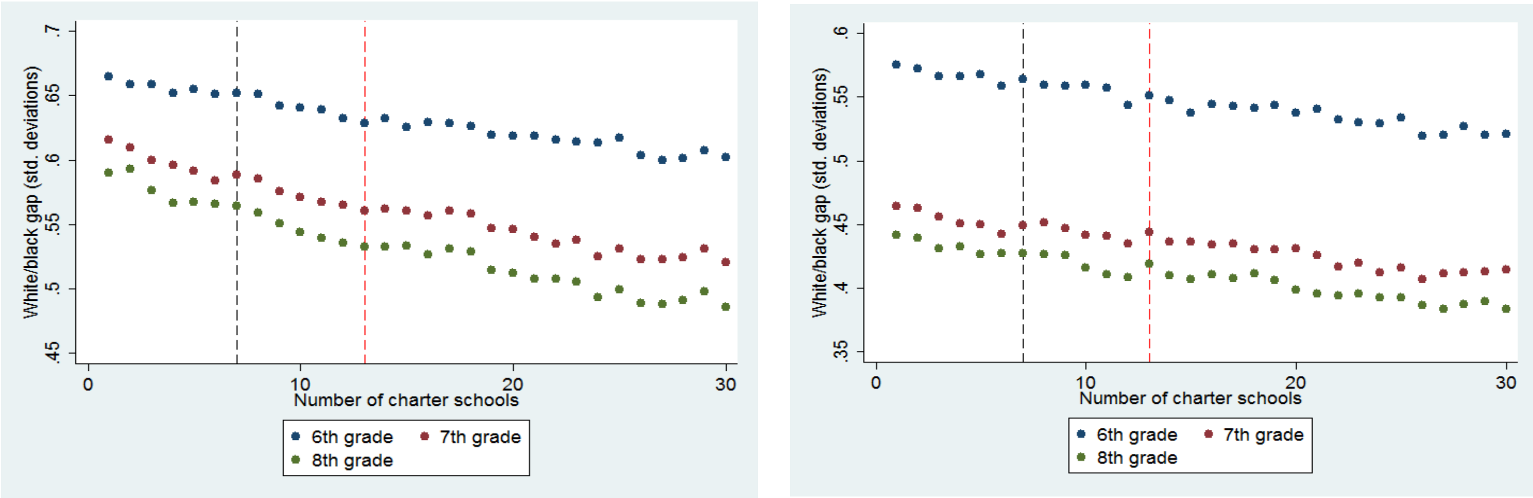


Panel A. Average math scores

Panel B. Average ELA scores

Notes: This figure displays simulated effects of charter school expansion on average test scores in Boston.The black dashed line in each panel corresponds to the existing number of charter schools, while the red dashed line corresponds to Boston's planned expansion. Panel A shows math scores, while Panel B shows ELA scores. The figures are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the structural sample.

Figure 9: Simulated Effects of Charter School Expansion -- Citywide Achievement Gaps



Panel A. Math achievement gap

Panel B. Average ELA achievement gap

Notes: This figure displays simulated effects of charter school expansion on white/black achievement gaps in Boston. The black dashed line in each panel corresponds to the existing number of charter schools, while the red dashed line corresponds to Boston's planned expansion. Panel A shows the gap in math, while Panel B shows the ELA gap. The figures are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the structural sample.

Table 1: Boston Charter Middle Schools

School name (1)	Grade coverage (2)	Years open (3)	Lotteries available (4)
Academy of the Pacific Rim	5-12	1997-present	Yes
Boston Collegiate	5-12	1998-present	Yes
Boston Preparatory	6-12	2004-present	Yes
Edward Brooke	K-8 (with 5th entry)	2002-present	Yes
Excel Academy	5-8	2003-present	Yes
Frederick Douglass	6-10	2000-2005	No
MATCH Middle School	6-8	2008-present	Yes
Smith Leadership Academy	6-8	2003-present	No
Roxbury Preparatory	6-8	1999-present	Yes
Uphams Corner	5-8	2002-2009	No

Notes: This table lists charter middle schools serving traditional student populations in Boston, Massachusetts from 1997-1998 through 2010-2011. Schools are included if they accept students in 5th or 6th grade. Column (3) lists the calendar years of operation for each school through 2010. Column (4) indicates whether lottery records were available for cohorts of applicants attending 4th grade between 2006 and 2009.

Table 2: Boston Charter Middle School Expansions

School name (1)	Grade coverage (2)	Year opening (3)	Linked schools (4)
Dorchester Preparatory	5-12	2012-2013	Roxbury Preparatory
Edward Brooke II	K-8 (with 5th entry)	2011-2012	Edward Brooke
Edward Brooke III	K-8 (with 5th entry)	2012-2013	Edward Brooke
Excel Academy II	5-12	2012-2013	Excel Academy
Grove Hall Preparatory	5-12	2011-2012	Roxbury Preparatory
KIPP Academy Boston	5-8	2012-2013	KIPP Academy Lynn (Lynn, MA)

Notes: This table lists Boston charter middle schools opening in 2011-2012 and 2012-2013. Schools are included if they planned to serve traditional student populations and accept students in 5th or 6th grade. Column (2) lists the planned grade coverage for each school. Column (3) shows the academic year in which the school opened. Column (4) lists Massachusetts charter middle schools operated by the same organization.

Table 3: Descriptive Statistics

	All students (1)	Charter applicants (2)
Applied to charter school	0.166	1.00
Attended charter school	0.100	0.571
Female	0.481	0.483
Black	0.461	0.520
Hispanic	0.390	0.306
Subsidized lunch	0.817	0.715
Special education	0.231	0.178
Limited English proficiency	0.206	0.130
Miles to closest charter school	1.84	1.64
4th grade math score	-0.544	-0.326
4th grade ELA score	-0.657	-0.418
N	10986	1822

Notes: This table shows descriptive statistics for students attending 4th grade at traditional public schools in Boston between 2006 and 2009. The sample excludes students with missing middle school test scores.

Table 4: Lottery-based Estimates of Charter Effects

Race	First stage (1)	Math effect (2)	ELA effect (3)
All	0.652***	0.502***	0.306***
	(0.024)	(0.087)	(0.090)
N		3792	
White	0.599***	0.139	-0.156
	(0.052)	(0.169)	(0.183)
N		638	
Black	0.674***	0.615***	0.377***
	(0.027)	(0.113)	(0.117)
N		1989	
Hispanic	0.634***	0.569***	0.534***
	(0.038)	(0.145)	(0.148)
N		1165	
p -value for racial equality	-	0.053	0.010

Notes: This table reports 2SLS estimates of the effects of attendance at Boston charter schools on test scores for lottery applicants. The sample stacks test scores in grades 6 through 8. The treatment variable is a dummy for attending any charter school after the lottery and prior to the test. The instrument is a dummy for receiving an offer from any charter school. Column (1) reports coefficients from regressions of charter attendance on the offer variable. Columns (2) and (3) report second stage estimates for math and ELA scores. All models control for lottery fixed effects. *P* -values are from Wald tests of the hypothesis that the 2SLS coefficients are the same across races. Standard errors are robust to heteroskedasticity and are clustered at the student level.

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

Table 5: The Distance Instrument

	Balance check: 4th grade scores			2SLS comparison		
	Math	ELA		First stage	Math	ELA
Controls	(1)	(2)	Instrument	(3)	(4)	(5)
None	0.031***	0.042***	Lottery	0.652***	0.502***	0.306***
	(0.010)	(0.011)		(0.024)	(0.087)	(0.090)
	10986				3792	
Student characteristics	-0.003	0.011	Distance	-0.022***	0.540**	0.190
	(0.007)	(0.007)		(0.003)	(0.251)	(0.240)
	10986				23638	

Notes: Columns (1) and (2) show regressions of 4th grade test scores on miles to the closest charter middle school. The first row includes no controls, while the second controls for student characteristics, including sex, race, free lunch status, special education status, limited English proficiency, and 4th grade score in the other subject. Columns (3) through (5) show 2SLS results for middle school test scores using the lottery and distance instruments. The lottery 2SLS models control for lottery fixed effects, while the distance models control for demographics and 4th grade test scores. Standard errors for the 2SLS models are clustered at the student level.

Table 6: School Practices

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

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

Table 7: Charter School Choices Among Applicants

Fraction of applicants choosing:	Fraction (1)	Mean distance (2)	Extra distance (3)
Closest charter	0.394	1.63	0.00
2nd closest	0.219	2.61	1.17
3rd closest	0.119	3.51	1.82
4th closest	0.115	4.98	3.22
5th closest	0.072	6.04	4.26
6th closest	0.066	7.66	5.89
7th closest	0.005	10.58	8.73

Notes: This table shows the fractions of applicants in the structural sample applying to charter schools by distance. Column (1) shows fractions of students making each choice. Column (2) shows mean distance to applicants' chosen schools. Column (3) shows extra distance relative to the closest charter school. If an applicant applied to multiple charters, the closest one is used in these calculations.

Table 8: Maximum Simulated Likelihood Estimates of Utility Parameters

Parameter	Description	Estimate (1)	Standard error (2)	Marginal effect (3)
γ^0	Constant	-2.10***	0.153	-
γ^x	Female	0.086	0.069	0.007
	Black	0.411***	0.115	0.031
	Hispanic	0.042	0.118	0.004
	Subsidized lunch	-0.769***	0.095	-0.068
	Special education	-0.162*	0.088	-0.011
	Limited English proficiency	-0.513***	0.101	-0.041
	Baseline math score	0.203***	0.050	0.017
	Baseline ELA score	0.108**	0.050	0.008
γ^d	Distance (miles)	-0.238***	0.010	-0.007
γ^a	Application cost	0.543***	0.006	-
σ_θ	Standard deviation of charter school tastes	2.45***	0.101	-
σ_τ	Standard deviation of school-specific tastes	0.773***	0.031	-
π	Acceptance probability	0.640***	0.046	-
N	Sample size		10986	

Notes: This table reports maximum simulated likelihood estimates of the parameters of the structural school choice model. Column (1) reports parameter estimates, while column (2) reports standard errors. The constant is the average of school-specific intercepts, and its standard error is computed using the delta method. The reported acceptance probability is an average across schools and cohorts. Column (3) reports average marginal effects of observed characteristics on the probability of applying to at least one charter school. The marginal effect for distance is the effect of a one-mile increase in road distance to a school on the probability of applying to that school, averaged across schools.

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

Table 9: Maximum Simulated Likelihood Estimates of Math Achievement Parameters

Parameter	Description	Charter school		Traditional public school		Charter effect	
		Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
		(1)	(2)	(3)	(4)	(5)	(6)
		Panel A: 6th grade					
$\alpha^0_{m6\ell}$	Constant	0.119	0.079	-0.249***	0.025	0.368***	0.083
$\alpha^x_{m6\ell}$	Female	0.007	0.038	-0.006	0.014	0.013	0.040
	Black	-0.068	0.063	-0.199***	0.022	0.131**	0.067
	Hispanic	0.078	0.067	-0.091***	0.023	0.169**	0.071
	Subsidized lunch	0.016	0.045	-0.131***	0.020	0.147***	0.049
	Special education	-0.355***	0.044	-0.337***	0.016	-0.018	0.047
	Limited English proficiency	-0.097*	0.053	0.042**	0.018	-0.139**	0.056
	Baseline math score	0.400***	0.026	0.564***	0.010	-0.164***	0.028
	Baseline ELA score	0.159***	0.025	0.105***	0.009	0.054**	0.026
$\alpha^\theta_{m6\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	-0.039**	0.017	0.027***	0.005	-0.066***	0.018
		Panel B: 7th grade					
$\alpha^0_{m7\ell}$	Constant	0.110	0.095	-0.140***	0.028	0.249**	0.099
$\alpha^x_{m7\ell}$	Female	0.095**	0.044	0.002	0.015	0.093**	0.046
	Black	0.041	0.074	-0.199***	0.025	0.240***	0.078
	Hispanic	0.136*	0.080	-0.086***	0.026	0.223***	0.085
	Subsidized lunch	0.010	0.051	-0.147***	0.023	0.157***	0.056
	Special education	-0.360***	0.059	-0.361***	0.018	0.001	0.061
	Limited English proficiency	0.018	0.063	0.080***	0.020	-0.061	0.066
	Baseline math score	0.372***	0.031	0.480***	0.011	-0.108***	0.032
	Baseline ELA score	0.100***	0.031	0.106***	0.010	-0.005	0.033
$\alpha^\theta_{m7\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	0.021	0.019	0.057***	0.006	-0.037*	0.020
		Panel C: 8th grade					
$\alpha^0_{m8\ell}$	Constant	0.230*	0.126	-0.164***	0.035	0.394***	0.130
$\alpha^x_{m8\ell}$	Female	-0.007	0.057	-0.029	0.019	0.023	0.060
	Black	0.048	0.082	-0.218***	0.031	0.266***	0.088
	Hispanic	0.156*	0.093	-0.107***	0.032	0.262***	0.098
	Subsidized lunch	0.043	0.068	-0.088***	0.028	0.131*	0.074
	Special education	-0.386***	0.070	-0.387***	0.022	0.001	0.073
	Limited English proficiency	0.003	0.080	0.083***	0.024	-0.080	0.084
	Baseline math score	0.347***	0.041	0.487***	0.013	-0.140***	0.042
	Baseline ELA score	0.026	0.040	0.066***	0.012	-0.039	0.041
$\alpha^\theta_{m8\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	-0.016	0.026	0.019***	0.006	-0.035	0.027

Notes: This table reports maximum simulated likelihood estimates of the parameters of the math achievement distribution. The constant in the charter school equation is the sum of the average of the school-specific effects and the relevant grade effect. Its standard error is computed using the delta method.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Maximum Simulated Likelihood Estimates of ELA Achievement Parameters

		Charter school		Traditional public school		Charter effect	
		Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Parameter	Description	(1)	(2)	(3)	(4)	(5)	(6)
				Panel A: 6th grade			
$\alpha^0_{e6\ell}$	Constant	-0.194**	0.090	-0.289***	0.026	0.095	0.094
$\alpha^x_{e6\ell}$	Female	0.094**	0.042	0.166***	0.015	-0.072	0.045
	Black	0.028	0.066	-0.144***	0.024	0.172**	0.070
	Hispanic	0.024	0.069	-0.088***	0.025	0.112	0.073
	Subsidized lunch	-0.049	0.049	-0.142***	0.022	0.093*	0.053
	Special education	-0.319***	0.050	-0.331***	0.016	0.011	0.053
	Limited English proficiency	-0.036	0.064	-0.051***	0.019	0.015	0.066
	Baseline math score	0.096***	0.028	0.176***	0.010	-0.080***	0.030
	Baseline ELA score	0.492***	0.028	0.462***	0.009	0.030	0.029
$\alpha^\theta_{e6\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	0.027	0.021	0.086***	0.006	-0.059***	0.022
				Panel B: 7th grade			
$\alpha^0_{e7\ell}$	Constant	0.005	0.113	-0.407***	0.030	0.411***	0.117
$\alpha^x_{e7\ell}$	Female	0.145***	0.050	0.211***	0.017	-0.065	0.053
	Black	0.027	0.079	-0.098***	0.027	0.125	0.083
	Hispanic	0.136	0.087	-0.009	0.028	0.146	0.091
	Subsidized lunch	0.033	0.059	-0.138***	0.025	0.171***	0.064
	Special education	-0.425***	0.067	-0.400***	0.019	-0.025	0.069
	Limited English proficiency	-0.067	0.072	-0.014	0.021	-0.054	0.075
	Baseline math score	0.094***	0.035	0.189***	0.010	-0.095***	0.036
	Baseline ELA score	0.352***	0.035	0.361***	0.010	-0.009	0.036
$\alpha^\theta_{e7\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	0.019	0.025	0.045***	0.006	-0.026	0.025
				Panel C: 8th grade			
$\alpha^0_{e8\ell}$	Constant	0.043	0.143	-0.344***	0.038	0.387***	0.148
$\alpha^x_{e8\ell}$	Female	0.146**	0.067	0.202***	0.021	-0.055	0.071
	Black	-0.016	0.097	-0.101***	0.034	0.085	0.102
	Hispanic	0.129	0.110	-0.009	0.036	0.137	0.116
	Subsidized lunch	0.064	0.079	-0.119***	0.031	0.183**	0.085
	Special education	-0.419***	0.085	-0.395***	0.024	-0.024	0.089
	Limited English proficiency	0.001	0.107	0.012	0.026	-0.011	0.110
	Baseline math score	0.080*	0.049	0.169***	0.013	-0.089*	0.050
	Baseline ELA score	0.272***	0.046	0.349***	0.013	-0.077	0.048
$\alpha^\theta_{e8\ell} \cdot \sigma_\theta$	Taste for charter schools (std. dv. units)	-0.018	0.031	0.085***	0.007	-0.104***	0.032

Notes: This table reports maximum simulated likelihood estimates of the parameters of the ELA achievement distribution. The constant in the charter school equation is the sum of the average of the school-specific effects and the relevant grade effect. Its standard error is computed using the delta method.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Maximum Simulated Likelihood Estimates of School-specific Parameters

School	Average utility (γ^0_k) (1)	Admission probability (π_k) (2)	Math effect (α^0_{mk}-α^0_{m60}) (3)	ELA effect (α^0_{ek}-α^0_{e60}) (4)
Charter school 1	-1.812***	0.544***	0.291***	-0.047
	(0.153)	(0.059)	(0.089)	(0.095)
Charter school 2	-1.603***	0.389***	0.275***	-0.004
	(0.153)	(0.052)	(0.082)	(0.096)
Charter school 3	-2.071***	0.642***	0.233**	-0.039
	(0.157)	(0.036)	(0.093)	(0.104)
Charter school 4	-2.554***	0.724***	0.346***	0.064
	(0.159)	(0.047)	(0.092)	(0.107)
Charter school 5	-0.880***	0.431***	0.329***	-0.178*
	(0.150)	(0.037)	(0.092)	(0.100)
Charter school 6	-3.308***	0.829***	0.467***	0.409***
	(0.172)	(0.049)	(0.102)	(0.118)
Charter school 7	-2.498***	0.871***	0.637***	0.461***
	(0.168)	(0.036)	(0.116)	(0.115)

Notes: This table reports maximum simulated likelihood estimates of the school-specific parameters from the structural model. The admission probabilities in column (2) are averages for 2006-2009. The school effects in column (3) correspond to 6th grade, as the 6th grade intercept is omitted from the model for charter achievement. Standard errors are in parentheses.

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

Table 12: Model Fit -- Choice Probabilities

	Application		Attendance	
	Data	Model	Data	Model
Choice	(1)	(2)	(3)	(4)
Apply/attend any charter	0.166	0.144	0.100	0.093
Among applicants/attenders:				
More than one charter	0.267	0.181	-	-
Charter school 1	0.179	0.145	0.154	0.115
Charter school 2	0.215	0.188	0.168	0.146
Charter school 3	0.283	0.225	0.170	0.204
Charter school 4	0.208	0.212	0.222	0.181
Charter school 5	0.282	0.240	0.128	0.191
Charter school 6	0.098	0.083	0.084	0.076
Charter school 7	0.093	0.089	0.076	0.087

Notes: This table compares empirical choice probabilities to simulated probabilities using the MSL estimates. Model statistics are produced by simulating the model 100 times for each observation in the sample, and then averaging over simulations and observations.

Table 13: Model fit -- Achievement Distributions

		Traditional public schools				Charter schools			
		Mean		Standard deviation		Mean		Standard deviation	
		Data	Model	Data	Model	Data	Model	Data	Model
Subject	Grade	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	6th	-0.56	-0.56	1.04	1.03	0.28	0.29	0.83	0.85
	7th	-0.50	-0.49	0.98	0.97	0.34	0.33	0.75	0.78
	8th	-0.46	-0.45	0.97	0.95	0.42	0.39	0.71	0.73
ELA	6th	-0.57	-0.56	1.07	1.06	-0.03	-0.03	0.92	0.94
	7th	-0.54	-0.56	0.99	0.99	0.10	0.07	0.83	0.86
	8th	-0.52	-0.53	0.98	0.97	0.17	0.12	0.79	0.81

Notes: This table compares empirical test score distributions to simulated distributions using the MSL estimates. Model statistics are produced by simulating the model 100 times for each observation in the sample, and then averaging over simulations and observations.

Table 14: Simulated Effects of Policy Changes -- Choice Behavior

Policy change	Fraction applying (1)	Fraction attending (2)	Avg. admission probability (3)	Fraction of seats filled (4)
None (7 charter schools)	0.147	0.087	0.632	1.000
Boston's planned expansion	0.223	0.148	0.799	0.854
(expand to 13 schools)	(51.4%)	(69.5%)	(26.4%)	(-14.6%)
Expand to 30 schools	0.331	0.239	0.912	0.598
	(125.0%)	(173.9%)	(44.3%)	(-40.2%)

Notes: This table reports simulated effects of expanding Boston's charter school network on charter application and attendance behavior. Numbers in parentheses are percentage changes relative to the existing policy environment. The predictions are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the structural sample.

Table 15: Simulated Effects of Policy Changes -- Math Scores

	6th grade			7th grade			8th grade		
Policy change	Boston average (1)	White-black gap (2)	TOT (3)	Boston average (4)	White-black gap (5)	TOT (6)	Boston average (7)	White-black gap (8)	TOT (9)
None (7 charter schools)	-0.507	0.651	0.327	-0.395	0.593	0.438	-0.403	0.565	0.537
All charter schools close	-0.534	0.671	-	-0.432	0.621	-	-0.448	0.602	-
	(-5.3%)	(3.1%)		(-9.5%)	(4.7%)		(-11.2%)	(6.7%)	
Boston's planned expansion	-0.482	0.628	0.360	-0.367	0.560	0.458	-0.371	0.535	0.559
(expand to 13 schools)	(4.9%)	(-3.6%)	(10.0%)	(7.2%)	(-5.6%)	(4.6%)	(7.9%)	(-5.2%)	(4.0%)
Expand to 30 schools	-0.450	0.610	0.397	-0.329	0.522	0.482	-0.321	0.488	0.583
	(11.3%)	(-6.3%)	(21.6%)	(16.8%)	(-12.0%)	(9.9%)	(20.2%)	(-13.6%)	(8.5%)
All students forced to attend	0.060	0.415	0.591	0.175	0.244	0.604	0.267	0.165	0.714
charter schools	(111.7%)	(-36.3%)	(81.0%)	(144.2%)	(-58.8%)	(37.7%)	(166.4%)	(-70.8%)	(32.8%)

Notes: This table reports simulated effects of modifying Boston's charter school network on the citywide math score distribution. Numbers in parentheses are percentage changes relative to the existing policy environment. Predictions are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the sample, and then averaging over the simulations.

Table 16: Simulated Effects of Policy Changes -- ELA Scores

	6th grade			7th grade			8th grade		
Policy change	Boston average (1)	White-black gap (2)	TOT (3)	Boston average (4)	White-black gap (5)	TOT (6)	Boston average (7)	White-black gap (8)	TOT (9)
None (7 charter schools)	-0.465	0.562	0.132	-0.485	0.454	0.581	-0.435	0.428	0.257
All charter schools close	-0.466	0.584	-	-0.523	0.476	-	-0.446	0.457	-
	(-0.2%)	(3.8%)		(-7.8%)	(4.7%)		(-2.4%)	(6.8%)	
Boston's planned expansion (expand to 13 schools)	-0.462	0.547	0.160	-0.457	0.440	0.598	-0.425	0.420	0.311
	(0.7%)	(-2.7%)	(21.4%)	(5.6%)	(-3.2%)	(2.9%)	(2.4%)	(-1.9%)	(20.8%)
Expand to 30 schools	-0.473	0.528	0.193	-0.430	0.407	0.616	-0.412	0.382	0.371
	(-1.7%)	(-6.2%)	(47.1%)	(11.3%)	(-10.4%)	(6.0%)	(5.4%)	(-10.8%)	(44.4%)
All students forced to attend charter schools	-0.204	0.347	0.362	0.092	0.222	0.714	0.139	0.190	0.684
	(56.2%)	(-38.3%)	(175.2%)	(119.1%)	(-51.0%)	(22.8%)	(132.0%)	(-55.6%)	(166.0%)

Notes: This table reports simulated effects of modifying Boston's charter school network on the citywide ELA score distribution. Numbers in parentheses are percentage changes relative to the existing policy environment. Predictions are produced by simulating the model 100 times for each of the 2,485 students in the 2009 cohort of the structural sample, and then averaging over the simulations.

References

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

[2] Abadie, A. (2003). "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113(2).

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

[4] Ajayi (2011). "A Welfare Analysis of School Choice Reforms in Ghana." Mimeo, Boston College.

[5] Angrist, J., Pathak, P., and Walters, C. (2011). "Explaining Charter School Effectiveness." NBER Working Paper no. 17332.

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

[7] Angrist, J., and Fernandez-Val, I. (2011). "ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework." NBER Working Paper no. 16566.

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

[9] Attanasio, O., Meghir, C., and Santiago, A. (2011). "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *Review of Economic Studies* 79.

[10] Barnett, W. (1992). "Benefits of Compensatory Preschool Education." *The Journal of Human Resources* 27(2).

[11] Berkovec, J., and Stern, S. (1991). "Job Exit Behavior of Older Men." *Econometrica* 59(1).

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

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

[14] Card, D., and Hyslop, R. (2005). "Estimating the Effects of a Time-Limited Earnings Subsidy for Welfare-Leavers." *Econometrica* 73(6).

[15] Carneiro, P., Heckman, J., and Vytlacil, E. (2010). "Estimating Marginal Returns to Education." IZA Discussion Paper no. 5275.

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

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

[18] Dobbie, W., and Fryer, R. (2011a). "Are High Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics* 3(3).

[19] Dobbie, W., and Fryer, R. (2011b). "Getting Beneath the Veil of Effective Schools: Evidence from New York City." NBER Working Paper no. 17632.

[20] Duflo, E., Hanna, R., and Ryan, S. (2012). "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102(4).

[21] Ferguson, R. (1998). "Can Schools Narrow the Black-White Test Score Gap?" In: Jencks, C., and Phillips, M., eds., The Black-White Test Score Gap. Washington, DC: Brookings Institution Press.

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

[23] Finkelstein and McGarry (2006). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." *American Economic Review* 96(4).

[24] Fryer, R. (2010). "Racial Inequality in the 21st Century: The Declining Significance of Discrimination." NBER Working Paper no. 16256.

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

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

[27] Fryer, R., and Levitt, S. (2004). "Understanding the Black-White Test Score Gap in the First Two Years of School." *Review of Economics and Statistics* 86(2).

[28] Fryer, R., and Levitt, S. (2006). "The Black-White Test Score Gap through Third Grade." *American Law and Economics Review* 8(2).

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

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

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

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

[33] Heckman, J. (2010). "Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy." *Journal of Economic Literature* 48(2).

[34] Heckman, J. (2011). "The American Family in Black and White: A Post-Racial Strategy for Improving Skills to Promote Equality." IZA Discussion Paper no. 5495.

[35] Heckman, J., and Vytlacil, E. (2001). "Policy-Relevant Treatment Effects." *The American Economic Review* 91(2).

[36] Heckman, J., and Vytlacil, E. (2005). "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73(3).

[37] Howell, W., and Peterson, P. (2002). The Education Gap: Vouchers and Urban Schools. Washington, DC: Brookings Institution Press.

[38] Hoxby, C., and Murarka, S. (2009). "Charter Schools in New York City: Who Enrolls and How They Affect Their Students' Achievement." NBER Working Paper no. 14852.

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

[40] Hotz, V., Imbens, G., and Mortimer, J. (2005). "Predicting the Efficacy of Future Training Programs Using Past Experiences at Other Locations." Journal of Econometrics 125(1).

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

[42] Imbens, G., and Rubin, D. (1997). "Estimating Outcome Distributions for Compliers in Instrumental Variables Models." *The Review of Economic Studies* 64(4).

[43] Krueger, A., and Whitmore, D. (2001). "Would Smaller Classes Help Close the Black-White Achievement Gap?" In: Chubb, J. and Loveless, T., eds., Bridging the Achievement Gap. Washington, DC: Brookings Institution Press.

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

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

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

[47] Massachusetts Department of Elementary and Secondary Education (2012c). "Opening Procedures Handbook: A Guide for Boards of Trustees and Leaders of New Charter Schools." http://www.doe.mass.edu/charter/guides/ophandbook.pdf.

[48] McFadden, D. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57(5).

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

[50] Neal, D. (2006). "Why has Black-White Skill Convergence Stopped?" In: Hanushek, E. and Welch, F., eds., Handbook of the Economics of Education, Volume 1. Amsterdam: Elsvier.

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

[52] Newey, W., and McFadden, D. (1986). "Large Sample Estimation and Hypothesis Testing." In: Engle, R. and McFadden, D., eds., Handbook of Econometrics. Amsterdam: Elsevier.

[53] Ravitch, D. (2010). The Death and Life of the Great American School System: How Testing and Choice are Undermining Education. New York: Basic Books.

[54] Reardon, S. (2011). "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations." In: Duncan, J., and Murnane, R., eds., Whither Opportunity? New York: Russell Sage Foundation.

[55] Rothstein, R. (2004). Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap. Washington, DC: Economic Policy Institute.

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

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

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

[59] Todd, P., and Wolpin, K. (2006). "Using Experimental Data to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico." *American Economic Review* 96(5).

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

[61] Vanneman, A., Hamilton, L., Anderson, J., and Rahman, T. (2009). "Achievement Gaps: How Black and White Students in Public Schools Perform in Mathematics and Reading on the National Assessment of Educational Progress." US Department of Education, NCES 2009-455.

[62] Vaznis, J. (2012). "Charter School Moratorium Lifted by State." *The Boston Globe*, May 8.

[63] Wald, A. (1940). "The Fitting of Straight Lines if Both Variables are Subject to Error." *The Annals of Mathematical Statistics* 11(3).

[64] Walker, J., Ben-Akiva, M., and Bolduc, D. (2004). "Identification of the Logit Kernel (or Mixed Logit) Model." Mimeo, MIT.

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

Appendix A: Complier Densities

This appendix shows that 2SLS estimation of equation (4) produces consistent estimates of potential outcome densities for lottery compliers. This result is an extension of the methods developed in Abadie (2002). Abadie's Lemma 2.1 implies that the 2SLS estimate of $\gamma(y)$ is a consistent estimator of the expectation of $K_h(y-Y_i(1))$ for compliers as $N \to \infty$. I show that as $h \to 0$ and $Nh \to \infty$, this estimator converges in probability to the complier density function. Imbens and Rubin (1997) outline an alternative method for estimating complier densities based on linear combinations of empirical densities for the four possible combinations of Z_i and S_i. The approach taken here allows these densities to be estimated in a simple one-step IV procedure.

Let $f_s^c(y)$, $f_s^{at}(y)$, and $f_s^{nt}(y)$ be the density functions of $Y_i(s)$ for compliers, always takers, and never takers, respectively, with $s \in \{0,1\}$. Define $K_h(u) \equiv \frac{1}{h}K\left(\frac{u}{h}\right)$, where $K(\cdot)$ is a function that satisfies $\int K(u)du = 1$, $\int uK(u)du = 0$, $\int u^2K(u)du < \infty$, and $\int K^2(u)du < \infty$. Consider the equation

$$K_h(y - Y_i) \cdot 1\{S_i = s\} = \alpha_s(y) + \gamma_s(y) \cdot 1\{S_i = s\} + \epsilon_{isy}$$

for $s \in \{0,1\}$. If Z_i is used as an instrument for $1\{S_i = s\}$ in this equation, the resulting IV estimator is

$$\hat{\gamma}_s(y) \equiv \frac{E_N[K_h(y - Y_i) \cdot 1\{S_i = s\}|Z_i = 1] - E_N[K_h(y - Y_i) \cdot 1\{S_i = s\}|Z_i = 0]}{E_N[1\{S_i = s\}|Z_i = 1] - E_N[1\{S_i = s\}|Z_i = 0]} \tag{13}$$

where $E_N[\cdot]$ is the empirical expectation operator. The following theorem shows that this estimator is consistent for $f_s^c(y)$.

Theorem: *Suppose that assumptions A1-A3 hold, and that the density functions $f_s^c(y)$, $f_1^{at}(y)$, and $f_0^{nt}(y)$ exist and are twice differentiable at y. Then*

$$\operatorname*{plim}_{h \to 0, Nh \to \infty} \hat{\gamma}_s(y) = f_s^c(y)$$

Proof: I demonstrate the result for $s = 1$. The proof for $s = 0$ is analogous. I begin by considering the expectation and variance of the each term in the numerator of (13). Define

$$\hat{E}_z(y) = E_N[K_h(y - Y_i) \cdot 1\{S_i = 1\}|Z_i = z]$$

$\hat{E}_z$ is a sample average, so it is unbiased for the corresponding population moment. For $z = 1$, we have

$$\begin{aligned} E\left[\hat{E}_1(y)\right] &= E[K_h(y - Y_i) \cdot 1\{S_i = 1\}|Z_i = 1] \\ &= E[K_h(y - Y_i(1))|S_i(1) > S_i(0)] \cdot Pr[S_i(1) > S_i(0)] \\ &+ E[K_h(y - Y_i(1))|S_i(1) = S_i(0) = 1] \cdot Pr[S_i(1) = S_i(0) = 1] \end{aligned}$$

which can be written

$$E\left[\hat{E}_1(y)\right] = \int K_h(y-t)\cdot\left(\phi^c f_1^c(t) + \phi^{at} f_1^{at}(t)\right) dt$$

where ϕ^c and ϕ^{at} are the fractions of compliers and always takers, respectively. Using the change of variables $u = \frac{y-t}{h}$ and taking a second-order Taylor expansion around $h = 0$ yields

$$E\left[\hat{E}_1(y)\right] = \phi^c f_1^c(y) + \phi^{at} f_1^{at}(y) + \frac{h^2}{2}\cdot\left(\phi^c f_1^{c''}(y) + \phi^{at} f_1^{at''}(y)\right)\cdot\int K(u)u^2du + o(h^2)$$

which implies

$$\lim_{h\to 0} E[\hat{E}_1(y)] = \phi^c f_1^c(y) + \phi^{at} f_1^{at}(y).$$

A similar argument shows that

$$\lim_{h\to 0} E[\hat{E}_0(y)] = \phi^{at} f_1^{at}(y).$$

Next, consider the variance of $\hat{E}_1(y)$. We have

$$Var\left(\hat{E}_1(y)\right) = \tfrac{1}{N_1}E[K_h^2(y-Y_i)\cdot 1\{S_i=1\}|Z_i=1] - \tfrac{1}{N_1}E[\hat{E}_1(y)]^2$$

where N_1 is the number of observations with $Z_i = 1$. The argument above shows that $E[\hat{E}_1(y)]$ is bounded as $h \to 0$, so as $N_1 \to \infty$ (which is implied by $N \to \infty$ together with (A2)) the second term is negligible. The first term is

$$\begin{aligned}\tfrac{1}{N_1}E[K_h^2(y-Y_i)\cdot 1\{S_i=1\}|Z_i=1] &= \tfrac{1}{N_1}\int K_h^2(y-t)\cdot\left(\phi^c f_1^c(t) + \phi^{at} f_1^{at}(t)\right)dt \\ &= \tfrac{1}{N_1 h}\cdot\left(\phi^c f_1^c(y) + \phi^{at} f_1^{at}(y)\right)\cdot\int K^2(u)du + o\left(\tfrac{1}{Nh}\right).\end{aligned}$$

Therefore, we have

$$\lim_{h\to 0, Nh\to\infty} Var\left(\hat{E}_1(y)\right) = 0.$$

A similar calculation shows that the variance of $\hat{E}_0(y)$ also converges to zero.

The arguments so far imply that as $h \to 0$ and $Nh \to \infty$, $\hat{E}_1(y)$ and $\hat{E}_0(y)$ converge in mean square, and therefore in probability, to $\left(\phi^c f_1^c(y) + \phi^{at} f_1^{at}(y)\right)$ and $\phi^{at} f_1^{at}(y)$, respectively. When $s = 1$, the probability limit of the denominator of (13) as $N \to \infty$ is ϕ^c. Then by the continuous mapping theorem we have

$$\begin{aligned}\operatorname*{plim}_{h\to 0, Nh\to\infty} \hat{\gamma}_1(y) &= \frac{\phi^c f_1^c(y) + \phi^{at} f_1^{at}(y) - \phi^{at} f_1^{at}(y)}{\phi_c} \\ &= f_1^c(y).\end{aligned}$$

This completes the proof.

Appendix B: Identification of Preference Coefficients

This appendix uses a simplified version of the structural model to show analytically that the combination of lottery and distance instruments identifies the coefficients on the charter preference θ_i in equations (7) and (8). Suppose there is a single charter school, and the utilities of charter and public school attendance are given by

$$U_{i1} = \gamma^0 + \gamma^d \cdot D_i + \theta_i + v_i - \gamma^a \cdot A_i$$
$$U_{i0} = -\gamma^a \cdot A_i$$

where D_i is distance to the charter school, A_i indicates charter application, $\theta_i \sim N(0, \sigma_\theta^2)$ is observed prior to the application decision, and $v_i \sim N(0,1)$ is observed after the application decision.[37] The charter school holds a lottery for applicants with acceptance probability π.

The expected utility of applying to the charter school is

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

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

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

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

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

where

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

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

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

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

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

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

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

The model for potential outcomes in charter and public school is

$$Y_i(1) = \alpha_1^0 + \alpha_1^\theta \cdot \theta_i + \epsilon_{i1}$$
$$Y_i(0) = \alpha_0^0 + \alpha_0^\theta \cdot \theta_i + \epsilon_{i0}$$

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

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

where

$$\begin{aligned}\mu_\theta^c(D) = &\sigma_\theta \cdot \Phi\left(\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right) \cdot \lambda\left(\frac{\theta^*(D)}{\sigma_\theta}\right) \\ &+ \sigma_\theta \cdot (1 - \Phi\left(\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right)) \cdot \int \lambda\left(\frac{-\gamma^0 - \gamma^d \cdot D - v_i}{\sigma_\theta}\right) dF\left(v_i | v_i < -\psi^{-1}\left(\frac{\gamma^a}{\pi}\right)\right).\end{aligned}$$

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

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

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

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

Appendix C: Relationship to Roy Model

This appendix shows that equations (5), (7), and (8) nest a Roy model of selection in which students seek to maximize achievement and have private information about their test scores in charter and public schools. For simplicity, I omit application costs and preferences for distance, and focus on scores in a single subject and grade. Achievement for student i at charter school k is given by

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

while public school achievement is

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

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

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

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

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

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

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

where

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

$$\gamma^x = \alpha_c^x - \alpha_p^x$$

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

and $U_{i0} \equiv 0$. These preferences are a special case of equation (5) with $\gamma^d = \gamma^a = 0$ and $Var(v_{ik}) = 0$.

Returning to the test score equation, we have

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

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

where

$$\alpha_c^\theta = \frac{\sigma_c^2 - \sigma_{cp}}{\sigma_c^2 + \sigma_p^2 - 2\sigma_{cp}}$$
$$\alpha_p^\theta = \frac{\sigma_{cp} - \sigma_p^2}{\sigma_c^2 + \sigma_p^2 - 2\sigma_{cp}}$$

This implies that potential test scores are given by

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

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

where $E[\epsilon_{ik}|X_i, \theta_i] = 0$, which is the specification for achievement in equations (7) and (8).

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

Appendix D: Equilibrium Admission Probabilities

Description of the Game

This appendix describes the determination of equilibrium admission probabilities in counterfactual simulations. These probabilities are determined in a Subgame Perfect Nash Equilibrium (SPE) in which students make utility-maximizing choices as described in Section 4, and schools set admission probabilities to maximize enrollment subject to capacity constraints. Enrollment at school k is given by

$$e_k = E\left[1\{S_i = k\}\right].$$

Q_k denotes school k's capacity, which is its maximum potential enrollment. In the simulations, I set capacities for existing schools equal to their average enrollment in years when they were oversubscribed. Capacity for new schools is set equal to average capacity for existing schools.

The timing of the game follows Figure 4. Strategies in each stage of the game are as follows:

1. Students first choose application vectors.
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, note that the distribution of students is atomless, so schools do not change their admission probabilities in the second stage in response to the application decisions of individual students in the first stage. Students therefore act as "price takers" in the first stage, in the sense that they do not expect schools to react to their application choices. Without loss of generality, therefore, 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 vector A_i^* for each student, a vector of admission probabilities π^*, and a rule for assigning school choices that satisfy the following conditions:

1. Student application choices are given by $A_i^* = A^*(X_i, D_i, \theta_i, \tau_i; \pi^*)$, where A^* is defined as in Section 4.2 and now explicitly depends on the vector of admission probabilities students expect to face in each lottery.

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

3. After receiving the offer vector z, student i chooses school k with probability $P_{ik}(z, \theta_i, \tau_i)$ as in Section 4.2.

School Best Response Functions

I begin by deriving a school's optimal admissions probability as a function of students' expected admission probabilities and the actions of other schools. 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 given by $A^*(X_i, D_i, \theta_i, \tau_i; \pi^e)$. In addition, suppose that schools other than k admit students with probability π_{-k}. If school k admits students with probability π_k in the second stage, its enrollment is given by

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

School k chooses π_k to solve

$$\max_{\pi_k\in[0,1]} e_k(\pi_k, \pi_{-k}, \pi^e) \quad s.t. \quad e_k(\pi_k, \pi_{-k}, \pi^e) \leq Q_k \tag{14}$$

The best response function $\pi^{BR}_k(\pi_{-k}, \pi^e)$ is the solution to problem (14). The optimal admission probability sets school k's enrollment equal to its capacity. The following equation implicitly defines π^{BR}_k at interior solutions:

$$E\left[\sum_{z\in\{0,1\}^K} f\left(z|A^*(X_i, D_i, \theta_i, \tau_i; \pi^e); \pi^{BR}_k, \pi_{-k}\right) P_{ik}(z, \theta_i, \tau_i)\right] = Q_k$$

This equation can be rewritten

$$E\left[\sum_{z_{-k}} f_{-k}\left(z_{-k}|A^*_{-k}(X_i, D_i, \theta_i, \tau_i; \pi^e); \pi_{-k}\right) \cdot f_k\left(1|A^*_k(X_i, D_i, \theta_i, \tau_i; \pi^e); \pi^{BR}_k\right) \cdot \tilde{P}_{ik}(1, z_{-k})\right] = Q_k$$

where the sum is over all values of z_{-k} in $\{0,1\}^{K-1}$. Here $\tilde{P}_{ik}(z_k, z_{-k})$ is $P_{ik}(z)$ with the k-th element of z set to z_k and the remaining elements set to z_{-k}, and I have used the fact that $\tilde{P}_{ik}(0, z_{-k}) = 0$ since school k is not in student i's choice set if she does not receive an offer. Substituting in f_k yields

$$E\left[\sum_{z_{-k}} f_{-k}\left(z_{-k}|A^*_{-k}(X_i, D_i, \theta_i, \tau_i; \pi^e); \pi_{-k}\right) \cdot A^*_k(X_i, D_i, \theta_i, \tau_i; \pi^e) \cdot \pi^{BR}_k \cdot \tilde{P}_{ik}(1, z_{-k})\right] = Q_k$$

which can be solved for π_k^{BR}:

$$\pi_k^{BR} = \frac{Q_k}{E[A_k^*\left(X_i, D_i, \theta_i, \tau_i; \pi^e\right) \cdot \sum_{z_{-k}} f_{-k}\left(z_{-k} | A_{-k}^*\left(X_i, D_i, \theta_i, \tau_i; \pi^e\right); \pi_{-k}\right) \tilde{P}_{ik}(1, z_{-k})]}$$
$$\equiv \Gamma_k(\pi_{-k}, \pi^e)$$

If the denominator of Γ_k is sufficiently small, it may exceed one, in which case school k cannot fill its capacity. In this case, the optimal action is to set $\pi_k = 1$ and fill as many seats as possible. Furthermore, note that the denominator of Γ_k is zero when the measure of students applying to school k is zero, in which case this function is undefined. If no students apply, school k's enrollment does not depend on its own admission probability, so every point in the unit interval is a best response. These arguments imply that the best response function is given by

$$\pi_k^{BR}\left(\pi_{-k}, \pi^e\right) = \begin{cases} \min\{\Gamma_k\left(\pi_{-k}, \pi^e\right), 1\}, & E\left[A_k^*(X_i, D_i, \theta_i, \tau_i; \pi^e)\right] > 0 \\ [0, 1], & E\left[A_k^*(X_i, D_i, \theta_i, \tau_i; \pi^e)\right] = 0 \end{cases}$$

Existence of Equilibrium

Define $\pi^{BR}(\pi) \equiv \left(\pi_1^{BR}(\pi_{-1}, \pi), ..., \pi_K^{BR}(\pi_{-K}, \pi)\right)'$. A vector π^* supports a Nash equilibrium if and only if it is a fixed point of $\pi^{BR}(\pi)$:

$$\pi^* \in \pi^{BR}\left(\pi^*\right) \tag{15}$$

A Nash equilibrium always exists, though it is not unique. Existence can be proved constructively by noting that the vector 0^K always supports an equilibrium. If students expect not to be admitted to any school, none will apply, and it is a best response for each school to set $\pi_k = 0$. This equilibrium is an artifact of the assumption that the distribution of students is atomless. If an individual students had positive mass, she could deviate and apply to a charter school, and the school would strictly prefer to raise its admission probability and admit her, breaking the equilibrium. The no-admission equilibrium is not economically interesting, and other equilibria with positive admission probabilities generally exist. In the counterfactual simulations, I compute equilibria by numerically solving the system of nonlinear equations implied by (15), restricting the π_k to be strictly positive.[38] I never found more than one equilibrium in any simulation. Table A4 lists the equilibrium admission probabilities used to simulate the effects of charter expansion.

[38] For expansions involving many charter schools, the student choice set becomes very large and solving this system becomes infeasible. For example, with 30 schools, there are $2^{30} = 1,073,741,824$ possible application choices. To limit the number of choices, I assume that no student would apply to more than two charter schools when computing the equilibrium probabilities and simulating counterfactuals. Less than 1 percent of students applied to more than two schools in the data.

Uniqueness of Non-trivial Equilibrium

As noted above, the equilibrium of this game is not unique, as there are many trivial equilibria. For example, $\pi = 0^K$ is always an equilibrium, there are K equilibia where $\pi_k > 0$ and $\pi_{-k} = 0^{K-1}$, and there are typically other similar equilibria where admission probabilities are zero for one or more schools. However, I ignore all equilibria with $\pi_k = 0$ for any k, and there may exist a unique equilibrium where each school has a positive measure of applicants. Define the function

$$\psi_k(\pi) \equiv \pi_k - \Gamma_k(\pi_{-k}, \pi)$$

Let $\psi(\pi) = (\psi_1(\pi), ..., \psi_K(\pi))'$. A vector $\pi^* \in (0,1)^K$ is a non-trivial, interior equilibrium if $\psi(\pi^*) = 0$. Moreover, π^* is the only such equilibrium if the Jacobean of $\psi(\cdot)$ is a positive dominant diagonal matrix. This requires the following two conditions:

$$1. \frac{\partial \psi_k}{\partial \pi_k} > 0 \quad \forall k$$
$$2. \left|\frac{\partial \psi_k}{\partial \pi_k}\right| \geq \sum_{j \neq k} \left|\frac{\partial \psi_k}{\partial \pi_j}\right| \quad \forall k$$

These conditions provide intuition for when a unique equilibrium is more likely. The first condition requires that $\frac{\partial \Gamma_k}{\partial \pi_k} < 1$. This is always satisfied: Γ_k is decreasing in π_k. As students' expected admission probability at school k rises, more choose to apply, and it is optimal for school k to lower π_k to maintain enrollment at capacity. The second condition requires the effect of π_k on ψ_k to dominate the effects of admission probabilities at other schools. This condition is more likely to hold when there is little substitution across charter schools. For example, if the pools of potential applicants at each charter school are disjoint sets, then $\frac{\partial \psi_k}{\partial \pi_j} = 0$ for $k \neq j$ and the second condition is satisfied. More generally, this condition is likely to be satisfied when few students apply to multiple charter schools, which is more likely when charter schools are more geographically disbursed, or preferences for spatial proximity are stronger.

Table A1: Covariate Balance for the Lottery Sample

Variable	All (1)	White (2)	Black (3)	Hispanic (4)
Black	0.008	-	-	-
	(0.029)			
Hispanic	0.015	-	-	-
	(0.027)			
Female	-0.042	0.020	-0.109**	0.039
	(0.031)	(0.077)	(0.043)	(0.058)
Free/reduced lunch	0.015	0.032	0.043	-0.053
	(0.027)	(0.068)	(0.035)	(0.047)
Special education	0.005	-0.011	0.022	-0.011
	(0.024)	(0.060)	(0.034)	(0.045)
Limited English proficiency	-0.001	0.025	0.040*	-0.093*
	(0.020)	(0.021)	(0.022)	(0.050)
Baseline math score	-0.066	-0.104	-0.074	0.076
	(0.061)	(0.105)	(0.091)	(0.112)
Baseline ELA score	-0.010	-0.033	-0.073	0.169
	(0.063)	(0.110)	(0.091)	(0.120)
p -value from joint test	0.755	0.945	0.117	0.567
N	1822	317	948	557

Notes: This table reports coefficients from regressions of baseline student characteristics on an offer dummy and lottery fixed effects. *P* -values are from joint tests of the hypothesis that the offer variable has a zero coefficient in all regressions. Robust standard errors are in parentheses.

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

Table A2: Attrition

		All students	Lottery applicants	
		Follow-up rate	Follow-up rate	Differential
Race		(1)	(2)	(3)
All		0.863	0.839	-0.002
				(0.020)
	N	27881	4579	
White		0.853	0.845	0.020
				(0.052)
	N	4067	774	
Black		0.865	0.830	-0.006
				(0.027)
	N	13020	2431	
Hispanic		0.864	0.853	0.000
				(0.038)
	N	10794	1374	

Notes: This table reports the fraction of follow-up test scores that are observed for students attending 4th grade in Boston between 2006 and 2009. Column (1) shows the follow-up rate for the full sample. Column (2) shows the follow-up rate for lottery applicants, while column (3) shows the difference in follow-up rates for lottery winners and losers. This differential is computed from a regression that controls for lottery fixed effects. Standard errors are robust to heteroskedasticity and are clustered at the student level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A3: Maximum Simulated Likelihood Estimates of Covariance Parameters

		Charter school		Traditional public school		Charter effect	
		Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
Parameter	Description	(1)	(2)	(3)	(4)	(5)	(6)
$\rho_{m\ell}$	Serial correlation of math scores	0.589***	0.016	0.679***	0.005	-0.089***	0.017
$\rho_{e\ell}$	Serial correlation of ELA scores	0.468***	0.021	0.589***	0.006	-0.121***	0.022
$Var(\zeta_{imgk})^{1/2}$	Standard deviation of math shocks	0.568***	0.009	0.658***	0.004	-0.090***	0.010
$Var(\zeta_{iegk})^{1/2}$	Standard deviation of ELA shocks	0.628***	0.009	0.681***	0.004	-0.053***	0.010
$Corr(\zeta_{imgk}, \zeta_{iegk})$	Correlation between math and ELA shocks	0.386***	0.018	0.435***	0.006	-0.049***	0.019

Notes: This table reports maximum simulated likelihood estimates of the covariance parameters of the math and ELA achievement distributions.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table A4: Equilibrium Admission Probabilities

	No change	Planned expansion	30 schools
	(1)	(2)	(3)
School 1	0.584	0.813	1.000
School 2	0.450	0.604	0.847
School 3	0.548	0.740	1.000
School 4	0.759	1.000	1.000
School 5	0.383	0.505	0.703
School 6	0.907	1.000	1.000
School 7	0.793	1.000	1.000
School 8	-	1.000	1.000
School 9	-	0.367	0.537
School 10	-	1.000	1.000
School 11	-	0.390	0.560
School 12	-	1.000	1.000
School 13	-	0.969	1.000
School 14	-	-	1.000
School 15	-	-	1.000
School 16	-	-	1.000
School 17	-	-	1.000
School 18	-	-	1.000
School 19	-	-	0.574
School 20	-	-	1.000
School 21	-	-	1.000
School 22	-	-	1.000
School 23	-	-	0.820
School 24	-	-	0.692
School 25	-	-	1.000
School 26	-	-	0.645
School 27	-	-	1.000
School 28	-	-	1.000
School 29	-	-	1.000
School 30	-	-	0.986

Notes: This table shows equilibrium admission probabilities for the counterfactual simulations. The procedure for determining these probabilities is described in Appendix B.