

NBER WORKING PAPER SERIES

ACCOUNTABILITY AND FLEXIBILITY IN PUBLIC SCHOOLS:
EVIDENCE FROM BOSTON'S CHARTERS AND PILOTS

Atila Abdulkadiroglu
Joshua Angrist
Susan Dynarski
Thomas J. Kane
Parag Pathak

Working Paper 15549
<http://www.nber.org/papers/w15549>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2009

The authors are grateful to the Boston Foundation for financial support and to Boston's charter schools, the Boston Public Schools, and Carrie Conaway, Mary Street, and the staff of the Massachusetts Department of Elementary and Secondary Education for data and assistance. Sarah Cohodes, Jon Fullerton and the staff of the Center for Education Policy Research at Harvard University played a critical role in this project. We also thank Bruno Ferman, Yuqiao Huang, Simone Schaner, Chris Walters, and Victor Vazquez for expert research assistance. We benefited from comments by Derek Neal and seminar participants at the Bank of Italy, the LSE, McGill, the New York Fed, UBC, UC Dublin, UCLA, UC Riverside, UIUC, ZEW, and the November 2009 New Directions in the Economic Analysis of Education Conference at the University of Chicago. Abdulkadiroglu acknowledges an NSF-CAREER award. Pathak thanks the Rappaport Institute for Greater Boston for research support. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2009 by Atila Abdulkadiroglu, Joshua Angrist, Susan Dynarski, Thomas J. Kane, and Parag Pathak. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots
Atila Abdulkadiroglu, Joshua Angrist, Susan Dynarski, Thomas J. Kane, and Parag Pathak
NBER Working Paper No. 15549
November 2009
JEL No. H52,I21,I28,J24

ABSTRACT

Charter schools are publicly funded but operate outside the regulatory framework and collective bargaining agreements characteristic of traditional public schools. In return for this freedom, charter schools are subject to heightened accountability. This paper estimates the impact of charter school attendance on student achievement using data from Boston, where charter schools enroll a growing share of students. We also evaluate an alternative to the charter model, Boston's pilot schools. These schools have some of the independence of charter schools, but operate within the school district, face little risk of closure, and are covered by many of same collective bargaining provisions as traditional public schools. Estimates using student assignment lotteries show large and significant test score gains for charter lottery winners in middle and high school. In contrast, lottery-based estimates for pilot schools are small and mostly insignificant. The large positive lottery-based estimates for charter schools are similar to estimates constructed using statistical controls in the same sample, but larger than those using statistical controls in a wider sample of schools. The latter are still substantial, however. The estimates for pilot schools are smaller and more variable than those for charters, with some significant negative effects.

Atila Abdulkadiroglu
Duke University
Department of Economics
Durham, NC 27708
atila.abdulkadiroglu@duke.edu

Joshua Angrist
Department of Economics
MIT, E52-353
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
angrist@mit.edu

Susan Dynarski
University of Michigan
Weill Hall
735 South State Street
Ann Arbor, MI 48109-3091
and NBER
dynarski@umich.edu

Thomas J. Kane
Harvard Graduate School of Education
Gutman Library, Room 455
Appian Way
Cambridge, MA 02138
and NBER
kaneto@gse.harvard.edu

Parag Pathak
Assistant Professor of Economics
E52-391C
MIT Department of Economics
Cambridge, MA 02142
and NBER
ppathak@mit.edu

I. Introduction

Charter schools operate with considerably more independence than traditional public schools. They are free to structure their curriculum and school environment; for instance, many charter schools fit more instructional hours into a year by running longer school days and providing instruction on weekends and during the summer (Matthews 2009, Wilson 2008, Hoxby et al, 2009). Because few charter schools are unionized, they can hire and fire teachers and administrative staff without regard to the collectively bargained seniority and tenure provisions that constrain such decisions in most public schools. Although charter students made up only 2.9 percent of U.S. public school enrollment in 2008-2009, charter enrollment has grown rapidly and seems likely to accelerate in the near future (NAPCS 2009). The growth of charter schools is an important component of the contemporary education reform movement's pursuit of accountability and flexibility in public education.

Proponents see charter schools' freedom from regulation as a source of educational innovation, with the added benefit of providing a source of competition that may prompt innovation and improvement in the rest of the public system. At the same time, charter schools are controversial because, after a transition period in which the state provides subsidies, they receive a tuition payment for each enrolled student paid by students' home (or "sending") districts. In Massachusetts, the site of our study, tuition payments are determined largely by the average per-pupil expenditure in sending districts. Not surprisingly, therefore, public school districts are concerned about the revenue lost when their students enroll in charter schools.

The purpose of this paper is to assess the causal effects of charter school attendance and a closely related alternative, called pilot schools, on student achievement. Pilot schools arose in Boston as a union-supported alternative to charter schools.¹ Boston's charter schools are legally constituted by the state as individual school districts and therefore operate independently of the Boston Public Schools (BPS). In contrast, Boston's pilot schools are legally part of the BPS district, and the extent to which they operate

¹ See Center for Collaborative Education (2006).

outside collective bargaining provisions is spelled out in school-specific election-to-work agreements signed by pilot faculty. In addition to these negotiated exemptions, pilot schools have generally greater flexibility and decision-making powers over school budgets, academic programs, and educational policies than do traditional BPS schools. This includes freedom with regard to policies related to student promotion, graduation, discipline, and attendance.²

In practice, pilot schools occupy a middle ground between charter schools and traditional public schools. Their teachers are part of the Boston Teachers Union (BTU), with their pay, benefits and working hours determined by the district-wide BTU contract. On the other hand, pilot schools can set their own policies with regard to curriculum, student promotion, and graduation. They also fit more instructional hours into a school year than traditional schools, but not as many as the charter schools. Accountability standards appear to bind less strictly for pilot schools than for charter schools: while nine Massachusetts charters have been lost, no pilot school has been closed.

The charter and pilot school models fit into a broader movement towards choice and decentralization in public education. Related policy experiments that have also been the subject of impact evaluations include vouchers (Rouse 1998; Angrist, Bettinger, Bloom, King and Kremer 2002), magnet schools (Cullen, Jacob, and Levitt 2006), and intra-district choice plans (Abdulkadiroglu and Sonmez 2003; Hastings, Kane and Staiger 2008; Abdulkadiroglu, Pathak and Roth 2009). Charter schools appear to be one of the most widely replicated members of this reform family.

The schools in our study are attended by students who would otherwise attend traditional Boston public schools. The Boston Public Schools (BPS) system serves a disproportionately Black and Hispanic student population. Like students in many urban schools, BPS students have lower test scores and lower rates of high school graduation and college attendance than students from nearby suburban districts.

² See the Boston Teachers Union website (http://www.btu.org/leftnavbar/HP_PilotSchools.htm), which also notes: "Pilot schools do not have to purchase a variety of services provided by the central office, such as substitute teachers, textbook, SPED contracted services, and academic coaches. By not purchasing these services pilot schools 'save' , typically, \$300 to \$400 per year per student. They are allowed to retain these funds and purchase these services privately if they wish."

Boston's charter schools also serve a high proportion of Black students, even relative to the majority non-white BPS district. The effects of charter schools in urban populations are of special interest since any gains in this context might help reduce the black-white achievement gap.

The major empirical challenge in any study of alternative school models is selection bias. Students who attend charter and pilot schools differ in a number of ways from the general pool of public school students, a fact that biases naive comparisons. We can hope to eliminate some of this bias by controlling for student characteristics such as free lunch status, but the possibility of bias from unobserved variables such as motivation or family background remains. An important aspect of our study, therefore, is the use of student admissions lotteries to estimate causal effects. These lotteries, which admit applicants randomly at over-subscribed schools, are used to construct a quasi-experimental research design that should generate unbiased estimates of the causal effects of charter and pilot attendance.

A charter or pilot school contributes application cohorts to our lottery estimates when the school is over-subscribed and therefore runs a lottery, has complete lottery records, and, in the case of pilots, uses a lottery to select students.³ In addition, the charter schools in our lottery sample were all operating at the time we were collecting lottery data (closed charter schools have often been under-subscribed). These selection criteria may have consequences for the external validity of our results. The over-subscription condition tilts our sample towards charter and pilot schools that parents find appealing, as does the requirement that schools still be open. From a policy perspective, however, this is an interesting set of schools. As it stands, Massachusetts currently limits both the number of charter schools and the proportion of a district's budget that can be lost due to charter enrollment. Were the supply of alternative school models allowed to freely vary, it seems reasonable to expect currently operating over-subscribed schools to expand and imitators to open similar schools.

The requirement that participating schools have complete lottery records also affects our selection of charter schools for the lottery sample. Specifically, the records requirement tilts the charter lottery

³ More precisely, a given *school-year-grade* cell contributes to the lottery analysis if entry at that point is over-subscribed and the resultant lottery records are available.

sample towards schools that have archived lottery records. Massachusetts law does not *require* charter schools to retain their lottery data. The net impact of the record-keeping constraint is unclear. On one hand, poor record-keeping may be a sign of disorganization that spills over into teaching. On the other hand, lottery record-keeping may be a distraction that takes time and energy away from instructional activity. In some cases, lost records are also a result of bad luck and the fact that the preservation of lottery data is not a priority once the school admissions process is complete.

Finally, on the pilot side, not all schools use the centralized lottery system that is embedded in the BPS school assignment mechanism. Some pilot schools opt out of the BPS assignment mechanism and chose students by a combination of admissions testing or audition. Non-lottery pilots share this feature with Boston's elite exam schools (the most famous of which is the Boston Latin School). In contrast, over-subscribed charters must use lotteries to select students.

In an effort to gauge the generality of our lottery-based findings, we complement the quasi-experimental lottery analysis with an observational analysis of the full set of charter and pilot schools. The observational analysis controls for demographic and background characteristics as well as students' lagged test scores (for example, the elementary school scores of middle school students). This investigation produces estimates remarkably similar to the lottery-based estimates when carried out in the sample of charter schools that have lotteries, lending some credence to the observational analysis. At the same time, the observational analysis suggests that the charter schools in our sample are better than others in the sense of generating larger treatment effects. We therefore think of our (mostly positive) charter estimates as indicative of what a certain charter model can accomplish, rather than an overall charter-school treatment effect.

For pilot middle schools, the observational analysis also generates results that are broadly consistent with the lottery-based analysis, but the observational results for pilot high schools are more positive than the corresponding lottery-based estimates. This leaves a somewhat muddled picture. Across all designs and samples, however, the estimate effects of attending a pilot school are smaller than

the corresponding charter estimates. In the subsample of pilot middle schools that serve only grades 6-8, lottery-based estimates of effects on middle school math scores are negative and significantly different from zero.

The next section describes Boston's charter and pilot schools in more detail and briefly discusses a few related studies and questions. Following that, Section III lays out our lottery-based estimation framework while Section IV discusses data and descriptive statistics. Section V presents the lottery analysis. This section also discusses the characteristics of the charter and pilot lottery compliers and results from models with treatment-effect interactions. Section VI reports observational results from a broader sample, and compares these with the lottery estimates. The paper concludes in Section VII.

II. Background

The 1993 Massachusetts Education Reform Act opened the door to charter schools in Massachusetts. Non-profit organizations, universities, teachers and parents can apply to the state's Board of Elementary and Secondary Education for a charter (there are no for-profit charter schools in Massachusetts). Massachusetts charter schools are generally managed by a board of trustees and are independent of local school committees. Like other public schools, charter schools charge no tuition and are funded mostly by sending districts according to formulas set by the state.

Massachusetts charter schools have a number of organizational features in common with charter schools in other states. First, they are typically outside local collective bargaining agreements. As a result, they have greater flexibility than traditional public schools when it comes to staffing, compensation, and scheduling. For example, the five Massachusetts charter schools studied by Merseeth (2009), four of which appear in our lottery sample, have a longer school day and year than public schools. Many charter schools offer extensive tutoring during and after school. Moreover, teachers in charter schools need not hold an active state license to begin teaching, though they must pass the Massachusetts Test for Educator Licensure within the first year of employment.

Wilson (2008) describes seven Boston charters, six of which are in our lottery sample, as a well as a charter school in Lynn, near Boston. Wilson identifies school practices prevalent at the schools in his sample. This collection of practices is sometimes said to characterize the “No Excuses” model, a term that probably originates with Thernstrom and Thernstrom (2003). No Excuses schools are characterized by small size, frequent testing, a long school day and year, selective teacher hiring, and a strong student work ethic. Other features include an emphasis on discipline and comportment, teacher-led whole-class instruction, and the absence of computer-aided instruction. Merseeth (2009) details account of the workings of five Boston-area charters, which she calls “high-performing schools,” suggests they share these features.

The first two columns of Table 1 compare some of the statistically measurable differences between Boston charter schools and traditional (BPS) public schools. This table shows student-weighted averages of teacher characteristics and student-teacher ratios by school type. The student-teacher ratio is substantially lower in charter schools, and charter teachers are less likely to be licensed or to be “highly qualified” as defined by NCLB. The latter is likely a consequence of the relative inexperience of many charter school teachers, a fact documented in the age distribution data show in the table.⁴ As shown in column 7 of Table 1, the features that distinguish the full roster of Boston charter schools are shared by the schools in our lottery sample.

Massachusetts charter schools appear to face more stringent accountability requirements than non-charter public schools. The state Charter School Office reviews and makes recommendations on charter applications, reviews the performance of existing charter schools, and decides whether charters should be renewed. Charter schools are held accountable via annual reports, financial audits, and site visits, and are required to file for renewal every five years. Renewal applications must show that a school’s academic program is successful, that the school is a viable organization, and that it has been

⁴ The definition of highly qualified has varied over time, but typically this designation is awarded to teachers who have a bachelor’s degree, full state certification or licensure, and have shown that they know the subject they teach (usually this requires some additional certification). Note that in Table 1, the denominators for the proportion licensed and the proportion highly qualified differ.

faithful to its charter. Since 1994, the state has received a total of 350 charter applications and has granted 76. Eight of the 76 Massachusetts charters ever granted were surrendered or revoked as of Fall 2009 (Massachusetts Department of Elementary and Secondary Education, 2009b). A ninth (Uphams Corner charter school) was revoked later in 2009.⁵

In the 2009-2010 school year, 26,384 Massachusetts students attended 62 operating charter schools, including 16 in Boston (Massachusetts Department of Elementary and Secondary Education 2009a). State law caps the number of charter schools at 72, and total enrollment at 30,034, so the statewide charter cap is not a constraint. However, a provision limiting local charter spending to nine percent of the district total generates binding or near-binding caps in specific districts, including Boston, where charter enrollment is already relatively high. The question of whether this local cap should be lifted is currently the subject of intense debate, fueled in part by the availability of federal stimulus money for states that facilitate new charters (Vaznis 2009).

Pilot schools

Pilot schools were developed jointly by BPS and the Boston Teachers Union (BTU) as an alternative to both charter schools and traditional public schools. Pilot schools are created as the result of a planning process currently funded by the Boston Foundation, a private charity, with technical assistance from the Center for Collaborative Education, a local nonprofit organization that runs the Boston Pilot Schools Network. New schools may be granted pilot status but most are conversions from traditional BPS schools. Pilot school conversions must be authorized by a two-thirds vote of the BTU membership employed at the school and authorized by the BTU Pilot School Steering Committee.⁶

⁵ Four of the eight charter losses through Fall 2009 occurred before school operations began. Two of the remaining four were revocations and two were non-renewals.

⁶ The pilot school model originated in Boston but other Massachusetts districts have begun to experiment with it. The Massachusetts Board of Elementary and Secondary Education recently adopted a Commonwealth Pilot School option for schools that otherwise would have been designated as underperforming under NCLB. Five

Like charter schools, pilot schools are answerable to independent governing boards. Also like charters, pilot schools determine their own budgets, staffing, curricula, and scheduling. Unlike charter schools, however, they remain part of the Boston school district and their teachers are BTU members covered by most contract provisions relating to pay and seniority. Pilot school teachers have no job protection within schools but remain in the BPS system if they choose to move or are removed by the pilot school principal.

Pilot teachers sign an election-to-work agreement which spells out the extent to which union contract provisions apply. These agreements vary by school.⁷ Pilot schools are subject to external review, but the review process to date appears to be less extensive and structured than the external state charter reviews. No pilot school has been closed or converted back to a traditional public school.⁸ Pilot schools are open to all students in the BPS district and operate as part of the district. In the 2007-8 school year, 6,337 BPS students were enrolled in 20 pilot schools. Assignment to all elementary and middle pilot schools, and to two of the seven pilot high schools, is through the centralized BPS choice plan, which includes a lottery when schools are over-subscribed.

Pilot teachers have characteristics between those of traditional BPS schools and charter schools, as can be seen in columns 3 and 8 of Table 1. For example, pilot teachers are younger than traditional BPS teachers but not as young as charter teachers. Many pilot schools share with charter schools longer school days and years. But the BTU agreement covering pilot schools limits uncompensated overtime, as do school-specific election-to-work agreements. This is reflected in statistics on hours of instruction that we collected from the schools in our lottery sample. The official BPS school year is 180 days, with a

Commonwealth Pilot Schools are now operating in Boston, Fitchburg, and Springfield. Versions of the pilot school model are also being tried in Los Angeles (Manzo, 2007).

⁷ See <http://www.ccebos.org/pilotschools/resources/index.html> for sample agreements.

⁸ For more on pilot structure, see http://www.ccebos.org/pilotschools/pilot_qa.doc and <http://www.ccebos.org/pilotguides/>. The current BTU contract allows for the creation of up to seven additional pilot schools. In 2007, two pilot conversions were voted down.

little over six hours of instruction per day, for a total of 1110 annual school hours. Annual school hours at pilot middle and high school hours are a little longer, but still under 1200 hours/year. In contrast, the average charter middle school in our sample provides 1500 hours of instruction, while charter high schools provide 1407 hours.⁹

Related Studies and Related Questions

A few recent studies use admissions lotteries to measure charter school achievement effects. These include Hoxby and Murarka (2009), which estimates charter effects in New York City (grades 3-8), Hoxby and Rockoff (2005), which estimates the effects of three Chicago charter schools, and Dobbie and Fryer (2009), which evaluates the effects of a charter middle and a charter elementary school located in the Harlem Children's Zone. All three studies report positive effects, though they vary by grade and subject. The achievement gains reported in Hoxby and Murarka (2009) and Hoxby and Rockoff (2005) are fairly modest, though Hoxby and Rockoff note problems with the Chicago charter lotteries that may invalidate some of their findings. The Dobbie and Fryer (2009) estimates for middle school math scores are the largest of the three, with impacts reaching almost half of a standard deviation per year in charter. Dobbie and Fryer also report IV estimates for an elementary school in the Harlem Children's Zone based on proximity to the Zone; these are imprecise. Farther afield, Clark (2009) uses a regression-discontinuity design to study the impact of attendance at Britain's grant-maintained schools, a charter-like model. Grant-maintained schools appear to have produced large achievement gains.¹⁰

A related study by Cullen, Jacob and Levitt (2006) uses admissions lotteries to estimate the effect of attending Chicago magnet schools. Like Boston pilot schools, Chicago magnet schools are part of the

⁹ Data on hours of instruction at charter and pilot schools come from the individual schools' web sites.

¹⁰ Charter evaluations that don't use lotteries have generally produced more mixed results. See, for example, Booker, Sass, Gill, and Zimmer (2008) for Chicago and Florida; Eberts and Hollenbeck (2002) for Michigan; Bifulco and Ladd (2006) for North Carolina; Booker, Gilpatric, Gronberg, and Jansen (2007) for Texas; Berends, Mendiburo, Nicotera (2008) for a Northwest Urban District; and CREDO (2009).

local (Chicago Public Schools) district. Chicago magnet schools are also like pilot schools in that they offer special programs and enjoy a degree of autonomy while operating within the Chicago Public School district. Cullen, Jacob and Levitt find no evidence that winning a magnet lottery raises test scores or any other measure of academic performance, though magnet students appear to be less likely to be involved in disciplinary action or to have been arrested.

The question of whether innovations in school organization generate achievement gains is also part of a debate over whether schools alone can address large racial achievement gaps. Positions on this point sometimes translate into views about the social significance of charter schools in the Knowledge is Power Program (KIPP). The students at KIPP schools are mostly from poor families and 95 percent are Black and Hispanic. Although there are no KIPP schools in our sample, most of the schools for which we have lottery data embrace the No Excuses model and serve a similar low-income, minority population.

The sense in which KIPP constitutes a marker for the role schools can play is clear from influential accounts on either side. In a study of the racial achievement gap, titled *No Excuses*, (Thernstrom and Thernstrom (2003, p. 43)) credit KIPP and similar No Excuses programs with impressive gains for low-income students:

Scattered across the American landscape are what some call “break-the-mold” schools — high-poverty public schools with students who score well on statewide tests. There aren’t many of them, and all are atypical within their own districts. Nevertheless, their record of success suggests that truly radical education innovation can change the lives of inner-city students, whatever their race or ethnicity.

An alternative view sees schools as only one part – arguably not the most important – in a broad struggle that requires social reform at many levels. Here is Rothstein (2004, p. 83), former New York Times education correspondent and an articulate spokesman for the systemic view:

... there is nothing illogical about a belief that schools, if well-operated, can raise lower-class achievement without investing in health, social, early childhood, after-school, and summer programs. But while the belief is not illogical, it is implausible, and the many claims made about instructional heroes or methods that close that gap are, upon examination, unfounded.

Although our limited study cannot be definitive, we see the results as relevant for the evaluation of these opposing points of view.¹¹

III. Empirical Framework

We're interested in the effects of charter or pilot school attendance on student achievement. Because the effects of attendance at different types of school seem likely to be an increasing function of the time spent in school, we model score effects as a function of years in pilot or years in charter. The causal relation of interest is captured using equations like this one for the scores, y_{igt} , of student i testing in year t in grade g :

$$(1) \quad y_{igt} = \alpha_t + \beta_g + \sum_j \delta_j d_{ij} + \gamma' X_i + \rho S_{igt} + \varepsilon_{igt}.$$

The variable S_{igt} is the years spent in a charter or pilot school as of the test date, counting any repeated grades, and counting time in all charter and pilot schools, not just the ones in our lottery samples. We also define a year to be a charter or pilot year if any portion of that year is spent in a charter or pilot school. The causal effect of S_{igt} is ρ . The terms α_t and β_g are year-of-test and grade-of-test effects, while X_i is a vector of demographic controls with coefficient γ , and ε_{igt} is an error term that reflects random fluctuation in test scores. The dummies d_{ij} are indicators for lottery-specific risk sets (indexed by j), described below.

If S_{igt} were randomly assigned, ordinary least squares (OLS) estimates of (1) would capture an average causal effect of years spent at a charter or pilot school. Because students and parents selectively chose schools, however, OLS estimates may be biased by correlation between school choice and unobserved variables related to ability, motivation, or family background. We therefore use an

¹¹ Researchers have also studied the effects of charter schools on traditional public schools. Examples include Bettinger (2005) and Booker, Gilpatric, Gronberg, and Jansen (2007). A related literature looks at the way parents make schooling choices (see, e.g., Hastings, Kane, and Staiger 2008; Hanushek, Kain, Rivkin, and Branch 2007; Jacob and Lefgren 2007). A few studies investigate general equilibrium effects of school choice on achievement (Hoxby 2000, 2007; Rothstein 2006, 2007).

instrumental variables (IV) strategy that exploits the partial random assignment of S_{igt} in school-specific lotteries. Assuming the applicant lotteries are fair, students who win and lose a given lottery should have similar characteristics. .

The first stage equations for IV estimation take the form:

$$(2) \quad S_{igt} = \lambda_t + \kappa_g + \sum_j \mu_j d_{ij} + \Gamma' X_i + \pi Z_i + \eta_{igt}$$

where λ_t and κ_g are year-of-test and grade effects in the first stage. The first-stage effect is the coefficient, π , on the instrumental variable, Z_i . The charter instrument is a dummy for having been offered a seat at one of the schools in the applicants' charter risk set. The pilot instrument is a dummy for having a BPS lottery number below the highest number offered a spot in the pilot risk set.

For a given charter applicant, the charter risk set is the list of all lotteries to which the student applied in a given year and entry grade, among the lotteries included in our charter lottery sample. Students who did not apply to any of the charter schools in the lottery sample are not in any charter risk set and are therefore omitted from the IV analysis. The relevant sample of pilot applicants includes only those students who listed a pilot school first on their BPS assignment form (few students who did not do so end up in a pilot school). The pilot risk set is defined by the identity of this first-choice school and the applicant's walk-zone status. Charter and pilot risk sets also vary by grade of entry and year of application (the entry cohort).¹²

IV. Data and Descriptive Statistics

Student Demographics, Schools Attended, and Test Scores

The Massachusetts Students Information Management System (SIMS) contains information on all Massachusetts public school students' race, ethnicity, sex, reduced-price lunch status, special education

¹² The relevant risk set for students in the pilot lottery is based on the BPS assignment mechanism. Among first-choice applicants to a given pilot school, admission priority is randomly assigned, with lotteries run separately for students who live inside and outside the school's walk-zone. In the pilot analysis, the risk set is therefore specified as the interaction of the four variables indicating the student's first-choice pilot school, walk-zone status for that school, and the year and grade of application.

status, English-language learner status, town of residence and current school. These data are collected in October and again at the end of the school year. We worked with SIMS files for the 2001-2002 through 2006-2007 school years. The SIMS data were used to determine how many years students spent in a charter, pilot or traditional BPS school. A student observed at any time during a school year in a charter or pilot school was classified as a charter or pilot student for that year.

To construct an analysis file, we used student identifiers to merge SIMS demographic and school history data with test scores from the Massachusetts Comprehensive Assessment System (MCAS) database, again for the 2001-2002 through 2007-2008 school years. The MCAS database contains raw scores for math, English language arts (ELA), and writing. MCAS is administered each spring, typically in grades 3-8 and 10. For the purposes of our analysis, scores were standardized by subject, grade, and year to have mean zero and unit variance in the population of students attending Boston schools.

Charter School Lotteries

Each charter school collects applications and holds its own lottery in years the school is over-subscribed. Siblings of students already attending the school are guaranteed a seat, as are students continuing on from earlier grades. Students can apply to as many charter schools as they like; the lotteries are statistically and administratively independent. Students may therefore be accepted or wait-listed at more than one school. When admitted students decline, slots open up for additional offers farther down the lottery list. Thus, some students may be offered spots immediately, while others may be offered seats closer to the beginning of the school year. Charter school lottery records were match to MCAS and SIMS data using applicants' names and year and grade of application.¹³

¹³ Gender, race, town of residence, and date of birth were used to resolve ambiguities. We matched 96.5 percent of charter applicants at the middle school level (97.3 percent of those admitted and 94.8 percent of those not admitted) and 92.2 percent of applicants at the high school level (92 percent of those admitted and 92.9 percent of those not admitted). Additional information related to the construction of analysis files appears in the data appendix.

Pilot School Lotteries

Students apply to pilot schools as part of the regular BPS assignment mechanism. BPS parents submit a rank order list of at least three schools in January to obtain a seat at a new school in September. At each school, admission priority is determined in part by whether the applicant is a continuing student who is guaranteed admission, currently has a sibling at the school, or lives in the school's walk zone. Within these priority groups, students are selected using an ordering determined by the BPS lottery number. The choice mechanism tries to assign as many students as possible to their top choice, using coarse priority rankings and lottery numbers when there are more applicants than capacity.¹⁴ This produces a system that induces random assignment (albeit with varying probabilities), conditional on priority groups such as sibling and walk-zone status.

Students were classified as pilot applicants if they listed a pilot school as their first choice. Because most pilot schools are oversubscribed, students who rank a pilot school as a second or lower choice are unlikely to be assigned to a pilot. The BPS assignment mechanism runs in multiple rounds but we use information only from the first round. Data on parents' choices and BPS lottery numbers came from the BPS applications data base. These data were matched to our merged MCAS-SIMS analysis file using state identifiers.

Lottery Sample Coverage

We asked all operating charter schools in Boston for current and past lottery records. The set of charter schools and applicant years included in the lottery sample are listed in Table A1. Of the four charter schools with elementary grades, one had lottery records for a period that we could use, but these are for 6th grade entrants and hence contribute to the middle school sample (this school added elementary grades only in 2006). Of the 11 charter schools that enroll middle school students, five contribute to the lottery analysis. Two of the omitted middle schools closed before or while our study was under way (one

¹⁴ For details, see Abdulkadiroglu and Sonmez (2003) and Abdulkadiroglu, Pathak, Roth and Sonmez (2006).

was under-subscribed), so that charter middle school coverage consists of 5 out of 9 currently operating schools. Two omitted middle schools admit primarily from their own elementary grades. In such cases, we were unable to recover lottery records because the entry date in this case precedes the outcome dates by at least seven years.

Of the eight charter schools that enroll regular (non-alternative) high school students, four contribute to the lottery analysis. Two of the omitted high schools closed before or during our study (one was under-subscribed) and two admit students at younger grades and are included in the middle school sample. Our charter high school sample therefore includes all four of Boston's currently operating regular 9-12 charter high schools.

All elementary and middle school pilots use the BPS assignment mechanism and lottery, but only two pilot high schools do. Four others use school-specific admissions criteria, such as musical auditions, to select their students. One is a 6-12 school that was not over-subscribed. Of the seven pilot schools that enroll elementary school students, five were over-subscribed and contribute to the lottery sample. Of seven pilot middle schools admitting 6th graders, six were over-subscribed and contribute to the lottery sample. Of the four K-8 pilot schools, our lottery middle school sample includes kindergarten applicants from three (the kindergarten entry grade is known as K2, the year after preschool, K1). One K-8 pilot school opened too late to contribute middle school test scores by K2 applicants.

One of the high schools in our charter lottery sample (Health Careers Academy or HCA) is what is known as a Horace Mann charter school. HCA is the only Horace Mann charter that serves the regular 9-12 population.¹⁵ HCA is unusual in that it began as a pilot school in 1995, but converted to a Horace Mann charter school in 1997.

Like pilot schools, a Horace Mann charter school's employees are members of the BTU bargaining unit and Horace Mann charters are established with the approval of BPS and the BTU. Otherwise, however, they operate like other charter schools (called Commonwealth charters). In

¹⁵ The only other Boston-area Horace Mann charter school, called Boston Day and Evening Academy, is an alternative school; it serves students aged 16-22 who are overage for grade level or have dropped out.

particular, HCA is subject to state (as opposed to district) supervision, and must submit to the same type of five-year review for charter renewal. HCA operates in rented space on the Northeastern University campus, does not use BPS facilities, and is funded like other charter schools.¹⁶ The elect-to-work agreements at HCA also appear to allow for more unpaid overtime than those at pilot schools. Because HCA has features in common with pilot schools, however, we look briefly at the implications of a reclassification of HCA as a pilot school in our sensitivity analysis.

Student Characteristics

Table 2 reports descriptive statistics for students at Boston's traditional schools, charter schools, and pilot schools, as well as a separate breakout for those included in the charter and pilot lottery samples. The racial and ethnic composition of the student bodies attending pilot elementary and middle schools is similar to that at traditional BPS schools: around 45 percent Black and 30 percent Hispanic. In contrast, charter schools have a higher proportion of Black students (about 70 percent) and a lower proportion of Hispanic students (about 20 percent). Differences in racial make-up across school types are similar at the high school level.

Roughly 85 percent of students at traditional Boston schools are eligible for free or reduced-price lunch, a measure of poverty status. Charter students are not as poor; about 70 percent fall into this category. The pilot school student body occupies a middle ground, with more poor students than the charter schools but fewer than the traditional schools. Relatively few English learners (also known as

¹⁶ The Massachusetts Department of Elementary and Secondary Education describes the two charter models as follows: A charter school is a public school that is governed by a board of trustees and operates independently of any school committee under a five year charter granted by the Board of Elementary and Secondary Education (Board). It has the freedom to organize around a core mission, curriculum, theme, and/or teaching method and to control its own budget and hire (and fire) teachers and staff. In return for this freedom, a charter school must attract students and produce positive results within five years or its charter will not be renewed. There are two types of charter schools, Commonwealth charter schools and Horace Mann charter schools. They operate under the same set of rules as described above, with the following exceptions. First, a Horace Mann charter school must have its charter approved by the local school committee and the local teacher's union in addition to the Board. Second, to the extent provided by their charters, Horace Mann charter schools may be exempt from certain provisions in local collective bargaining agreements. Third, employees of a Horace Mann charter school remain members of the local collective bargaining unit; continue to accrue seniority; and receive, at a minimum, the salary and benefits established by the local collective bargaining agreement. See also <http://www.doemass.org/charter/qanda.doc>.

limited English proficiency or LEP students) attend charter schools. For example, only seven percent of charter middle schools students are LEP, while the traditional Boston population is 22 percent LEP (pilot schools are close at 21 percent). Charter schools also enroll fewer special education students than traditional schools and pilot schools. Girls are over-represented at charter schools and, to a lesser extent, at pilot schools; this is particularly striking at the high school level, where nearly 60 percent of charter school students are female, compared to 52 percent at the pilot schools and 50 percent at traditional schools. Importantly, however, the demographic make-up of the charter and pilot lottery samples, described in columns 6 and 7 of Table 2, is similar to that of the total charter and pilot samples.

Table 2 also reports pre-treatment test scores - measured in elementary school for the middle school sample and in middle school for the high school sample. We think of these lagged score variables as baseline ability measures. For middle school students, baseline scores come from tests taken in fourth grade while for high school students baseline scores come from tests taken in eighth grade. There are no baseline scores for elementary school students, since MCAS testing starts in third grade. At the middle school level, pilot school students have somewhat lower baseline scores than students at traditional schools, while baseline scores of the charter students are closer to those for students in traditional BPS schools. At the high school level, charter school students have higher baseline scores, averaging about .4 standard deviations above those of students in traditional schools and a tenth of a standard deviation above those of students attending pilot schools. Among charter school students applying to lotteried middle schools, there is a baseline advantage of about .2 standard deviations.

Do the Charter and Pilot Lotteries Look Random?

As a measure of lottery quality, Table 3 reports covariate differences between winners and losers. The numbers reported in the table are regression-adjusted differences by win/loss status, where a win means students were offered a spot in a charter or pilot school in the relevant risk set. The only covariates in the regression are dummies for risk sets (year of application and the set of schools applied to for

charters; first-choice school, year of application, and walk zone status for pilots). Conditional on these covariates, lottery wins should be randomly assigned.

With a few exceptions, the differences in Table 3 are small and statistically insignificant. Among middle school applicants, charter lottery winners have baseline math scores a tenth of a standard deviation above losers' scores; at the high school level, winners are two percentage points more likely to be Asian and two percentage points more likely to be English learners. These differences are only marginally significant and could be due to chance. Among elementary pilot schools applicants, lottery winners are eight percentage points less likely to be eligible for a subsidized lunch, while at the high school level, they are six percentage points *more likely* to be eligible. Again, these and the other scattered marginally significant contrasts in the table seem likely to be chance findings. This conclusion is reinforced by the F statistics at the bottom of each column, which test the joint hypothesis that all differences in baseline test scores and background characteristics in the column are zero. Only two of these are close to a .05 p-value.

V. Lottery Results

Charter Schools

Charter middle school applicants who were offered a spot at one of the schools to which they applied spent about a year longer attending a charter school than applicants who were not offered a spot. This difference, which is shown in column 1 of Table 4 (labeled “first stage”) is measured as of the MCAS test date. The first stage is smaller than the three-year span of middle school because about a fifth of lottery winners never attend a charter school, while some lottery losers end up in charter school later, either because they enter the admissions lottery in a future year, gain sibling preference when a sibling wins the lottery, or move off a wait list after the offers coded by our instrument. In addition, some tests were taken as early as one year into middle school (the MCAS tests math and ELA proficiency in every middle school grade).

Middle-school students who win the charter lottery score almost $.2\sigma$ higher on ELA and $.4\sigma$ higher in math, as shown in column 2 of Table 4 (labeled “reduced form”). The 2SLS estimate of the effect of an additional year in a charter school is the ratio of the reduced-form estimates to the first-stage coefficients. Since the first stage coefficients are close to one, the 2SLS estimates (reported in column 3) are similar to the reduced form estimates, though their interpretation differs. When estimated without demographic controls, the 2SLS estimates imply that ELA scores increase by about $.2\sigma$ for *each year* in a charter, while the per-year math effect is $.4\sigma$. These estimates are reasonably precise, with standard errors around $.06$ -. $.08$, showing that our research design has the power to detect more modest effects as well. The addition of demographic controls changes this result little, as can be seen in column 4.

Although the reduced form effects on high school ELA and math scores are smaller than the corresponding reduced form effects in middle school, the high school first stage is also smaller. As a consequence, the ELA and math score gains generated by time spent in charter high schools are estimated to be similar to the corresponding 2SLS estimates for middle school. High school students also take a writing test; here the 2SLS estimates show gains of $.17\sigma$ -. $.28\sigma$. As with the results for middle school, the addition of demographic controls leaves the estimates largely unchanged.

Pilot Schools

Our lottery-based analysis of pilot effects looks at elementary-grade outcomes as well as test scores from middle and high school. The impact of a pilot school offer on time spent in elementary school is almost three years, as can be seen at the top of column 5 in Table 4. The relatively large elementary-level pilot first stage is driven by the fact that elementary school applicants apply to enter in kindergarten, while they are not tested until third or fourth grade. The reduced form effect of a pilot school offer on elementary school applicants is a little under $.2\sigma$, but this translates into a much smaller per-year effect of $.06\sigma$ -. $.07\sigma$, reported in column 7 for models without demographic controls.

The estimated effect of a pilot offer on time spent in high school is similar to the corresponding first stage for charter applicants, while the pilot middle school first stage is somewhat larger. On the other hand, the estimated effects on ELA and math scores – both reduced form and 2SLS – are much smaller (one is negative) and not significantly different from zero. Here too, it's worth pointing out that the standard errors are of a size that modest effects, say on the order of $.1\sigma$, would be detectable in middle school, though the high school design has less power. The only significant post-elementary pilot school estimate is for one of the writing scores (This is $.18\sigma$, a marginally significant result). Finally, as with the charter lottery results, the estimates using pilot school lotteries are similar with and without demographic controls. The pilot school estimates with demographic controls appear in column 8.

Robustness Checks

The strong charter school effects on middle school scores reported in Table 4 are readily apparent in a visual representation of IV estimates based on a version of equations (1) and (2). Averaging equation (1) conditional on treatment status and risk set (and dropping covariates), we have

$$(3) \quad E[y_{igt} | d_{ij}=1, Z_i] = \alpha_t + \beta_g + \delta_j + \rho E[S_{igt} | d_{ij}, Z_i]$$

Differencing (3) by offer status within risk sets, this becomes

$$(4) \quad E[y_{igt} | d_{ij}, Z_i=1] - E[y_{igt} | d_{ij}, Z_i=0] = \rho(E[S_{igt} | d_{ij}, Z_i=1] - E[S_{igt} | d_{ij}, Z_i=0]).$$

In other words, the slope of the line linking offer-status differences in test scores within risk sets to the corresponding offer-status differences in average years at a charter or pilot school should be the causal effect of interest, ρ .

The empirical counterpart of equation (4) for charter applicants' middle school math scores appears in Panel A of Figure 1. The unit of observation here is a charter risk set. The regression line fits the scatterplot well, with no highly influential or outlying risk sets. The slope of the line in the figure is $.44$ when weighted by the size of risk sets. The corresponding 2SLS estimate of ρ using a full set of

offer×risk set dummies as instruments in a model without out covariates is identical.¹⁷ In contrast, the analogous plot for pilot schools, plotted in Panel B, shows a flatter line, with a slope of -.045. The pilot x-axis has a wider range than that for charters because some of the pilot risk sets are small, and because applicants to pilot K-8 schools may spend a longer time in a pilot school than applicants to grade 6-8 schools. But omitting the small risk sets and outlying first stage values leaves the weak negative relation captured in the figure essentially unchanged.

Most of the lottery results shown in Table 4 are insensitive to the inclusion of controls for applicants' baseline test scores. This is documented in Table 5, which reports results with additional controls and alternative samples. For reference, Table 5 also repeats the 2SLS estimates with demographic controls from Table 4 (these regressions also include dummies for year of test, year of birth, and risk sets). Baseline test scores are available for students who have taken the MCAS before applying to the lottery; the sample used to construct pilot lottery estimates with baseline score controls therefore omits elementary school students and K2 applicants to pilot K-8 middle schools.

The addition of baseline score controls has almost no effect on the charter lottery estimates, a result documented in column 2 of Table 5. On the other hand, the pilot lottery estimates for middle school math scores turn negative and at least marginally significant with baseline score controls, as can be seen in column 6. There is little relation between the pilot lottery instruments and baseline scores, so this change in middle school math estimates cannot be attributed to omitted variables bias. Rather, it stems from the loss of K-8 pilot schools in the lagged-score sample. We confirmed this by estimating middle school pilot effects with demographic controls in a sample that includes grade 6-to-8 middle schools only. These results, reported in column 7 of Table 5, show estimates similar to those with lagged score controls. The complementary sample that includes only K-8 schools generates small positive effects (not reported

¹⁷ The sample analog of equation (4) can be thought of as a strategy that combines Wald estimates based on Z_i across risk sets. Generalized Least Squares (GLS) estimation of the sample analog of equation (3) can be shown to be the same as 2SLS using a full set of offer×risk set dummies as instruments (Angrist, 1991). Size-weighted least squares estimation of (4) is not algebraically the same as this 2SLS estimator, but should be close (and is indistinguishable from 2SLS in Panel A of the figure).

in the table. Thus, there is considerable heterogeneity in the pilot middle school math effects: grade 6-to-8 schools appear to be weaker than K-8 schools, at least as measured by their impact on MCAS math scores. It should also be noted, however, that the omission of K2 applicants to K-8 schools leads to a large drop in precision.

The results of adding an additional instrument to the charter lottery analysis are reported in column 3. The maintained instrument is a dummy for having been offered a charter seat any time after the relevant lottery, including off the waiting list (“eventual offer”), while the additional instrument is a dummy for having received an “initial offer” – that is, an offer of a seat in a charter school made immediately after the lottery was held.¹⁸ The multi-instrument models are over-identified so we can expect the resulting IV estimates to be more precise. In addition, the comparison of one- and two-instrument estimates provides an over-identification test for the underlying exclusion restrictions. In practice, the two-instrument results are close to those using a single instrument, while precision gains are modest at best.

The final alternative specification reported in Table 5 is motivated by the unusual history of Health Careers Academy (HCA), a Horace Mann charter high school. As noted in the background section, HCA was born as a pilot school in 1995 but converted to a Horace Mann charter school in 1997. Horace Mann charter schools are like other charter schools in most respects, but their employees remain part of the BTU bargaining unit. For the reasons detailed in our earlier discussion, we believe HCA should be seen as a charter school, albeit with union staff. Nevertheless, as a robustness check, we explored the implications of moving HCA from the charter to the pilot ledger.¹⁹

Reclassification of HCA as a pilot school has a modest effect on the charter lottery estimates: The estimated impact on high school math increases from $.23\sigma$ to $.3\sigma$ while the ELA effect falls from $.19\sigma$ to

¹⁸ The initial offer instrument can be constructed only for some schools; see appendix Table A1 for details.

¹⁹ Like other charter schools, HCA uses a stand-alone lottery while pilot schools are part of the Boston assignment mechanism. When HCA is introduced into the sample of pilot high schools, we define risk sets according to whether applicants picked one of the other two pilot high schools as a first choice, picked HCA as a first choice, or applied to HCA with one or the other of the other two schools as a first choice in the BPS mechanism.

.11 σ , as shown in column 4 of Table 5. The latter estimate is no longer significantly different from zero, but this is due more to a drop in precision than to the changed estimate (reclassification of HCA shifts 300 observations). The charter writing impacts also change little. Addition of HCA to the sample of pilot high schools produces a marginally significant ELA effect of about .11 σ and an insignificant math effect of .09 σ (HCA becomes the third pilot high school in this case). The effects of pilot high school attendance on writing increase and are more precisely estimated with HCA included. The increase in ELA and math estimates when HCA is treated as a pilot school suggests that HCA may be stronger than the two pilot high schools that are otherwise in our lottery sample. On the other hand, the changes observed in this context are modest enough to be chance fluctuations.

Interpreting Magnitudes

A natural benchmark for the estimates reported here is the black-white test score gap. Among students attending regular BPS middle schools, blacks score about .7 σ below whites on ELA and .8 σ below whites in math. The charter middle school effects of .15 σ on middle school ELA scores is therefore enough to essentially close two-thirds of the black-white reading gap for students who remain enrolled in a charter middle school in grades 6-8. The even larger math gains of about .4 σ are more than enough to eliminate the racial gap in mathematics achievement while students are in middle school. The effects of roughly .2 σ estimated for high school ELA and math are also large enough to close the black-white high school gap of about .8 σ in both subjects (assuming four years of charter high school enrollment). Of course, these are extrapolations based on our per-year average effects and not direct estimates. In practice, we should expect some reduction in the returns to charter attendance over a long enough period. Moreover, these calculations assume that the effects reported in Table 4 apply equally to blacks and whites. On the other hand, separate estimation by race (not reported here in detail) suggests that this is indeed the case.

Selective Attrition

Tables 4 and 5 use samples of students who participate in charter school and pilot school lotteries and for whom we have post-lottery test scores. Lottery winners and losers should be similar at the time the lotteries are held, since a coin flip is all that distinguishes them. Subsequent attrition may lead to differences in the follow-up sample, however, unless the attrition process itself is also random. In other words, we worry about differential and selective attrition by win/loss status. For example, losers may be less likely to be found than winners, since students who lose the opportunity to attend a charter or pilot school may be more likely to leave the public schools altogether. Differential attrition generates selection bias (although those who leave Boston for another Massachusetts public school district should turn up in our sample). A simple test for selection bias looks at the impact of lottery offers on the probability lottery participants contribute MCAS scores to our analysis sample. If differences in follow-up rates are small, selection bias from differential attrition is also likely to be modest.²⁰

Table 6 reports the mean follow-up rate for lottery participants in columns 1 and 4, and win-loss differentials in columns 2-3 and 5-6. Roughly 80 percent of charter lottery losers and about 70 percent of pilot lottery losers contribute a post-randomization test score. These high follow-up rates are due to the fact that our extract is limited to those enrolled in BPS at baseline and to our use of a statewide MCAS data set. Follow-up differentials by win/loss status were estimated using regression models that parallel the reduced forms reported in Table 4. Positive coefficients indicate that lottery winners are more likely to contribute an MCAS score.

The estimated follow-up differentials for charter high school applicants are small and not significantly different from zero. For example, high school charter lottery winners are about 3 percentage points more likely to contribute an ELA score than losers (when the differential is estimated with demographic controls; see column 2). The follow-up differentials for charter middle school outcomes are

²⁰ More formally, if attrition can be described by a latent-index model of the sort commonly used to model discrete choice in econometrics, then selection bias in lottery comparisons arises only if winning the lottery affects the probability of MCAS participation. See, e.g., Angrist (1997).

a little larger, on the order of 4-5 percentage points. Although these small differences are marginally significant, they seem unlikely to impart substantial selection bias.

There are virtually no attrition differentials for pilot middle schools. The largest differentials turn up for participants in pilot high school lotteries, as can be seen in columns 5 and 6 of Table 6. For example, controlling for demographic characteristics, high school winners are roughly 5 percentage points more likely to have taken the ELA test than losers, a significant effect with an estimated standard error of 2.6. But this too seems unlikely to explain our results, which show no effect on pilot lottery winners in high school. First, the most likely scenario for selective attrition has relatively high achieving losers dropping out. Second, the attrition differentials in this case are still fairly small. Nevertheless, as a check on the main findings, we discarded a few of the most imbalanced cohorts to construct a sample of charter middle school and pilot high school applicants with close-to-balanced attrition. We then re-estimated treatment effects using this balanced sample. Attrition differentials for balanced cohorts are reported in Appendix Table A3, while the corresponding lottery-based estimates of treatment effects are reported in A4. These results are similar to those reported in Table 4.

Charter and pilot compliers

Our lottery-based research design uses random assignment to construct causal effects for the subpopulation that spends time in a charter or pilot school as a consequence of winning a lottery. What sort of schools do these students go to? Do the losers in this group remain in schools that look like the BPS average? Or, are the schools they come from unusual in some way? On the other hand, do winning lottery participants end up in schools with a different socioeconomic and ability mix or go to schools with smaller classes? We address this set of questions using a causal IV framework to describe the school characteristics experienced by lottery applicants in alternative states of the world.

We begin by defining the causal relationship of interest to be the link between characteristics of schools and classmates and charter or pilot attendance. Let the attendance variable be D_i for applicant i ,

where this is recorded in the year immediately after application to schools in the relevant risk set. Let X_i be a characteristic of these schools, such as the fraction nonwhite or class size. To capture the fact that D_i affects school choice and therefore changes the characteristics of the school that applicant i attends, we write

$$X_i = X_{0i}(1-D_i) + X_{1i}D_i,$$

where X_{0i} and X_{1i} are the potential school characteristics that applicant i would be exposed to if he or she attends a regular school or a charter or pilot school. The causal effect of charter or pilot attendance on i is the difference in potential characteristics by school type, $X_{1i}-X_{0i}$.

Charter and pilot lottery compliers are applicants who would attend a charter or pilot school in the year after application if they win the lottery, but not otherwise. To describe this group formally, we write

$$D_i = D_{0i}(1-Z_i) + D_{1i}Z_i,$$

where D_{0i} and D_{1i} are *potential assignments* that tell us the type of school applicant i would attend if he or she wins or loses (as before, the instrument is Z_i , a dummy indicating lottery winners). Following Imbens and Angrist (1994), we impose monotonicity:

$$D_{1i} \geq D_{0i}.$$

In other words, winning the lottery can only make charter or pilot attendance more likely. Given this restriction, $D_{1i}-D_{0i}$ is equal to zero or one.

In the causal IV framework developed by Angrist, Imbens, and Rubin (1996), the set of charter or lottery *compliers* consists of applicants with $D_{1i}-D_{0i}=1$. This happens when $D_{0i}=0$ and $D_{1i}=1$. Compliers are those who do not attend a charter or pilot school if they lose the lottery but do attend if they win. The rest of the population has $D_{1i}-D_{0i}=0$, because either $D_{1i}=D_{0i}=1$ (indicating applicants who attend whether they win or not, the *always-takers*) or $D_{1i}=D_{0i}=0$ (indicating applicants who do not attend whether they win or not, the *never-takers*). Always-takers' and never-takers' are unaffected by the lottery, so IV estimates using lottery instruments do not reflect the impact of charter or pilot attendance on these two

groups. Replacing years-in-charter or pilot with a dummy for charter or pilot attendance, and using Y_{1i} and Y_{0i} to denote potential test score outcomes, IV identifies the average causal effect on compliers,

$$E[Y_{1i} - Y_{0i} | D_{1i} \geq D_{0i}].$$

This is the local average treatment effect (LATE) captured by lottery instruments.²¹

Here we are interested in the school environment or school characteristics rather than test scores. Moreover, we focus on the average type of schools attended by compliers in alternative states if the world rather than the treatment effect on school type. In other words, we focus on

$$\alpha_{1c} \equiv E[X_{1i} | D_{1i} \geq D_{0i}]$$

$$\alpha_{0c} \equiv E[X_{0i} | D_{1i} \geq D_{0i}].$$

We think of α_{0c} as the fallback position for compliers unable to attend the type of school to which they applied (because they lost the lottery); α_{1c} describes the schools attended by compliers whose lottery applications were successful. Abadie (2002) shows that these marginal means can be estimated by using Z_i as instrument for the following second-stage equations (ignoring covariates):

$$(5) \quad D_i X_i = \lambda_{1c} + \alpha_{1c} D_i + \xi_{1i}$$

$$(6) \quad (1 - D_i) X_i = \lambda_{0c} + \alpha_{0c} (1 - D_i) + \xi_{0i},$$

where λ_{1c} and λ_{0c} are terms that incorporate controls – in this case dummies for risk sets.²²

Compliers who lose the lottery end up in schools with a demographic mix similar to the system-wide BPS average. This can be seen in columns 1 and 5 of Table 7 for charter applicants and columns 3 and 7 for pilot applicants. For example, the average proportion black in BPS middle schools is 47 percent (a statistic from Table 2). Non-treated charter compliers attend middle schools that are also 47 percent Black while non-treated pilot compliers attend middle schools that are 46 percent Black. Likewise, the

²¹ LATE can be generalized to covered ordered treatments such as years in charter or pilot, and models with covariates; see Angrist and Pischke (2009) for details.

²² Abadie (2002) uses equations like (5) and (6) to estimate the marginal distribution of potential *outcomes* for compliers, but his approach works for characteristics as well. Because we use a saturated model for controls (risk sets), 2SLS estimates of α_{1c} and α_{0c} are the same as those that would be generated using Abadie's (2003) kappa-weighting scheme to estimate compliers' marginal means.

average proportion Black in BPS high schools is .51, while non-treated compliers attend high schools that are 55 percent Black.

In some respects, the complier fallback position is somewhat better than the BPS average. This suggests that charter treatment effects cannot be explained by arguing that the regular BPS schools they would otherwise attend have unusually low-achieving peers. At losers' middle schools, for example, the mean baseline scores of peers are positive while the overall BPS average is negative. In high school, peer means for losers are higher than the BPS average for math though about the same for ELA scores. Socioeconomic status as measured by the proportion getting a free or reduced price lunch is also higher at the fallback schools than in traditional BPS schools (e.g., compare .77 and .67 in columns 1 and 5 of Table 7 with BPS means of .89 and .85). Class sizes at these schools are smaller than the traditional BPS average. Also noteworthy is the fact that the characteristics of fallback schools are similar for pilot and charter lottery applicants.

Successful charter and pilot compliers attend middle schools that have fewer minorities than the schools attended by losers. In high school, however, winning compliers have more Black and female classmates (though fewer Hispanic classmates) than losers. Other differences in school environment include fewer SPED and LEP students and higher baseline scores at the charter middle schools attended by winners. For example, the peer mean middle school ELA score for winners is .35, versus .11 for losers. For high school compliers, the peer mean math score is .05, versus -.385 for losers. On the other hand, pilot compliers experience similar differentials: among both middle and high school compliers, peer mean baseline scores are higher, there are fewer LEP and SPED students, and the average proportion receiving a free or reduced price lunch is lower for winners than for losers. This weighs against a pure peer effect interpretation of charter students gains, since a similar peer differential does not produce gains for pilot compliers. Its worth noting, however, that charter middle school compliers end up in smaller classes if they win the lottery. Pilot compliers, by contrast, attend schools with larger classes.

Ability Interactions and Peer Effects

The distribution of school characteristics described in Table 7, as well as the fact that charter applicants are positively selected, motivates our analysis of treatment effect heterogeneity. Specifically, we explore treatment effect interactions with an applicants' own ability and interactions with the ability of peers. The interaction with own ability addresses the question of whether charter and pilot schools cater to a relatively high-ability group since charter applicants (and pilot high school applicants) have higher baseline scores. The interaction with peer ability provides some evidence on the extent to which peer effects might explain our findings.

The equation used to estimate models with interaction terms looks like this:

$$(7) \quad y_{igt} = \alpha_t + \beta_g + \sum_j \delta_j d_{ij} + \gamma' X_i + \rho_0 S_{igt} + \rho_1 A_i (S_{igt} - s_g) + \varepsilon_{igt},$$

where A_i is student i 's baseline score and s_g is the average score in the sample, so that the main effect, ρ_0 , is evaluated at the mean.²³ The set of covariates, X_i , includes baseline scores. The interaction term of interest is ρ_1 . The corresponding first stage equations are

$$(8a) \quad S_{igt} = \lambda_{1t} + \kappa_{1g} + \sum_j \mu_{1j} d_{ij} + \Gamma_1' X_i + \pi_{10} Z_i + \pi_{11} A_i Z_i + \eta_{1igt}$$

$$(8b) \quad A_i (S_{igt} - s_g) = \lambda_{2t} + \kappa_{2g} + \sum_j \mu_{2j} d_{ij} + \Gamma_2' X_i + \pi_{20} Z_i + \pi_{21} A_i Z_i + \eta_{2igt},$$

so that equation (7) is identified by adding an interaction between A_i and Z_i to the instrument list.

The effects of charter middle school attendance are larger for students with *lower* baseline scores, though the estimated interaction terms are small. This can be seen in the second column of Table 8, which reports 2SLS estimate of ρ_1 in equation (7). For example, a lottery applicant with a baseline score $.2\sigma$ below the mean is estimated to have an ELA score gain that is $.02\sigma$ higher ($.094 \times .2 = .019$) and a math score gain that is $.027\sigma$ higher ($.137 \times .2$) than an applicant with a baseline score at the mean. None of the estimated own-ability interaction terms for applicants to charter high school are significantly different from zero. These results weigh against the view that charter schools focus on high achieving applicants.

²³The variables A_i and s_g are fixed in any given sample, including those that pool grades, but vary for a given applicant from middle to high school.

The one (marginally) significant own-ability interaction for pilot schools is that for high school math scores. This interaction term is of the same order of magnitude as the charter interaction terms for middle school. On the other hand, no significant own-ability interactions emerge from the analysis of treatment effects in pilot middle schools. Thus, the evidence in column 4 suggests that to the extent that effects in either charter or pilot schools vary with students' baseline scores, those with lower scores benefit more, especially in the charter case. Overall, however, the variation in treatment effects with pre-treatment scores is modest.

Table 8 also shows little evidence of peer effects of the usual sort. In particular, columns (6) and (8) report 2SLS estimates of ρ_1 in equation (7) when A_i is mean baseline score in the risk set. We might think, for example, that one of the benefits of charter attendance is the opportunity to attend schools with higher-achieving peers. In practice, however, the score gain from charter middle school attendance varies *inversely* with peer achievement. For example students, who apply to schools in a risk set with peer means $.1\sigma$ above the sample mean are estimated to have an ELA gain that is reduced by $.07\sigma$ and a math gain that is reduced by $.1\sigma$ for each year in charter. These negative interactions are noteworthy since the charter effect on middle school math scores is the most dramatic result reported here.

A marginally significant positive peer interaction appears for charter effects on high school math scores, but this is the only evidence we have for interactions with peer means in charter high schools. The corresponding interaction term for ELA, for example, is only $.05\sigma$ with a much larger standard error. There is one significant peer interaction for pilot applicants in the effect of middle school on math scores. It should be noted, however, that many of the peer-mean interactions are imprecisely estimated, especially for pilot high schools.

VI. Observational Findings

Our lottery analysis use a sample of applicants and schools for which lotteries are both relevant and documented. To gauge the generality of the lottery findings, we also estimated the effects of charter

and pilot school attendance in a larger sample that includes all Boston schools. In this case, identification of causal effects is based on statistical controls for family background and earlier achievement. We refer to this analysis as an “observational study”, to distinguish it from the quasi-experimental lottery-based identification strategy. The observational results consist of OLS estimates of equations like (1), controlling for student demographics and baseline scores.

The data structure for the observational analysis is similar to that used for the quasi-experimental study. Baseline scores and demographics for middle school come from 4th grade data, while baseline scores and demographics for high school come from 8th grade data. The regressors of interest count years spent attending a charter or pilot school at the relevant level (e.g., years in a charter middle school, in grades 6-8), as well as time spent in an exam or alternative school. Specifically, the observational estimates were constructed by fitting

$$(9) \quad y_{igt} = \alpha_t + \beta_g + \gamma'X_i + \rho_c C_{igt} + \rho_p P_{igt} + \rho_e E_{igt} + \rho_{alt} ALT_{igt} + \varepsilon_{igt},$$

where C_{igt} , P_{igt} , E_{igt} , and ALT_{igt} denote years in a charter, pilot, exam, or alternative school, with corresponding effects, ρ_c , ρ_p , ρ_e , and ρ_{alt} . To parallel the exclusion of K-8 pilot school applicants in the lottery analysis with baseline scores, the observational analysis that controls for baseline scores is estimated on samples that omit students who attended elementary grades in a K-8 pilot.

The observational charter results, reported in Table 9, are positive and significantly different from zero for every score outcome except elementary school math. Moreover, at every grade level and for every test, the charter estimates are more positive than the pilot estimates. Control for baseline scores has surprisingly little effect on the charter estimates. In models with baseline score controls, each year spent at a charter middle school is estimated to increase ELA scores by about $.1\sigma$ and to increase math scores by about $.18\sigma$. In contrast, the corresponding estimates for pilot middle schools are significant and negative, on the order of $-.08\sigma$ for ELA and $-.1\sigma$ for math. For both pilot schools and charter schools, the high school effects in Table 9 are positive and significant for all score outcomes, though the estimated charter effects are consistently larger than the parallel estimates for pilots. The charter effects in

elementary school are also larger than those for elementary school pilots; neither of the pilot estimates in this case is significant.

The observational results in Table 9 are broadly consistent with the lottery results, but the magnitudes differ. For example, the charter middle school effects in Table 9, while still positive and economically and statistically significant, are smaller than those reported in Table 4. Is this difference due to selection bias or to the broader coverage of the observational sample? Table 10 addresses this question by showing observational results in and out of the lottery sample, alongside the corresponding lottery estimates. These estimates were constructed by replacing the terms $\rho_c C_{\text{igt}}$ and $\rho_p P_{\text{igt}}$ in equation (9) with terms that allow for separate effects for time spent in schools in and out of the set of schools in the lottery sample. For example, we replace $\rho_c C_{\text{igt}}$ with $\rho_{lc} C_{\text{ligt}} + \rho_{nc} C_{\text{nigt}}$ where C_{ligt} counts years spent in a charter school that is included in the lottery sample and C_{nigt} counts non-lottery charter years (note that $C_{\text{igt}} = C_{\text{ligt}} + C_{\text{nigt}}$). The observational estimates reported in Table 10 control for baseline scores.

Allowing the effects of time spent in schools in and out of the lottery sample to differ goes a long way towards reconciling the observational and lottery estimates, especially for charter schools. For example, the observational estimates of the effects of time attending a charter middle school in the lottery sample are $.16\sigma$ for ELA and $.31\sigma$ for math. These estimates are not too far from the corresponding lottery estimates with baseline scores ($.14\sigma$ for ELA and $.39\sigma$ for math). The high school estimates are also a good match: compare, for example, observational estimates of ELA effects of about $.19\sigma$ using both designs.

Because the lottery and observational estimates of charter effects line up reasonably well when estimated for schools in the same sample, the results in Table 10 support the notion that the observational study design does a good job of controlling for selection bias in this case. On the other hand, the table also suggests that the charter schools in our lottery sample are among the best in the city. Observational estimates of the effects of time spent in the charter schools not included in our lottery sample, while still

economically and statistically significant for every outcome, are uniformly smaller than the corresponding estimates for the effects of time spent in lottery-sample schools.

The observational and lottery-based analyses of pilot middle schools are also broadly consistent in that both produce negative effects in the sample that includes lagged scores. The observational results for ELA are more negative than the corresponding lottery estimates, while the opposite is true for math. The match across designs is not as good for the high school results, where the observational analysis using the lottery sample produces substantial and significant positive effects for all outcomes. In contrast, the lottery results for high school ELA and math are small and not significantly different from zero (though the match for writing is good). The variance across designs may be due to the fact that the lottery estimates for pilot high schools are not very precise. Also noteworthy is the fact that small positive effects of pilot attendance on elementary school achievement turn up using both the observational and lottery designs.

On balance, as with the charter results, Table 10 mostly reinforces the lottery findings for pilots. Moreover, with the exception of middle schools, the observational estimates of pilot-school treatment effects are also larger for schools used in the lottery sample than for other pilot schools. Like the middle school effects, the estimated impacts of time spent at a non-lottery elementary school pilot are negative.

VII. Summary and Conclusions

Lottery-based estimates of the impact of charter attendance on student achievement in Boston show impressive score gains in middle and high school. We also evaluate an alternative to the charter model, Boston's pilot schools. Lottery-based estimates for pilot schools are small and mostly insignificant. In an effort to gauge the generality of our lottery-based findings, we complement the quasi-experimental lottery analysis with an observational study of the full set of charter and pilot schools. This approach produces estimates similar to the lottery-based estimates for the sample of charter schools that have lotteries, lending credence to the observational analysis.

Across study designs and samples, the estimate effects of attending a pilot school are smaller than the corresponding charter estimates. Although we cannot say pinpoint the source of differential impacts, a number of factors seem likely to be important. The student-teacher ratio is much smaller in the charter schools, and the school day and school year are considerably longer. On both of these dimensions, pilot schools are closer to traditional Boston public schools than they are to charter schools. These differences may originate in collective bargaining constraints that make it more expensive for pilot schools to expand instructional hours and staff their schools with more teachers per student. In addition, many of the charter schools in our sample use elements of the No Excuses model, an instructional paradigm that is not common in public schools, pilot or otherwise.

A natural benchmark for our estimates is the black-white test score gap. Among students attending regular BPS middle schools, blacks score about .7 standard deviations below whites in language arts and .8 standard deviations below whites in math. Our estimated charter school effects are large enough to reduce the black-white reading gap in middle school by two-thirds. The even larger math gains (about .4 standard deviations) are more than enough to eliminate the racial gap in math while students are in middle school. The effects of roughly .2 standard deviations estimated for high school ELA and math are also large enough to close the black-white high school gap of about .8 standard deviations in both subjects (assuming four years of charter high school enrollment).

It bears repeating that the lottery results reported here do not estimate overall charter and pilot school effects, even for Boston. Rather, the lottery results are generated by schools that parents find most appealing and, in the case of charter schools, those that have well-documented lotteries and that continue to operate. Perhaps not surprisingly, our observational analysis suggests that the schools in our charter lottery sample are among the most effective charter schools. On the other hand, results for this relatively effective set of schools seem likely to be of considerable policy interest. Massachusetts, along with a number of other states, caps the number of charter schools, while the U.S. Department of Education is pressing states to lift these caps. It seems likely that over-subscribed and well-run schools will be the first

to expand if caps are lifted. State regulators might also reasonably choose to encourage replication of schools that have documented effectiveness and high demand.

Table 1: Teacher Characteristics by School Type

	Traditional	Pilot, Charter, Exam or Alternative School				Lottery Sample	
	BPS Schools (1)	Charter (2)	Pilot (3)	Exam (4)	Alternative (5)	Charter (7)	Pilot (8)
<i>I. Elementary School (3rd and 4th grades)</i>							
Teachers licensed to teach assignment	86.0%	60.0%	73.2%	-	70.6%	-	71.9%
Core classes taught by highly qualified teachers	90.6%	61.3%	78.2%	-	56.6%	-	77.8%
Student/Teacher ratio	15.7	11.4	15.9	-	6.9	-	15.8
Proportion of teachers 32 and younger	26.6%	64.5%	51.8%	-	27.3%	-	50.4%
Proportion of teachers 49 and older	39.9%	8.0%	11.9%	-	31.6%	-	11.1%
Number of schools	72	3	7	-	2	-	5
<i>II. Middle School (6th, 7th, and 8th grades)</i>							
Teachers licensed to teach assignment	77.8%	53.9%	65.8%	90.8%	48.6%	54.4%	65.5%
Core classes taught by highly qualified teachers	84.8%	70.4%	70.2%	94.5%	45.4%	73.1%	69.8%
Student/Teacher ratio	16.1	11.9	19.5	21.1	5.2	11.9	19.6
Proportion of teachers 32 and younger	27.1%	74.5%	55.0%	30.0%	28.6%	81.1%	54.4%
Proportion of teachers 49 and older	36.0%	4.8%	13.6%	43.3%	27.8%	1.3%	13.9%
Number of schools	29	11	7	3	4	5	6
<i>III. High School (10th grade)</i>							
Teachers licensed to teach assignment	80.9%	57.6%	64.1%	90.7%	75.8%	57.7%	73.5%
Core academic teachers identified as highly qualified	85.7%	78.6%	72.7%	94.3%	80.6%	82.1%	83.6%
Student/Teacher ratio	17.6	10.9	16.0	21.1	8.9	10.6	17.5
Proportion of teachers 32 and younger	31.9%	66.9%	44.7%	30.0%	29.7%	64.3%	41.3%
Proportion of teachers 49 and older	40.3%	6.9%	15.0%	43.9%	25.3%	8.2%	7.7%
Number of schools	22	8	7	3	4	4	2

Notes: This table reports student-weighted average characteristics of teachers and schools using data for 2004-2007 posted on the Mass DOE website at http://profiles.doe.mass.edu/state_report/teacherdata.aspx. The variable *teachers licensed to teach assignment* is the fraction of teachers who are licensed with provisional, initial, or professional licensure to teach in the area(s) in which they are teaching. The variable *core classes taught by highly qualified teachers* is the percent of core academic classes (defined as English, reading or language arts, mathematics, science, foreign languages, civics and government, economics, arts, history, and geography) taught by highly qualified teachers (defined as teachers not only holding a Massachusetts teaching license, but also demonstrating subject matter competency in the areas they teach). For more information on the definition and requirements of highly qualified teachers, see http://www.doe.mass.edu/nclb/hq/hq_memo.html.

Table 2: Descriptive Statistics

	Traditional BPS Schools (1)	Enrolled in Pilot or Charter		Applicants in Lottery Sample		Applicants in Lottery Sample with Baseline Scores	
		Charter (2)	Pilot (3)	Charter (4)	Pilot (5)	Charter (6)	Pilot (7)
I. Elementary School (3rd and 4th grades)							
Female	48.3%	52.4%	48.4%	-	50.6%	-	-
Black	43.4%	72.0%	43.3%	-	54.3%	-	-
Hispanic	34.5%	15.9%	32.0%	-	22.0%	-	-
Special education	10.4%	6.3%	10.7%	-	9.9%	-	-
Free or reduced price lunch	83.2%	67.8%	69.1%	-	66.5%	-	-
Limited English proficiency	28.8%	3.8%	19.3%	-	7.0%	-	-
Years in charter	0.016	4.558	0.009	-	0.222	-	-
Years in pilot	0.031	0.023	3.801	-	1.803	-	-
Number of students	10,443	656	823	-	573	-	-
Number of schools	75	3	7	-	5	-	-
II. Middle School (6th, 7th, and 8th grades)							
Female	47.0%	48.9%	49.9%	48.3%	52.6%	48.2%	54.9%
Black	46.9%	69.4%	50.5%	59.2%	49.8%	59.1%	50.6%
Hispanic	37.3%	19.0%	28.3%	19.4%	31.2%	19.6%	35.0%
Special education	24.5%	18.5%	21.4%	19.3%	17.5%	19.1%	18.2%
Free or reduced price lunch	89.3%	73.1%	85.6%	69.1%	79.3%	69.1%	87.8%
Limited English proficiency	21.8%	7.1%	21.0%	7.7%	15.0%	7.8%	18.0%
4th Grade Math Score	-0.119	-0.069	-0.196	0.167	-0.077	0.167	-0.077
4th Grade ELA Score	-0.113	0.080	-0.127	0.235	-0.018	0.235	-0.018
Years in charter	0.018	2.458	0.012	1.480	0.220	1.469	0.221
Years in pilot	0.023	0.033	2.149	0.120	0.954	0.119	1.054
Number of students	12,257	2,382	2,696	1,355	1,917	1,331	1,298
Number of schools	33	11	7	5	7	5	6
III. High School (10th grade)							
Female	50.1%	59.9%	52.2%	59.1%	44.7%	59.0%	44.8%
Black	50.9%	65.8%	53.8%	67.6%	58.1%	67.6%	57.9%
Hispanic	36.1%	15.6%	26.7%	23.0%	24.9%	22.9%	25.0%
Special education	22.8%	14.8%	17.5%	15.5%	12.7%	15.3%	12.6%
Free or reduced price lunch	84.7%	66.7%	77.1%	75.7%	78.5%	76.1%	79.0%
Limited English proficiency	18.9%	3.9%	7.2%	4.2%	5.5%	4.1%	5.6%
8th Grade Math Score	-0.288	0.131	0.059	0.092	0.163	0.092	0.163
8th Grade ELA Score	-0.187	0.231	0.148	0.193	0.209	0.193	0.209
Years in charter	0.013	2.012	0.023	0.483	0.971	0.484	0.970
Years in pilot	0.006	0.009	1.951	0.719	0.269	0.718	0.271
Number of students	9,135	1,149	1,949	1,957	1,010	1,934	1,003
Number of schools	23	8	7	4	2	4	2

Notes: The table reports sample means in baseline years by school type. The numbered notes below describe each sample. Demographic characteristics are taken from grade K2 for elementary school students, grade 4 for middle school students, and grade 8 for high school students. All students in the sample reside in Boston and were enrolled in BPS or a charter school in the baseline year. To be included, students must have at least one MCAS score among the following: ELA and Math for elementary and middle school students; ELA, Math, Writing Topic, and Writing Composition for high school students.

1. BPS students excluding exam, alternative, charter and pilot students from 2004-2008.

2. Students enrolled in charter schools from 2004-2008.

3. Students enrolled in pilot schools from 2004-2008.

4. Charter applicant cohorts in randomized lotteries: middle school students in 2002-2007, and high school students in 2002-2006.

5. Pilot applicant cohorts: elementary school students in 2002-2004, middle school students in 2002-2007, and high school students in 2003-2006.

Table 3: Covariate Balance with Lottery Winners minus Lottery Losers at Charter and Pilot Schools

	Charter Schools				Pilot Schools					
	Middle School		High School		Elementary School		Middle School		High School	
	Lotteries with		Lotteries with		Lotteries with		Lotteries with		Lotteries with	
	All Lotteries	Baseline Scores	All Lotteries	Baseline Scores	All Lotteries	Baseline Scores	All Lotteries	Baseline Scores	All Lotteries	Baseline Scores
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hispanic	-0.004 (0.024)	-0.005 (0.024)	-0.029 (0.023)	-0.028 (0.023)	-0.032 (0.038)	-	-0.016 (0.025)	-0.034 (0.038)	0.021 (0.028)	0.016 (0.028)
Black	-0.014 (0.029)	-0.013 (0.030)	0.026 (0.026)	0.027 (0.026)	0.016 (0.042)	-	0.007 (0.027)	0.008 (0.040)	-0.002 (0.031)	0.006 (0.031)
White	0.018 (0.023)	0.018 (0.024)	-0.010 (0.012)	-0.011 (0.012)	0.028 (0.036)	-	0.001 (0.019)	0.026 (0.022)	-0.017 (0.017)	-0.020 (0.017)
Asian	0.003 (0.008)	0.003 (0.008)	0.021** (0.011)	0.019* (0.011)	-0.031* (0.018)	-	0.001 (0.014)	-0.001 (0.021)	0.000 (0.015)	0.001 (0.016)
Female	0.025 (0.031)	0.030 (0.032)	-0.004 (0.026)	-0.004 (0.026)	0.013 (0.049)	-	0.017 (0.030)	0.030 (0.043)	0.015 (0.031)	0.009 (0.031)
Free or Reduced Price Lunch	-0.010 (0.029)	-0.008 (0.029)	0.007 (0.023)	0.008 (0.023)	-0.080* (0.043)	-	-0.013 (0.023)	-0.019 (0.029)	0.059** (0.026)	0.065** (0.026)
Special Education	-0.017 (0.025)	-0.017 (0.025)	-0.011 (0.020)	-0.013 (0.020)	-0.026 (0.026)	-	0.000 (0.020)	0.022 (0.034)	-0.025 (0.023)	-0.021 (0.023)
Limited English Proficiency	0.021 (0.015)	0.019 (0.015)	0.021* (0.011)	0.022* (0.011)	-0.018 (0.026)	-	-0.033** (0.016)	-0.051* (0.030)	0.015 (0.015)	0.007 (0.015)
Baseline ELA Test Score	-	0.029 (0.053)	-	0.022 (0.043)	-	-	-	0.031 (0.077)	-	0.013 (0.054)
Baseline Math Test Score	-	0.095* (0.055)	-	0.076 (0.048)	-	-	-	0.076 (0.078)	-	-0.092 (0.057)
Baseline Writing Composition Test Score	-	-	-	0.006 (0.044)	-	-	-	-	-	0.046 (0.053)
Baseline Writing Topic Test Score	-	-	-	-0.079 (0.048)	-	-	-	-	-	0.028 (0.055)
p-value, from F-test	0.829	0.713	0.113	0.061*	0.046**	-	0.714	0.775	0.470	0.611

Notes: This table reports coefficients from regressions of the dependent variables listed on an indicator variable equal to one if the student won the lottery. Charter regressions include dummies for (combination of schools applied to)*(year of application) and exclude students with sibling priority. Pilot regressions include dummies for (first choice school)*(year of application)*(walk zone) and exclude students with sibling priority or guaranteed admission. Samples in columns (1), (3), (5), (7), and (9) are restricted to students from cohorts where we should observe at least one test score. Samples in columns (2), (4), (6), (8), and (10) are restricted to students who also have baseline test scores. F tests are for the null hypothesis that the coefficients on winning the lottery in all regressions are all equal to zero. These tests statistics are calculated for the subsample that has non-missing values for all variables tested. Standard errors, in parentheses, are clustered as described in the Table 4 notes

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

Table 4: Lottery Results

		Charter				Pilots			
Level	Subject	First Stage (1)	Reduced Form (2)	2SLS (3)	2SLS w/demos (4)	First Stage (5)	Reduced Form (6)	2SLS (7)	2SLS w/demos (8)
Elementary School									
	ELA	-	-	-	-	2.852*** (0.193)	0.196** (0.078) 876	0.069** (0.027)	0.064*** (0.024) 876
	N								
	Math	-	-	-	-	2.858*** (0.194)	0.177** (0.078) 874	0.062** (0.027)	0.060** (0.026) 874
	N								
Middle School									
	ELA	0.965*** (0.114)	0.181*** (0.061) 2,416	0.187*** (0.062)	0.149*** (0.052) 2,416	1.378*** (0.189)	0.021 (0.065) 3,390	0.016 (0.047)	0.006 (0.043) 3,390
	N								
	Math	0.918*** (0.108)	0.397*** (0.067) 2,582	0.432*** (0.075)	0.405*** (0.066) 2,582	1.291*** (0.181)	-0.078 (0.069) 3,851	-0.060 (0.053)	-0.059 (0.048) 3,851
	N								
High School									
	ELA	0.742*** (0.139)	0.147*** (0.053) 1,947	0.199*** (0.062)	0.187*** (0.055) 1,947	0.644*** (0.108)	0.013 (0.049) 1,007	0.020 (0.075)	0.007 (0.073) 1,007
	N								
	Math	0.738*** (0.137)	0.235*** (0.059) 1,929	0.318*** (0.077)	0.274*** (0.071) 1,929	0.632*** (0.108)	0.013 (0.065) 996	0.021 (0.100)	-0.011 (0.101) 996
	N								
	Writing Topic	0.741*** (0.138)	0.206*** (0.069) 1,931	0.278*** (0.083)	0.267*** (0.078) 1,931	0.636*** (0.109)	0.119** (0.059) 997	0.187** (0.093)	0.173* (0.093) 997
	N								
	Writing Composition	0.741*** (0.138)	0.127** (0.052) 1,931	0.171*** (0.064)	0.168*** (0.062) 1,931	0.636*** (0.109)	0.077 (0.054) 997	0.120 (0.084)	0.111 (0.086) 997
	N								

Notes: This table reports first stage, reduced-form, and 2SLS estimates using lottery-offer instruments. The sample is restricted to students with baseline demographic characteristics. Regressions include year of test and year of birth dummies. Middle school and elementary school regressions pool grade outcomes and include dummies for grade level. Charter regressions include dummies for (combination of schools applied to)*(year of application) and exclude students with sibling priority. Pilot regressions include dummies for (first choice)*(year of application)*(walk zone status) and exclude students with sibling priority or guaranteed admission. Columns (4) and (8) report 2SLS coefficients from specifications that add demographic controls, which include dummies for female, black, hispanic, asian, other race, special education, limited english proficiency, free/reduced price lunch, and a female*minority interaction. Elementary school tests are given in 3rd and 4th grade; middle school tests are given in 6th, high school tests are given in 10th grade. The table reports robust standard errors, clustered on year by 10th grade school for high school. Since the middle and elementary school regressions pool students across grades, the associated standard errors allow for two-way clustering on school-by-year and student identifier. Standard errors are reported in parentheses.

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

Table 5: Lottery Results, Robustness Checks

Table 3: Lottery Results, Robustness Checks									
		Charter Lotteries				Pilot Lotteries			
				Overidentified	High school w/HCA			No K-8 pilot	High school w/HCA
		Demo controls	Demo & baseline	model, demo	as pilot, demo		Demo & baseline	applicants, demo	as pilot, demo
Level	Subject	(1)	(2)	controls	controls	Demo controls	score controls	controls	controls
				(3)	(4)	(5)	(6)	(7)	(8)
Middle School									
	ELA	0.149*** (0.052)	0.144*** (0.044)	0.134*** (0.051)	-	0.006 (0.043)	-0.035 (0.112)	-0.079 (0.110)	-
	N	2,416	2,365	2,416		3,390	2,414	2,645	
	Math	0.405*** (0.066)	0.386*** (0.054)	0.370*** (0.061)	-	-0.059 (0.048)	-0.251** (0.106)	-0.233** (0.119)	-
	N	2,582	2,528	2,582		3,851	2,733	3,075	
High School									
	ELA	0.187*** (0.055)	0.186*** (0.049)	0.162*** (0.053)	0.112 (0.076)	0.007 (0.073)	-0.053 (0.059)	-	0.111* (0.065)
	N	1,947	1,629	1,947	1,683	1,007	949		1,367
	Math	0.274*** (0.071)	0.226** (0.060)	0.251*** (0.065)	0.303** (0.084)	-0.011 (0.101)	0.007 (0.070)	-	0.086 (0.077)
	N	1,929	1,892	1,929	1,664	996	983		1,355
	Writing Topic	0.267*** (0.078)	0.281** (0.083)	0.248*** (0.070)	0.225** (0.112)	0.173* (0.093)	0.151* (0.089)	-	0.214** (0.075)
	N	1,931	1,616	1,931	1,670	997	934		1,354
	Writing Composition	0.168*** (0.062)	0.132** (0.059)	0.146*** (0.055)	0.156* (0.089)	0.111 (0.086)	0.097 (0.080)	-	0.131** (0.065)
	N	1,931	1,616	1,931	1,670	997	934		1,354

Notes: This table reports modifications on the specifications reported in Table 4. Columns 1 and 5 are repeated from Table 4 models (with demographic controls). Columns 2 and 6 show the results of adding baseline scores. Column 3 presents results using both initial offer and eventual offer dummies as instruments for charter school attendance, so the model is overidentified. Columns 4 and 8 show the results of classifying HCA as a pilot high school. Column 7 shows pilot results for grade 6-8 middle schools only. All regressions also include year of test and year of birth dummies and use the same clustering scheme as used for Table 4. Standard errors are reported in parentheses.

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

Table 6: Attrition

Level	Subject	Charter			Pilot		
		Prop of non-offered with MCAS	Differential		Prop of non-offered with MCAS	Differential	
			Demographic Controls	Demographics + Baseline Scores		Demographic Controls	Demographics + Baseline Scores
		(1)	(2)	(3)	(4)	(5)	(6)
Elementary School							
	ELA	-	-	-	0.796	0.033 (0.037)	-
	N				686	1,085	-
	Math	-	-	-	0.796	0.032 (0.037)	-
	N				686	1,085	
Middle School							
	ELA	0.805	0.042* (0.021)	0.040* (0.022)	0.699	0.029 (0.024)	0.008 (0.026)
	N	923	2,869	2,801	2,625	4,596	2,778
	Math	0.811	0.046** (0.021)	0.046** (0.021)	0.702	0.030 (0.023)	0.009 (0.026)
	N	968	3,034	2,958	2,874	5,130	3,124
High School							
	ELA	0.776	0.027 (0.022)	0.020 (0.024)	0.749	0.053** (0.026)	0.074*** (0.026)
	N	825	2,433	2,026	786	1,300	1,210
	Math	0.767	0.029 (0.023)	0.028 (0.023)	0.740	0.048* (0.026)	0.064** (0.026)
	N	825	2,433	2,375	786	1,300	1,271
	Writing Topic and Writing Composition	0.768	0.028 (0.023)	0.026 (0.024)	0.743	0.048* (0.026)	0.073*** (0.026)
	N	825	2,433	2,019	786	1,300	1,200

Notes: This table reports effects of lottery offers on a dummy variable that indicates when the outcome test score is non-missing. The specification is the same as that used to construct the estimates in Table 4. The sample in this case includes all students in the baseline-demographics universe and not just those with test scores. Standard errors, in parentheses, are clustered as described in the Table 4 notes.

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

Table 7: Characteristics of Treated and Non-treated Schools for Compliers

School Characteristic	Middle Schools				High Schools			
	Charter		Pilot		Charter		Pilot	
	Non-Treated (1)	Treated (2)	Non-Treated (3)	Treated (4)	Non-Treated (5)	Treated (6)	Non-Treated (7)	Treated (8)
Fraction female	0.464	0.545	0.465	0.477	0.494	0.652	0.473	0.406
Fraction black	0.469	0.361	0.455	0.361	0.547	0.652	0.544	0.507
Fraction hispanic	0.273	0.192	0.371	0.220	0.322	0.242	0.256	0.254
Fraction with limited English proficiency	0.123	-0.002	0.109	0.103	0.147	0.001	0.110	0.021
Fraction special ed	0.227	0.103	0.267	0.186	0.188	0.089	0.194	0.143
Fraction free or reduced price lunch	0.767	0.501	0.786	0.762	0.668	0.683	0.631	0.502
Fraction with first language not English	0.310	0.143	0.390	0.382	0.368	0.229	0.346	0.315
Mean baseline ELA MCAS score	0.110	0.353	-0.043	0.262	-0.211	0.202	-0.168	0.039
Mean baseline Math MCAS score	0.098	0.380	-0.019	0.293	-0.385	0.050	-0.276	-0.078
Fraction of teachers licensed to teach assignment	0.904	0.496	0.889	0.857	0.842	0.776	0.864	0.898
Student/teacher ratio	12.680	10.605	12.639	13.084	14.644	13.372	14.221	14.786

Notes: This table reports mean treated and non-treated characteristics for compliers in the charter and pilot lotteries. The non-treated means are produced by estimating models of the form: $X_i(1-D_i) = \lambda_{0c} + \alpha_{0c}(1-D_i) + \xi_{0i}$, where X_i is the school characteristic of interest observed at the school actually attended by each student in the year immediately after the lottery, D_i is a dummy for whether the student attended charter/pilot in this year, λ_{0c} represents controls (risk set dummies), and $(1-D_i)$ is instrumented using the lottery win/loss dummy. The IV estimate of α_{0c} gives an estimate of the mean of X for the compliers in the non-treated state. The treated means are produced by estimating models of the form $X_i D_i = \lambda_{1c} + \alpha_{1c} D_i + \xi_{1i}$, where D is instrumented by the lottery win/loss dummy. The IV estimate of α_{1c} gives an estimate of the mean of X for the compliers in the treated state.

Table 8: Interaction Models

Level		Subject	Own baseline score				Mean baseline score in risk set			
			Charters		Pilots		Charters		Pilots	
			main effect	interaction	main effect	interaction	main effect	interaction	main effect	interaction
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Middle School										
	ELA		0.145*** (0.044)	-0.094* (0.051)	-0.030 (0.114)	-0.044 (0.053)	0.185*** (0.050)	-0.716** (0.346)	-0.009 (0.138)	-0.269 (0.453)
		N	2,365		2,414		2,365		2,414	
	Math		0.386*** (0.052)	-0.137** (0.060)	-0.250** (0.107)	-0.002 (0.041)	0.441*** (0.059)	-1.015*** (0.279)	-0.341*** (0.130)	0.535* (0.291)
		N	2,528		2,733		2,528		2,733	
High School										
	ELA		0.189*** (0.050)	0.092 (0.083)	-0.053 (0.058)	-0.026 (0.087)	0.185*** (0.050)	0.052 (0.419)	-0.040 (0.059)	0.792 (0.672)
		N	1,629		949		1,629		949	
	Math		0.236*** (0.061)	0.086 (0.066)	0.010 (0.065)	-0.102* (0.052)	0.218*** (0.055)	0.609** (0.297)	0.023 (0.070)	0.606 (0.565)
		N	1,892		983		1,892		983	
	Writing Topic		0.282*** (0.084)	0.037 (0.082)	0.153* (0.090)	-0.081 (0.087)	0.272*** (0.086)	0.385 (0.942)	0.164* (0.090)	0.143 (0.629)
		N	1,616		934		1,616		934	
	Writing Composition		0.137** (0.059)	0.078 (0.073)	0.117 (0.082)	-0.079 (0.091)	0.140** (0.068)	-0.238 (0.614)	0.129 (0.087)	0.428 (0.557)
		N	1,616		934		1,616		934	

Notes: This table shows results analogous to those reported in the 2SLS lottery results in Table 4, but the specifications include interaction terms. The models estimated are of the form: $y_{igt} = \alpha_t + \beta_g + \sum \delta_j d_{ij} + \gamma X_i + \rho_0 S_{igt} + \rho_1 A_i (S_{igt} - s_g) + \varepsilon_{igt}$, where y_{igt} is the outcome of interest for student i in grade g in year t , S_{igt} is years spent in charter (or Pilot), A_i is own baseline test score or mean baseline score in the risk set, the d_{ij} are risk set dummies, X_i is a vector of controls, and s_g is the average score in the sample so that the main effects ρ_0 are at the mean. X_i includes year of birth dummies, year of test dummies, demographic controls, and an own baseline score main effect. Standard errors, in parentheses, are clustered as described in the Table 4 notes.

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

Table 9: Observational Analysis for Charter and Pilot

Level	Subject	Demographics		Demographics & Baseline Scores	
		Charter (1)	Pilot (2)	Charter (3)	Pilot (4)
Elementary School	ELA ¹	0.055*** (0.017)	0.015 (0.020)	-	-
	N	20,058			
	R ²	0.134			
	Math ²	0.038 (0.023)	0.024 (0.023)	-	-
	N	17,356			
	R ²	0.131			
Middle School	ELA ³	0.116*** (0.014)	-0.072*** (0.014)	0.104*** (0.012)	-0.078*** (0.011)
	N	34,301		31,620	
	R ²	0.339		0.538	
	Math ⁴	0.176*** (0.020)	-0.096*** (0.017)	0.180*** (0.018)	-0.100*** (0.013)
	N	38,583		35,764	
	R ²	0.350		0.576	
High School	ELA ⁵	0.228*** (0.020)	0.155*** (0.018)	0.166*** (0.018)	0.094*** (0.016)
	N	16,609		12,347	
	R ²	0.487		0.623	
	Math ⁶	0.247*** (0.040)	0.126*** (0.026)	0.151*** (0.031)	0.052** (0.023)
	N	16,350		15,868	
	R ²	0.509		0.700	
	Writing Topic ⁷	0.228*** (0.028)	0.154*** (0.023)	0.206*** (0.031)	0.141*** (0.024)
	N	16,289		12,181	
	R ²	0.308		0.354	
	Writing Composition ⁸	0.204*** (0.021)	0.148*** (0.019)	0.178*** (0.022)	0.129*** (0.018)
	N	16,289		12,181	
	R ²	0.348		0.394	

Notes: This table reports OLS estimates of the effects of time spent in schools of various types. The excluded category is traditional BPS schools. The sample restricted to students with baseline demographic characteristics. Demographics include female, black, hispanic, asian, other race, special education, limited english proficiency, free/reduced price lunch, and a female*minority dummy. Regressions also include year of test and year of birth dummies. Middle school and elementary school regressions pool grade outcomes and include dummies for grade level. Regressions use robust standard errors and are clustered as in Table 4. Standard errors are reported in parentheses.

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

1. Elementary school ELA is for Grade 3 (2005-08) and Grade 4 (2005-08).

2. Elementary school Math is for Grade 3 (2005-08) and Grade 4 (2005-08).

3. Middle school ELA is for Grade 6 (2005-08), Grade 7 (2005-08), and Grade 8 (2005-08)

4. Middle school Math is for Grade 6 (2004-08), Grade 7 (2005-08), and Grade 8 (2006-08)

5. High school ELA is for Grade 10 (2004-08).

6. High school Math is for Grade 10 (2004-08).

7. High school Writing Topic is for Grade 10 (2004-08).

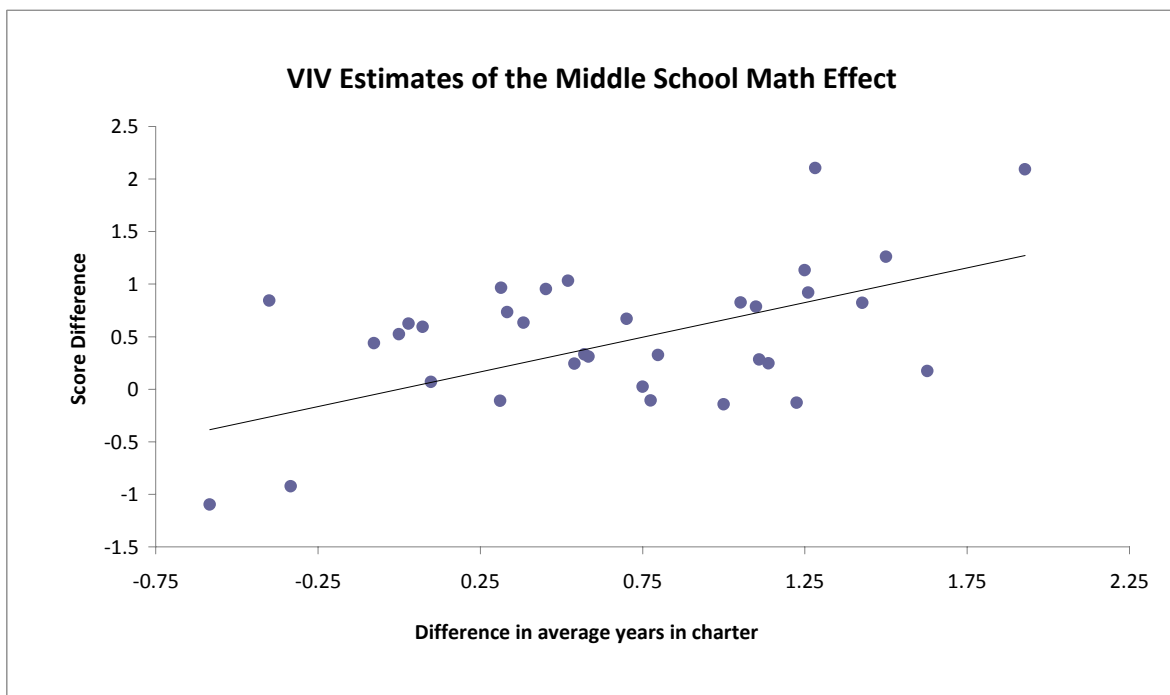
8. High school Writing Composition is for Grade 10 (2004-08).

Table 10: Estimates in and Out of the Lottery Sample

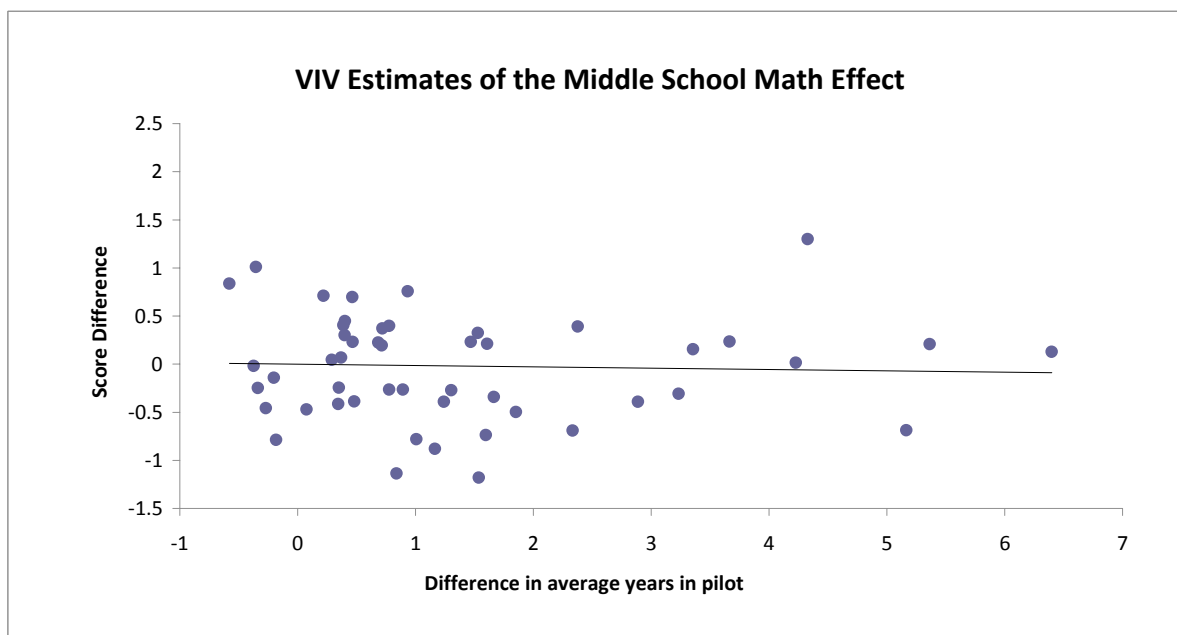
Level	Subject	Charters				Pilots			
		Lottery		Observational		Lottery		Observational	
		With Demographics (1)	With Baseline Scores (2)	In Lottery Sample (3)	Not in Lottery Sample (4)	With Demographics (5)	With Baseline Scores (6)	In Lottery Sample (7)	Not in Lottery Sample (8)
Elementary School	ELA	-	-	-	0.055*** (0.018)	0.064*** (0.024)	-	0.051* (0.026)	-0.033 (0.025)
		N			20,058	876		20,058	
	Math	-	-	-	0.037 (0.023)	0.060** (0.026)	-	0.079*** (0.028)	-0.059* (0.032)
		N			17,356	874		17,356	
Middle School	ELA	0.149*** (0.052)	0.144*** (0.044)	0.158*** (0.017)	0.082*** (0.014)	0.006 (0.043)	-0.035 (0.112)	-0.076*** (0.015)	-0.079*** (0.016)
		N	2,416	2,365	31,620	3,390	2,414	31,620	
	Math	0.405*** (0.066)	0.386*** (0.054)	0.312*** (0.028)	0.129*** (0.020)	-0.059 (0.048)	-0.251** (0.106)	-0.116*** (0.015)	-0.078*** (0.019)
		N	2,582	2,528	35,764	3,851	2,733	35,764	
High School	ELA	0.187*** (0.055)	0.186*** (0.049)	0.188*** (0.023)	0.134*** (0.022)	0.007 (0.073)	-0.053 (0.059)	0.141*** (0.018)	0.077*** (0.017)
		N	1,947	1,629	12,347	1,007	949	12,347	
	Math	0.274*** (0.071)	0.226** (0.060)	0.158*** (0.045)	0.140*** (0.032)	-0.011 (0.101)	0.007 (0.070)	0.139*** (0.036)	0.024 (0.023)
		N	1,929	1,892	15,868	996	983	15,868	
	Writing Topic	0.267*** (0.078)	0.281** (0.083)	0.253*** (0.041)	0.136*** (0.032)	0.173* (0.093)	0.151* (0.089)	0.242*** (0.019)	0.103*** (0.025)
		N	1,931	1,616	12,181	997	934	12,181	
	Writing Composition	0.168*** (0.062)	0.132** (0.059)	0.207*** (0.029)	0.134*** (0.024)	0.111 (0.086)	0.097 (0.080)	0.195*** (0.021)	0.104*** (0.018)
		N	1,931	1,616	12,181	997	934	12,181	

Notes: Columns (1) and (5) of this table report 2SLS coefficients from Table 4. Columns (2) and (6) report 2SLS coefficients from Table 5. These models include demographic and baseline test score controls. Observational models include separate variables for years in lottery sample pilot schools, lottery sample charter schools, non-lottery sample pilot schools, and non-lottery sample charter schools. These models also include the same covariates as used to construct the estimates in Table 9. For a given school level and test, columns (3), (4), (7), and (8) report coefficient estimates from the same regression. As in Table 9, observational models restrict the sample to students who were in a BPS or Boston charter school in the year of the relevant test. Standard errors, in parentheses, are clustered as described in the Table 4 notes.

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



A. Charter Schools



B. Pilot Schools

Figure 1. This figure plots treatment-control differences in test score means against treatment-control differences in years in charter (Panel A) or pilot (Panel B). The unit of observation is a charter or pilot application risk set ($N=34$ for charters and $N=51$ for pilots). The charter slope (unweighted) is .66, and the corresponding 2SLS estimate is .44. The pilot slope (unweighted) is -.006, while the corresponding 2SLS estimate is -.007. The pilot graph is produced after dropping one small imbalanced risk set.

Data Appendix

Data Sources

The data for this study come from four sources: the BPS assignment system, lottery records from individual charter schools, and two state-wide datasets, the SIMS and MCAS files. The BPS assignment data are documented in Abdulkadiroglu, Pathak, Roth and Sonmez (2006). Lottery coverage is documented below and in Table A1. Following the discussion of lottery data, we describe the state data sets and our matching process in detail.

Lottery Sample Coverage

Charter School Lotteries

- Of the four charters that enroll elementary school students, three had no usable lottery data. One K-8 school contributes to the middle school sample but is not used in the elementary school analysis because it added elementary grades only in 2006.
- Of the 11 charters that enroll middle school students, five contribute to the lottery analysis: Academy of the Pacific Rim, Boston Collegiate, Boston Preparatory, Edward Brooke and Roxbury Preparatory. Two non-included schools closed before or while the study was under way. Lottery data from four other schools were unavailable, either because they were not archived or lost.
- Of the eight (non-alternative) charter schools that enroll high school students, four contribute to the lottery analysis: City on a Hill, Codman Academy, MATCH, and Health Careers Academy (a Horace Mann Charter). Two non-included schools closed before or while the study was under way. Two schools do not admit via lottery in grade nine but rather admit students in grade five; they contribute to the middle school analysis.

Pilot School Lotteries

- Of the seven pilot schools that enroll elementary school students, five were over-subscribed in our sample window and contribute to the lottery sample: Lyndon, Mason Elementary, Mission Hill, Orchard Gardens, and Young Achievers.
- Of seven pilot middle schools admitting 6th graders, six were over-subscribed and contribute to our lottery sample: Harbor School, Josiah Quincy, Lilla G. Frederick, Lyndon, Mission Hill, Orchard Gardens.
- Of the six (non-alternative) 9-12 pilot high schools, two use lotteries and contribute to the lottery sample: Another Course to College, and TechBoston Academy. One 6-12 pilot was not over-subscribed. A single alternative high school that serves an older population was omitted from the study.

The applicant cohorts contributed by each school are detailed in Appendix Table A1.

State Data

The first state dataset is the Massachusetts Students Information Management System (SIMS). This contains snapshots of all students in public school in Massachusetts in October and at the end of the school year. The SIMS file includes demographic information on students, their current school, residence, and days of attendance. We work with the SIMS files for the 2001-2002 through the 2007-2008 school years. Schools are classified as pilot schools, charter schools, public schools, exam schools, and alternative/special education schools using a grade-specific algorithm described below.

The second state-wide dataset is the Massachusetts Comprehensive Assessment System (MCAS) database. The MCAS files have raw test scores for Math, English Language Arts (ELA), and Writing for grades 3-8 and grade 10 from spring 2002-2007. We standardized MCAS scores by subject, grade, and year to a combined BPS and Boston charter schools reference population. MCAS scores were merged to the SIMS file using unique student identifiers (the SASID). Students in the MCAS file but not in the SIMS file are excluded because they cannot be assigned to schools.

Cohorts and Test Years Analyzed

The cohorts and test years analyzed in this study are summarized in the table below. Baseline years are defined as the year a student attended 8th grade for the high school sample, the year a student attended 4th grade for the middle school sample, and the year a student attended kindergarten for the elementary school and K2-applicants-to-K-8-middle-school samples. There are no baseline test score for this last group of students, but we pulled demographics from the baseline years listed in the table.

	Elementary School	Middle (non-K2 applicants)	Middle (K2 lottery applicants)	High School
Outcome years	2005-2008	2003-2008	2004-2008	2004-2008
Lottery application Years	2002-2004	2002-2007	1997-2001	2002-2006
Baseline years	2002-2005	2001-2006	1998-2002	2002-2006

Notes: All years refer to Spring. Baseline grades are kindergarten for elementary school and K2 middle school applicants, 4th grade for all other middle school students and 8th grade for high school students. The 2003 middle school outcome year is used only for 2002 pilot middle school 6th grade applicants, as we use BPS demographics for these students from 2001 and we do not have state demographics for 2001.

Analysis Samples

Three matched files constitute the extracts used for statistical analysis. The first (the charter lottery extract) was constructed by linking charter applicants to SIMS and MCAS data; the second (the pilot lottery extract) was constructed by linking students with a first-choice pilot in the BPS assignment mechanism to SIMS and MCAS records; and the third and largest (the observational extract) uses the SIMS and MCAS matched file without regard to charter and pilot application. Table A2 shows how these three extracts were constructed. All three extracts are limited to students enrolled in BPS or a Boston charter school as of the relevant baseline date.

Classification of Schools

We define pilot schools as those listed on the pilot school network website at <http://www.ccebos.org/pilotschools/schools.html>. We define charter schools as those schools identified as such by the Massachusetts Department of Elementary and Second Education website (<http://www.profiles.doe.mass.edu>). As discussed in the text, Health Careers Academy is a Horace Mann charter high school. Boston Day and Evening Charter, also a Horace Mann charter, serves over-age students. This school, along with Greater Egleston Community High School, a pilot school that also

serves older students, was designated as an alternative/special education school rather than a charter or pilot and therefore omitted from our lottery analysis.

Placing Students at Schools

In the Massachusetts State SIMS file, school designation is grade-specific. When a student repeats grades, we kept the first school in grade. Students attending multiple schools in a given school year were assigned as follows:

- A student was assigned to a Boston charter or pilot school if any record for the grade shows attendance at a Boston charter or pilot.
- If a student attends multiple Boston charters or pilots, or both a Boston charter and pilot, the student was assigned to school for which the attendance duration is longest. Ties are broken randomly.
- Students attending multiple traditional BPS schools were assigned to the school of the longest attendance duration. Ties were broken randomly.

Table A.1: Charter and Pilot School Study Participation

	Years in Lottery Study	Initial offer data	Grade Range	Notes
	(1)	(2)	(3)	(4)
<i>I. Elementary School (3rd and 4th grades)</i>				
Charter Schools				
Boston Renaissance Charter Public School	x	x	K-6	School lottery data are unavailable for elementary school.
Conservatory Lab Charter School	x	x	K-5	School lottery data are unavailable for elementary school.
Edward Brooke Charter School	x	x	K-8	Became K-8 in 2006, so test scores are not used in elementary analyses.
Neighborhood House Charter School	x	x	K-8	School lottery data are unavailable for elementary school.
Pilot Schools				
Gardner Elementary School	x	x	K-5	All applicants in 2002-2004 are accepted.
Lee Academy	x	x	K-5	New school in 2004, all applicants in 2004 are accepted.
Lyndon K-8 School	2002-2004	x	K-8	
Mason Elementary School	2003-2004	x	K-5	All applicants in 2002 are accepted.
Mission Hill School	2002-2004	x	K-8	
Orchard Gardens K-8 School	2003-2004	x	K-8	No Kindergarten class in 2002 (new school).
Young Achievers K-8 School	2002-2004	x	K-8	
<i>II. Middle School (6th, 7th, and 8th grades)</i>				
Charter Schools				
Academy of the Pacific Rim Charter Public School	2005-2007	2005-2007	5-12	Insufficient lottery records from 2002-2004.
Boston Collegiate Charter School	2002-2006	2002-2006	5-12	
Boston Preparatory Charter Public School	2005-2007	2005, 2007	6-10	Under-subscribed in 2004. Initial offer only 2005, ever offer only 2006.
Boston Renaissance Charter Public School	x	x	K-6	School lottery data are unavailable.
Edward Brooke Charter School	2006	x	K-8	Became K-8 in 2006.
Excel Academy Charter School	x	x	5-8	Insufficient lottery records.
Frederick Douglass Charter School	x	x	5-12	Closed.
Neighborhood House Charter School	x	x	K-8	School lottery data are unavailable for elementary school.
Roxbury Preparatory Charter School	2002-2007	x	6-8	Ever offer only 2002-2007.
Smith Leadership Academy Charter Public School	x	x	6-8	Insufficient lottery records (appears under-subscribed in 03-04 and 04-05).
Uphams Corner Charter School	x	x	5-8	Insufficient admissions lottery records. Closed (under-subscribed).
Pilot Schools - 6th Grade Applicants				
Harbor School	2002, 2006-2007	x	6-8	All applicants in 2003-2005 are accepted.
Josiah Quincy Upper School	2004-2007	x	6-12	No Grade 6 class in 2002 and 2003 (new school).
Lilla G. Frederick Pilot Middle School	2004-2007	x	6-8	No Grade 6 class in 2002 and all Grade 6 applicants in 2003 are accepted.
Lyndon K-8 School	2004-2007	x	K-8	All Grade 6 applicants in 2002 and 2003 guaranteed admission.
Mission Hill School	2002-2003, 2007	x	K-8	All Grade 6 applicants in 2004-2006 guaranteed admission.
Orchard Gardens K-8 School	2003-2007	x	K-8	No Grade 6 class in 2002 (new school).
Young Achievers K-8 School	x	x	K-8	All Grade 6 applicants in 2002-2007 guaranteed admission.
Pilot Schools - K2 applicants				
Lyndon K-8 School	1997-1998, 2000-2001	x	K-8	All K2 applicants in 1999 were admitted.
Mission Hill School	1998-2001	x	K-8	
Orchard Gardens K-8 School	x	x	K-8	
Young Achievers K-8 School	1999-2001	x	K-8	Error in records 1998, all K2 applicants in 1997 guaranteed admission.
<i>III. High School (10th grade)</i>				
Charter Schools				
Academy of the Pacific Rim Charter Public School	x	x	5-12	No admissions lottery for students in Grade 9.
Boston Collegiate Charter School	x	x	5-12	No admissions lottery for students in Grade 9.
City On A Hill Charter Public School	2002, 2004-2006	2002, 2004-2006	9-12	Initial only for 2002 and 2004, records unclear in 2003.
Codman Academy Charter Public School	2004	2004	9-12	Initial offer data only for 2004.
Frederick Douglass Charter School	x	x	5-12	Closed.
MATCH Charter Public High School	2002-2006	2002-2004	9-12	Ever offer only 2005-2006.
Roxbury Charter High Public School	x	x	9-12	Closed (under-subscribed).
Horace Mann Charter Schools				
Boston Day and Even Academy Charter School	x		9-12, Alt	Included in observational study as alternative school.
Health Careers Academy Charter School	2003-2006	2003-2006	9-12	
Pilot Schools				
Greater Egleston Community High	x	x	9-12, Alt	Included in observational study as alternative school.
Another Course to College	2003-2006	x	9-12	No Grade 9 class in 2002 (new school).
Boston Arts Academy	x	x	9-12	Selective admissions.
Boston Community Leadership Academy	x	x	9-12	Selective admissions.
Fenway High School	x	x	9-12	Selective admissions.
Josiah Quincy Upper School	x	x	6-12	All students guaranteed admission.
New Mission High School	x	x	9-12	Selective admissions.
TechBoston Academy	2003-2006	x	9-12	No Grade 9 class in 2002 (new school).

Notes: Schools are excluded from the lottery study (those marked with x in column (1)) if they are undersubscribed, only have seats for students with guaranteed admissions, did not keep sufficient records, or if they do not have an entry lottery at the school level start grade as indicated in the notes. Frederick Douglass Charter School and Roxbury Charter High Public School were closed by the State Board of Education in the 2004-2005 school year. A school that spans K-8 is listed as both elementary and middle. K-8 pilot schools contribute to both the elementary and middle school samples.

Table A.2: Sample Selection

	Elementary (1)	Middle School (2)	High School (3)
<i>A. Charter Lottery Sample</i>			
Application cohorts	-	2002-2007	2002-2006
Applications to charter	-	3,940	4,726
Excluding applications not matched to state dataset	-	3,816	4,425
Excluding applications with sibling priority	-	3,593	4,309
Excluding applications that applied to a charter not subject to randomization	-	2,768	4,067
Transforming applications into one observation per student	-	2,275	3,028
Excluding students with no follow-up test score	-	2,020	2,418
Excluding students not in BPS or in a charter school in Boston at baseline (this also drops those without baseline demographics)	-	1,355	1,957
Excluding students with no baseline test scores	-	1,331	1,934
Number of follow-up LEA scores for students in BPS or a Boston charter at baseline	-	2,416	1,947
Number of follow-up Math scores for students in BPS or a Boston charter at baseline	-	2,582	1,929
<i>B. Pilot Lottery Sample</i>			
		<i>6th grade applicants</i>	<i>K2 applicants</i>
Application cohorts	2002-2004	2002-2007	1997-2001
Applied to pilot school as first choice	1,259	4,865	1,273
Excluding guarantees and students with sibling priority	1,051	3,545	1,007
Excluding students that applied to a pilot not subject to randomization	934	1,722	752
Excluding students with no follow-up test score	700	1,588	444
Excluding students not in BPS or in a charter school in Boston at baseline (this also drops those without baseline demographics)	573	1,474	443
Excluding students with no baseline test scores	-	1,298	-
Number of follow-up ELA scores for students in BPS or a Boston charter at baseline	876	2,645	745
Number of follow-up Math scores for students in BPS or a Boston charter at baseline	874	3,075	776
<i>C. Regression Sample</i>			
Baseline years	2002-2005	2002-2006	2002-2006
Students in BPS or in a charter school in Boston and demographics at baseline	17,293	23,976	25,588
Students in BPS or in a charter school in Boston and demographics at outcome	12,552	20,105	18,925
Excluding students with no follow-up test score	12,093	19,187	16,703
Excluding students with no baseline test scores	-	18,584	16,200
Number of follow-up test scores for students in BPS or a Boston charter at baseline and outcome	20,058	34,301	16,609
Number of follow-up test scores for students in BPS or a Boston charter at baseline and outcome	17,356	38,583	16,350

Table A.3: Differential Attrition Omitting Unbalanced Cohorts

		Charter			Pilot		
		Prop of non-offered with MCAS	Differential		Prop of non-offered with MCAS	Differential	
			Demographic Controls	Demographics + Baseline Scores		Demographic Controls	Demographics + Baseline Scores
Level	Subject	(1)	(2)	(3)	(4)	(5)	(6)
Elementary School	ELA	-	-	-	0.796	0.033 (0.037)	-
		N			686	1,085	-
	Math	-	-	-	0.796	0.032 (0.037)	-
		N			686	1,085	
Middle School	ELA	0.807	0.032 (0.022)	0.030 (0.022)	0.699	0.029 (0.024)	0.008 (0.026)
		N 819	2,663	2,600	2,625	4,596	2,778
	Math	0.817	0.033 (0.022)	0.032 (0.022)	0.702	0.030 (0.023)	0.009 (0.026)
		N 845	2,751	2,687	2,874	5,130	3,124
High School	ELA	0.776	0.027 (0.022)	0.020 (0.024)	0.754	0.029 (0.032)	0.048 (0.032)
		N 825	2,433	2,026	504	887	820
	Math	0.767	0.029 (0.023)	0.028 (0.023)	0.750	0.020 (0.033)	0.041 (0.032)
		N 825	2,433	2,375	504	887	864
	Writing Topic and Writing Composition	0.768	0.028 (0.023)	0.026 (0.024)	0.746	0.024 (0.033)	0.045 (0.033)
N 825		2,433	2,019	504	887	814	

Notes: This table has the same notes as Table 6, but uses samples that excludes the following cohorts: Charter 6th grade ELA (2005 5th grade applicants), Charter 7th grade ELA (2004 6th grade applicants), Charter 6th grade Math (2005 5th grade applicants and 2002 5th grade applicants), Charter 7th grade Math (2004 6th grade applicants), Pilot 10th grade ELA (2006 high school applicants), Pilot 10th grade Math (2006 high school applicants), Pilot 10th grade writing (2006 high school applicants). Standard errors, in parentheses, are clustered as described in the Table 4 notes.

Table A.4: Charter and Pilot Lottery Results Omitting Unbalanced Cohorts

Level	Subject	Charter				Pilots			
		First Stage (1)	Reduced Form (2)	2SLS (3)	2SLS w/demos (4)	First Stage (5)	Reduced Form (6)	2SLS (7)	2SLS w/demos (8)
Elementary School	ELA	-	-	-	-	2.852*** (0.193)	0.196** (0.078) 876	0.069** (0.027)	0.064*** (0.024) 876
		N							
	Math	-	-	-	-	2.858*** (0.194)	0.177** (0.078) 874	0.062** (0.027)	0.060** (0.026) 874
		N							
Middle School	ELA	0.950*** (0.122)	0.171*** (0.064) 2,238	0.179*** (0.066)	0.147*** (0.056) 2,238	1.385*** (0.190)	0.023 (0.065) 3,364	0.017 (0.047)	0.007 (0.043) 3,364
		N							
	Math	0.906*** (0.115)	0.389*** (0.071) 2,340	0.429*** (0.080)	0.408*** (0.070) 2,340	1.291*** (0.182)	-0.078 (0.069) 3,825	-0.061 (0.053)	-0.057 (0.048) 3,825
		N							
High School	ELA	0.724*** (0.131)	0.148*** (0.053) 1,839	0.204*** (0.066)	0.197*** (0.058) 1,839	0.591*** (0.101)	0.020 (0.053) 686	0.034 (0.087)	0.023 (0.091) 686
		N							
	Math	0.718*** (0.129)	0.233*** (0.061) 1,823	0.324*** (0.080)	0.282*** (0.075) 1,823	0.577*** (0.101)	-0.062 (0.075) 681	-0.107 (0.127)	-0.149 (0.127) 681
		N							
	Writing Topic	0.722*** (0.130)	0.185*** (0.066) 1,823	0.256*** (0.086)	0.248*** (0.081) 1,823	0.582*** (0.102)	0.091 (0.073) 678	0.156 (0.121)	0.146 (0.130) 678
		N							
	Writing Composition	0.722*** (0.130)	0.121** (0.053) 1,823	0.167** (0.069)	0.167** (0.066) 1,823	0.582*** (0.102)	0.058 (0.065) 678	0.099 (0.107)	0.092 (0.118) 678
		N							

Notes: This table repeats the analysis of Table 4, using the balanced-cohorts sample studied in Table A3. The sample is restricted to students with baseline demographic characteristics and excludes the following cohorts: Charter 6th grade ELA (2005 5th grade applicants), Charter 7th grade ELA (2004 6th grade applicants), Charter 6th grade Math (2005 5th grade applicants and 2002 5th grade applicants), Charter 7th grade Math (2004 6th grade applicants), Pilot 10th grade ELA (2006 high school applicants), Pilot 10th grade Math (2006 high school applicants), Pilot 10th grade writing (2006 high school applicants). Standard errors, in parentheses, are clustered as described in the Table 4 notes.

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

References

- Abadie, A. (2002). "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models." *Journal of the American Statistical Association*, 97(457): 284-292.
- Abadie, A. (2003). "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics*, 113(2): 231-263.
- Abdulkadiroglu, A. and Sönmez, T. (2003). "School Choice: A Mechanism Design Approach." *American Economic Review*, 93(3): 729-747.
- Abdulkadiroglu, A. Pathak, P. A., Roth A. E. and Sönmez, T. (2006). "Changing the Boston School Choice Mechanism." NBER Working Paper 11965.
- Abdulkadiroglu, A, Pathak, P. A., and Roth, A. E. (2009). "Strategyproofness versus Efficiency in Matching with Indifferences: Redesigning the NYC High School Match." *American Economic Review*, 99(5), forthcoming.
- Angrist, J. (1991). "Group-data Estimation and Testing in Simple Labor Supply Models." *Journal of Econometrics*, 47(2-3): 243-266.
- Angrist, J. (1997). "Conditional Independence in Sample Selection Models." *Economics Letters*, 54(2): 103-112.
- Angrist, J., Bettinger, E., Bloom, E., King, E. and Kremer, M. (2002). "Vouchers for Private Schooling in Columbia: Evidence from a Randomized Natural Experiment." *American Economic Review*, 92(5): 1535-1558.
- Angrist, J., Imbens, G., and Rubin, D. (1996). "Identification of Causal Effects Using Instrumental Variables." *JASA Applications* invited paper, with comments and authors' rejoinder, *Journal of the American Statistical Association*, 91 (434): 444-472.
- Angrist, J. and Pischke, J.S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.
- Berends, M., Mendiburo, M. and Nicotera, A. (2008). "Charter Schools Effects in an Urban School District: An Analysis of Student Achievement Growth," Working Paper, Vanderbilt University
- Bettinger, E. (2005). "The Effect of Charter Schools on Charter Students and Public Schools." *Economics of Education Review*, 24(2): 133-147.

- Bifulco, R. and Ladd, H. (2006). "The Impact of Charter Schools on Student Achievement: Evidence from North Carolina." *American Education Finance Association*, 1(1): 50-90.
- Booker, K., Gilpatric, S. M., Gronberg, T., and Jansen, D. (2007). "The Effect of Charter Schools on Traditional Public School Students in Texas: Are Children Who Stay Behind Left Behind?" *Journal of Urban Economics*, 64(1): 123-145.
- Booker, K., T.R. Sass, B. Gill, and R. Zimmer (2008). "Going Beyond Test Scores: Evaluating Charter School Impact on Educational Attainment in Chicago and Florida," RAND Working Paper WR-610-BMG, August.
- Center for Collaborative Education (2006). "Description of the Boston Pilot Schools Network." Accessed November 25, 2009 at http://www.ccebos.org/pilotschools/pilot_qa.doc.
- Clark, D. (2009). "The Performance and Competitive Effect of School Autonomy." *Journal of Political Economy*, 117(4): 745-783.
- CREDO. (2009). *Multiple Choice: Charter Performance in 16 states*. Stanford University: Center for Research on Education Outcomes.
- Cullen, J. B., Jacob, B. A. and Levitt, S. (2006). "The Effect of School Choice on Participants: Evidence from Randomized Lotteries." *Econometrica*, 74(5): 1191-1230.
- Dobbie, W. and Fryer, R. (2009). "Are High Quality Schools Enough to Close the Achievement Gap? Evidence from a Social Experiment in Harlem." NBER Working Paper 15473.
- Eberts, R. W. and Hollenbeck, K. (2002). *Impact of Charter School Attendance on Student Achievement in Michigan*. Upjohn Institute Staff Working Paper No. 02-080
- Hanushek E. A., Kain, J. F., Rivkin, S. G., and Branch, G. F. (2007). "Charter School Quality and Parental Decision Making with School Choice." *Journal of Public Economics*, 91(5-6): 823-48.
- Hastings, J., Kane, T., and Staiger, D. (2008). "Heterogeneous Preferences and the Efficacy of Public School Choice." NBER Working Paper 2145 and Working Paper 11805 combined.
- Hoxby, C. M. (2000). "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90(5): 1209-1238.
- Hoxby, C. M. (2007). "Does Competition Among Public Schools Benefit Students and Taxpayers? Reply." *American Economic Review*, 97(5): 2038-2055.
- Hoxby, C. M. and Murarka, S. (2009). "Charter Schools in New York City: Who Enrolls and How They Affect Student Achievement," NBER Working Paper 14852.

- Hoxby, C. M., Murarka, S., and Kang, J. (2009). *How New York City's Charter Schools Affect Achievement, August 2009 report*. Second report in series. Cambridge, MA: New York City Charter Schools Evaluation Project.
- Hoxby, C. M. and Rockoff, J. (2005). "The Impact of Charter Schools on Student Achievement," *Harvard Institute of Economic Research Working Paper Series*.
- Imbens, G. and Angrist, J. (1994). "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467-475.
- Jacob, B. and Lefgren, L. (2007). "What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers." *Quarterly Journal of Economics*, 122(4): 1603-1637.
- Manzo, K. K. (2007). "Arts Education Building Steam in L.A. Area Schools." *Education Week*, 02774232, 26(41), 6/13.
- Massachusetts Department of Elementary and Secondary Education. (2009a). *Massachusetts charter schools fact sheet*. Retrieved online on September 30, 2009 from: <http://www.doe.mass.edu/charter/factsheet.pdf>.
- Massachusetts Department of Elementary and Secondary Education. (2009b). *2008-2009 Enrollment by Grade Report*. Retrieved online on September 30, 2009 from: http://profiles.doe.mass.edu/state_report/enrollmentbygrade.aspx?mode=school&orderBy=.
- Matthews, J. (2009). *Work Hard. Be Nice. How Two Inspired Teachers Created the Most Promising Schools in America*. Chapel Hill: Alonguin Press.
- Merseeth, K., Robert, J., Cooper, K., Casey, M. C., and Valant, J. (2009). *Inside Urban Charter Schools: Promising Practices and Strategies in Five High-performing Charter Schools*. Cambridge: Harvard Education Publishing.
- National Alliance for Public Charter Schools. (2009). *Public Charter School Dashboard*. Washington D.C. Retrieved on September 21, 2009 from: <http://www.publiccharters.org/files/publications/DataDashboard.pdf>.
- Rothstein, J. (2006). "Good Principals or Good Peers? Parental Valuations of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *American Economic Review*, 96(4): 1333-1350.
- Rothstein, Jesse. (2007). "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment on Hoxby (2000)." *American Economic Review*, 97(5): 2026-2037.

- Rothstein, R. (2004). *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*. Columbia: Teachers College Press.
- Rouse, C. (1998). "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics*, 113(2): 553-602.
- Thernstrom, A. and Thernstrom, S. (2003). *No Excuses: Closing the Racial Gap in Learning*. New York: Simon & Schuster.
- Vaznis, J. (2009). "Backers Seek End to Charter Cap," *The Boston Globe*, August 5.
- Wilson, S. F. (2008). "Success at Scale in Charter Schooling," Working Paper 2008-02 from the American Enterprise Institute Future of American Education Project.