Who Benefits from KIPP?

Joshua D. Angrist Susan M. Dynarski Thomas J. Kane
Parag A. Pathak and Christopher R. Walters*

May 2010

Abstract

Charter schools run by the Knowledge is Power Program (KIPP) are emblematic of the No Excuses approach to public education. These schools feature a long school day, an extended school year, selective teacher hiring, strict behavior norms and a focus on traditional reading and math skills. We use applicant lotteries to evaluate the impact of KIPP Academy Lynn, a KIPP school that is mostly Hispanic and has a high concentration of limited English proficiency (LEP) and special education students, groups that critics have argued are served poorly by charter schools. The results show overall achievement gains of 0.35 standard deviations in math and 0.12 standard deviations in reading for each year spent at KIPP Lynn. Boys and girls realize similar math gains, while reading gains are largest for boys. LEP students, special education students, and those with low baseline test scores benefit more from time spent at KIPP than do other students. The average reading gains are driven almost completely by special education and LEP students, whose reading scores rise by roughly 0.35 standard deviations for each year spent at KIPP Lynn.

*Angrist and Pathak, MIT and NBER, angrist@mit.edu, ppathak@mit.edu. Walters: MIT, crwalt@mit.edu. Dynarski: University of Michigan and NBER, dynarski@umich.edu. Kane: Harvard University and NBER, kaneto@gse.harvard.edu. We are grateful to Josh Zoia and Esther Vargas at KIPP Academy Lynn and Carrie Conaway at the Massachusetts Department of Elementary and Secondary Education for data, patience and assistance. Sarah Cohodes, Jon Fullerton, and Emily House provided valuable assistance. Eric Bettinger and seminar participants at Australia National University, Brown University, the Federal Reserve Bank of New York, Simon Fraser University, the University of British Columbia, and the University of Wollongong provided helpful comments.

I. Introduction

Charter schools are publicly funded but operate with almost as much autonomy as private schools. Whether charter schools improve academic achievement, and what types of students benefit most from charter attendance, are controversial questions among both researchers and policymakers. Recent quasi-experimental evaluations using school lotteries show charter schools have generated large overall achievement gains in Boston and New York (Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak 2009; Hoxby and Murarka 2009), while an instrumental variables analysis exploiting charter proximity offers evidence for increased high school graduation rates (Zimmer, et al, 2009). Dobbie and Fryer's (2009) study of the Harlem Children's Zone also shows large achievement effects using lotteries for the two charter schools in the Zone, though these results may arise in part from other social programs and services. Evaluations based solely on statistical comparisons have been more mixed (see, e.g., CREDO 2009).

This body of evidence suggests that some charter schools have produced impressive results. But charter schools are heterogeneous by nature. Their culture and curriculum often reflect the unique personalities of founders and teachers. From both an economic and policy standpoint, a key unanswered question is whether the positive results that have been detected by the existing research can be replicated in other places and populations. A second question often raised in the charter debate is whether struggling students benefit from a charter education even where overall effects appear to be positive.

We address the first of these issues by evaluating a widely-replicated charter model. Charter management organizations (CMOs) are networks of charter schools that provide curriculum development, staff training, and operational support, usually in the context of a specific program or philosophy. The nation's largest charter management organization is the Knowledge is Power Program (KIPP), which serves a primarily low-income and minority student population. KIPP is expanding steadily: since 1995, KIPP has opened 80 schools in 19 states, and now serves 21,000 students. KIPP emphasizes traditional math and reading skills, the development of a strong student work ethic, strict behavior norms, long school days and an extended school year. KIPP schools exhibit a large measure of program standardization, maintained through central and regional offices that provide staff training, operational support, and curricular materials (Whitman, 2008).

This paper reports results from a quasi-experimental evaluation of the effects of attending

a KIPP school. The setting for our study is KIPP Academy Lynn, a middle school founded in 2004 in Lynn, Massachusetts. KIPP Lynn has been substantially over-subscribed since 2005. As required of all over-subscribed Massachusetts charter schools, KIPP Lynn uses a lottery to select its incoming class. The KIPP admissions lottery generates instrumental variables that we use eliminate selection bias in estimates of the causal effects of KIPP attendance. As far as we know, ours is the first study to use lotteries to evaluate a KIPP school. Results for KIPP Academy Lynn are of special interest because they may be relevant for the broader population of KIPP schools and students. These results may also apply to other charter schools that follow the KIPP model, which some have called "No Excuses" (Carter 2000; Thernstrom and Thernstrom 2003). Like KIPP, No Excuses schools focus on traditional math and reading skills and the development of a strong student work ethic. They feature a long school day and year, selective teacher hiring, and strict behavior norms.

Motivated by the concern that charter schools benefit relatively high-achieving and motivated students, leaving those who need the most support behind (a criticism voiced recently in United Federation of Teachers, 2010), we look at differential effects for special education students and those with limited English proficiency. KIPP Lynn is unusual among charter schools in that it enrolls a high proportion of limited English proficient (LEP) and special education students, and so affords the opportunity to evaluate achievement gains for these important groups. We also take a look at the impact of KIPP attendance on the distribution of student test scores.[1]

Our results show overall reading score gains of about 0.12 standard deviations (hereafter, σ) for each year a student spends at KIPP, with significantly larger gains for special education and LEP students of about 0.3-0.4σ. Students attending KIPP gain an average of 0.35σ per year in math; these effects are slightly larger for LEP and special education students. Results from a specification that interacts KIPP attendance with students' baseline (4th grade) scores suggests that effects are largest for those who start out behind their peers. Male and female students gain about equally in math, while boys benefit slightly more than girls in reading. Finally,

[1]A few studies have looked at overall effects of KIPP. See, for example, the studies linked at http://www.kipp.org/01/independentreports.cfm. Hoxby and Murarka (2009), a lottery-based evaluation of charter schools that includes some from KIPP in New York City, does not report KIPP results separately. Our recent AER Papers and Proceedings article (Angrist, *et al.* 2010) presents a brief overview of some of the results reported here, but does not explore the KIPP schooling model in depth or estimate effects among subgroups that have been at the heart of the charter debate.

an examination of Massachusetts Comprehensive Assessment System (MCAS) performance categories (similar to quartiles) shows that KIPP Lynn boosts achievement primarily by moving students up from the lowest group. Together, therefore, the findings reported here suggest that KIPP Lynn benefits the weakest students most.

The next section provides some background on KIPP schools and Lynn. Following that, Section III describes the data and our lottery-based estimation framework. Section IV presents the results, including estimates in LEP and special education subgroups and from models with baseline score interactions. The paper concludes in Section V.

II. Background

A. The KIPP Schools

KIPP was started in Houston and New York City in 1995 by veterans of Teach for America, a program that recruits graduates of elite colleges to teach in low-performing districts (Mathews, 2009). The first KIPP schools operated as alternative programs within traditional public school districts, with KIPP staff employed by the district and covered by its seniority rules and salary schedules. In New York City, KIPP staff were initially covered by the district's collectively bargained union contract. In both Houston and New York, the KIPP founders negotiated with district leadership for limited autonomy in curriculum, staffing, and hours of instruction. This model – negotiated autonomy within a traditional school district – is still followed by a few KIPP schools that operate under contract with a state or district. Today, however, most KIPP schools are charters (Childress and Marino, 2008). KIPP Houston became a charter school in 1998, and KIPP NYC did the same in 1999 (Leschly, 2008).

KIPP has expanded steadily over the last fifteen years, opening an additional 80 schools in 19 states. Most are middle schools, covering grades five through eight. Recently, eleven KIPP high schools and 16 elementary schools have been added to the network, usually attached to a pre-existing middle school. KIPP now serves 21,000 students, mostly low-income minority students who qualify for a free or reduced-price lunch. KIPP is currently the nation's largest CMO.

KIPP schools are characterized by a high level of standardization. For example, all KIPP students are expected to adhere to a behavioral code which governs comportment within and

between classes. Parents or guardians, students, and teachers are asked to sign a "Commitment to Excellence," a promise to come to school on time and work hard, among other things. KIPP schools use modest financial incentives, distributed weekly, to reward effort and good behavior, with payments for completed homework, and passing test scores. Paychecks in the form of "K dollars" can be redeemed for items from a school store (notebooks, t-shirts, snacks). Recent experiments suggest that short-term rewards of this form can improve student achievement (Fryer, 2010).[2]

Schools in the KIPP network, including KIPP Academy Lynn, have an unusually long school year. School starts in August, and runs on many Saturday mornings. The school day starts at 7:30 am and ends at 5:00 pm. This works out to about 1,900 hours of instruction a year at KIPP Lynn, as compared to about 1,250 in Lynn's traditional public schools. KIPP students are encouraged to call teachers at night with questions about homework.

As the number of KIPP schools has grown, the founders have tried to develop and maintain the KIPP brand, while still allowing local schools a high degree of autonomy. The current KIPP organizational structure, resembling a retail franchise, developed with the financial and logistical support of the founders of The Gap, Inc. Central and regional offices train school principals, recruit teachers, and provide operational support. KIPP schools pay a licensing fee to the central organization, amounting to one percent of revenue in the first year of a school's operation and three percent in subsequent years (Leschly, 2003). KIPP retains the right to withdraw the use of its name if it determines that a school is not meeting the network's standards. Between 2001 and 2007, five KIPP schools closed and two left the KIPP network, while continuing to operate as independent charter schools (Childress and Marino, 2008).

The KIPP Foundation bears many of the costs of starting a school, scouting out new locations and training new principals. KIPP headquarters also provides operational support, in the form of advice on human resources management, legal issues, procurement and budgeting (Leschly, 2008). KIPP principals receive a year of salaried training from KIPP. In 2002, there were 420 applicants for 20 slots in the training program. Principals-in-training spend six weeks at the Haas School at Berkeley, trained by instructors from the business schools at Berkeley and Stanford. They then "shadow" a principal at an established KIPP school. The remainder of the year is spent preparing their new school. New principals receive coaching and mentoring

[2]See http://www.kipp.org/about-kipp, accessed May 2, 2010.

in the first few years of operation (Leschly, 2003).

B. KIPP Academy Lynn and the Lynn Public Schools

Lynn, Massachusetts is a city of 90,000 located ten miles northeast of Boston. The city was a manufacturing center from its earliest days, with footwear driving the economy until the nineteenth century, when electrical manufacturing took center stage. General Electric's Lynn plant built the country's first jet engines during World War II, and a GE plant is still located in the city. Even in its heyday, the city had a colorful reputation. (A well-known New England ditty begins: "Lynn, Lynn, city of sin, you never come out the way you came in.") Crime rates in Lynn are among the highest in the state. When manufacturing declined, poverty rose and income fell. The city's 2007 poverty rate stood at 21 percent, more than twice the Massachusetts average.

Lynn Public Schools (LPS) enroll about 13,000 students a year in its nineteen elementary schools, four middle schools and five high schools. KIPP is the only charter school in the city. About 1,600 children, or 11 percent of the school-age population, attend private schools in Lynn. While the population of Lynn is more than two-thirds white, most of the 13,000 schoolchildren in the city are nonwhite. Nearly 80 percent of Lynn's students are eligible for a free or reduced-price lunch. All of the city's public schools fell short of the achievement standards laid out in the federal No Child Left Behind (NCLB) act in 2009. Middle school students are failing to meet state standards for adequate yearly progress (AYP) in every subgroup tracked by NCLB. Fifth-grade students in LPS score about a third of a standard deviation below the Massachusetts average on standardized tests, a fact documented in the first column of Table 1, which reports standardized scores of LPS students using a state reference population along with other descriptive statistics.

KIPP Lynn, which opened in the Fall of 2004, currently serves about 300 students in grades five through eight (The state recently approved KIPP's request to open a high school in Lynn.) KIPP Academy Lynn is governed by a board drawn from the local community. The principal ("school leader") serves at the will of the board. None of the school staff are employed by the national KIPP organization. The KIPP name is licensed from the national organization, which can revoke the license if it considers the school out of compliance with KIPP goals and standards. KIPP staff attend an annual nationwide conference, and participate in teacher and

principal training provided by the national organization.

Like many KIPP schools, KIPP Lynn uses a formal incentive system to encourage academic effort and compliance with behavioral norms. These awards, called paychecks, serve to keep parents informed since a paycheck cannot be cashed until a parent has signed it. The school's 2003 application for charter notes: "While students can earn up to K$50 each week, a minimum of K$35 on their paycheck tells a parent that the student is meeting the minimum behavioral standards required by the school (KIPP Academy Lynn Charter School, 2003)."

The school closely tracks students' academic performance. KIPP students take the Stanford 10, a widely used standardized test, each summer before school starts. These tests are used to assess the curricular needs of a cohort and to plan interventions for individual students. Student performance throughout the year is discussed in staff meetings. Students are also tested at the end of the year, again with the Stanford 10 (KIPP Academy Lynn Charter School, 2003). The state of Massachusetts tests students in all grades covered by KIPP, using the MCAS. Individual student MCAS scores are not immediately made available to schools, however, which limits their usefulness for planning.

Many KIPP Lynn teachers are graduates of Teach for America. They are recruited through the national KIPP organization, which provides a centralized clearinghouse for teachers interested in teaching at a KIPP school. Resumes are screened and forwarded to KIPP Lynn, which interviews applicants by phone and invites finalists to teach a sample lesson at the school. To fill five teaching slots at KIPP Lynn in 2007, the national organization screened 5,000 resumes and asked 250 applicants to complete a detailed questionnaire. Forty applicants were interviewed by phone and 25 invited to teach a sample lesson (Massachusetts Department of Elementary and Secondary Education, 2008).

KIPP Lynn teachers are much younger than their LPS colleagues: 88 percent are 40 or under, compared to 29 percent in LPS. Perhaps reflecting their age, KIPP teachers are far less likely to be licensed in their teaching assignment (26 percent, compared to 98 percent in LPS). While KIPP Lynn teachers are younger and less likely to be licensed that LPS teachers, they are paid more: in 2008, average teacher salaries were $69,353 at KIPP Lynn and $60,523 in LPS (KIPP teachers work longer hours). The KIPP Lynn and LPS student-teacher ratios are similar, at around 14, implying that KIPP Lynn spends a higher proportion of its budget on instructional salaries than LPS. The school employs four full-time special education teachers, as well as a

part-time, licensed ELL teacher. Eight staff members are fluent in Spanish (Massachusetts Department of Elementary and Secondary Education, 2008).[3]

KIPP Lynn applicants have (pre-lottery) math and reading scores that are 0.39σ and 0.44σ below the state average and somewhat below the LPS average. This can also be seen in Table 1, the first two columns of which show average characteristics (including 4th grade scores) of students in grade five, the entry grade for KIPP Lynn. About a fifth of both the LPS and KIPP Lynn populations are designated LEP, while a fifth are categorized as special education students.

Most KIPP Lynn students live in Lynn and would otherwise attend an LPS school. A handful of applicants come from Catholic schools or charter schools outside the city. Like most other Massachusetts charter schools, KIPP Lynn is funded primarily through tuition paid by students' sending districts. Tuition is typically set to match sending districts' average per-pupil expenditure. This amount is offset by state subsidies to the sending district when a student first transfers out of the regular public school district. KIPP Lynn spent about $11,500 per pupil in fiscal year 2008, net of rental and capital costs that add another $2,000. KIPP Lynn is currently located in a rented former church. KIPP Lynn received about $11,000 per pupil from the town of Lynn, with the remaining expenditures covered by donations and grants. Average per pupil expenditure in LPS schools was about $13,000 in 2008. Like all new KIPP schools, KIPP Lynn received substantial logistical support from the KIPP Foundation at startup.[4] As at LPS schools, the state and LPS cover busing costs for transportation-eligible students at KIPP.

Statewide regulations require that Massachusetts charter schools use a lottery when over-subscribed. KIPP Lynn was under-subscribed when it opened in the Fall of 2004, and only marginally over-subscribed in 2005. More recently, however, more than 200 students have applied for about 90 seats. The 2005-8 admissions lotteries are used here to develop a quasi-

[3]Statistics in the paragraph are calculated from data available at http://profiles.doe.mass.edu, accessed January 28, 2010. MA charter school teacher salaries are available at http://finance1.doe.mass.edu/charter. Public school salaries are at http://finance1.doe.mass.edu/statistics/.

[4]KIPP Lynn financial statistics are from Massachusetts Department of Elementary and Secondary Education. 2009. "FY2008 Charter School End of Year Financial Report Summary." Available at http://finance1.doe.mass.edu/charter/08CSEOYFR_Summary.xls, accessed May 3, 2010. LPS financial statistics are from Massachusetts Department of Elementary and Secondary Education. 2010. "FY08 Expenditures Per Pupil, All Funds, Summary By Function." Available at http://finance1.doe.mass.edu/schfin/statistics/function08.xls, accessed May 3, 2010.

experimental research design. These randomized lotteries allow us to estimate the causal effect of KIPP Lynn on achievement, solving the problem of selection bias that plagues most studies of school effectiveness.

C. Previous Research

The question of KIPP effectiveness often arises in the debate over whether schools alone can address achievement gaps between racial and income groups. In a study of the racial achievement gap titled *No Excuses*, Thernstrom and Thernstrom (2003, p. 43) credit KIPP with impressive gains among low-income students, and conclude that "truly radical education innovation can change the lives of inner-city students, whatever their race or ethnicity." Others argue that KIPP's rigorous requirements attract mostly families whose children would succeed anyway: "KIPP students . . . enter with substantially higher achievement than the typical achievement of schools from which they came. ...[T]eachers told us either that they referred students who were more able than their peers, or that the most motivated and educationally sophisticated parents were those likely to take the initiative to . . . enroll in KIPP (Carnoy, *et al.*, 2005)

A long literature attempts to understand the effect of schools on the achievement of low-income, nonwhite students. As suggested by the concerns of Carnoy, *et al.*, (2005), the key challenge in this literature is selection bias: students at different types of schools likely differ in many ways, both observed and unobserved. Catholic schools, which (like charter schools) have traditionally served a low-income, urban student population, have received particularly intense scrutiny. Research on Catholic schools has followed an arc similar to that on charter schools, with initial studies using statistical controls to control for selection (Coleman 1966) and more recent research employing instrumental variables methods.[5]

Charter school evaluations that rely on statistical controls to deal with selection bias have produced mixed results. A study using propensity score methods to match charter school students to similar students in nearby traditional public schools concludes that charter schools are no better on average (CREDO 2009). A recent study of KIPP schools in the Bay Area, which also used a propensity score approach, concluded that KIPP appears to increase test scores but this may be driven by high exit rates by weaker students. Following the discussion of achievement results, we look at the impact of attrition on our evaluation of KIPP.

[5]Altonji, Elder and Taber (2005a, 2005b) summarize and critique this literature.

The few charter evaluations that have used lotteries to identify causal effects report substantial achievement gains for charter students. Lottery estimates for middle and high schools in Boston, many of which use the No Excuses model, show test score gains on the order of 0.2 to 0.4 standard deviations for each year a child spends in a charter school (Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak 2009). Hoxby and Murarka (2009) find smaller though still substantial effects from a more heterogeneous sample of schools in New York City. Dobbie and Fryer (2009) uses lottery data to examine two of the Harlem Children's Zone (HCZ) charter schools, with results similar to those we find here for KIPP. However, children in the Harlem Children's Zone receive a barrage of unusually intensive and expensive social services, and Dobbie and Fryer (2009) are unable to disentangle the effect of these programs from that of the charter schools. By contrast, KIPP Lynn students receive non-school services typical of those received by children in any urban district.

III. Data and Empirical Framework

A. Data

Data set construction for this study began with lists of KIPP Lynn applicants for the Fall of 2005 through 2008. Most of these applicants are fourth graders coming from LPS, applying to enter KIPP Lynn in fifth grade. The applicant records include an applicant's name, date of birth, previous school and grade, and contact information for a parent, guardian, or the name of a sponsoring organization such as the Lynn Boys Club. The first five rows of Appendix Table A1 summarize the raw applicant data. If a student applied to KIPP Academy Lynn more than once, only the first application is included in our analysis. We exclude late applicants (who applied after the lotteries), as well as siblings and any students who went directly onto the waiting list (these are mostly 6th grade applicants in early cohorts). Imposing these restrictions reduces the number of applicant records from 629 to 542.

We matched KIPP applicant lottery records to the Massachusetts Student Information Management System (SIMS), a database with demographic and attendance information for all public school students in the state. As shown in Appendix Table A2 (and discussed below), 91 percent of KIPP applicants were located in the SIMS database. The SIMS variables of interest include grade, year, name, town of residence, date of birth, sex, race, special education

(SPED) and limited English proficiency (LEP) status, free/reduced price lunch, and school attended. The SIMS records capture data at multiple points within the school year. If a student is classified as SPED, LEP, or qualified for free/reduced price lunch at any time within a school-year-grade, then he or she is coded that way for the entire school-year-grade record.

KIPP lottery winners were more likely to be matched to the SIMS than losers, a difference of about 11 percent. Among students whose application indicates that they attended an LPS school at the time of application, however, the differential is only 0.029 (s.e.=0.027). The SIMS match differential is therefore mostly driven by the tendency of private (mostly Catholic) school applicants to remain in private school if they lose the KIPP lottery (about 15 percent of Lynn middle school students attend private schools). As we show later, limiting the analysis to LPS applicants yields effects very similar to those in the full set of applicants. This suggests that the difference in SIMS match rates between winners and losers does not bias our findings.

In Massachusetts, third through eight graders take MCAS exams in math and English language arts (ELA). Our analysis uses math and ELA scores from the Spring of 2006-9. We normalized these scores to a statewide mean of zero and standard deviation of one by subject, grade and year and then matched them to KIPP applicants' SIMS records. The MCAS and SIMS files were merged by grade, year and a state student identifier known as the SASID. In grades 4-8, 99.3 percent of MCAS scores were matched to a student in the SIMS. Scores that could not be matched to the SIMS were dropped.

The resulting dataset includes only applicants subject to random assignment, omitting those with siblings already enrolled in KIPP (who are guaranteed a slot), late applicants (who miss the lotteries), older applicants (who enter late if at all), as well as a few students with missing demographic data. There are 457 matched students with demographic data subject to random assignment. Dropping a further 38 applicants for whom we would not expect to have scores by the Spring of 2009 (by virtue of their grade and date of application) leaves 419 with post-lottery MCAS scores.

Table 2 lists the lottery cohorts contributing to this study, the share of randomized applicants who won entry to KIPP, and the share that attended. Of the 457 matched students subject to random assignment, 69 percent were offered a spot at KIPP Lynn and 54 percent enrolled. As the table makes clear, the earliest cohorts to apply to KIPP Lynn contribute more data to our analysis than do more recent cohorts, since earlier cohorts have taken more tests.

For example, from 2006 through 2009, the 2005 applicant cohort was tested in 5th through 8th grade, while the 2008 applicant cohort was tested in 5th grade only. Overall, lottery winners spent an average of about 1.85 years at KIPP Lynn in our sample period (as shown in last column of Table 2), but the 2005 cohort spent an average of 2.6 years in KIPP while the 2008 cohort spent only 0.7 years at KIPP. Since our test score data is weighted toward the earliest KIPP cohorts, the results reported here should be interpreted as the effect of a relatively new school on its first cohorts of students.

Table 1 reports descriptive information for 5th graders attending LPS and KIPP Lynn during our study period, as well as for the estimation sample of KIPP Lynn applicants. KIPP Lynn applicants look much like the students who ultimately enroll in KIPP. Importantly, lottery winners and losers appear to be similar, a fact documented in column (4) of Table 1, which reports differences in demographic characteristics and baseline scores between those who win and lose the lottery.[6] By virtue of random assignment, we would expect these differences to be small. In practice, however, our reconstruction of the KIPP lotteries is based partly on spreadsheets with some missing and incomplete information. In particular, for some applicants, lottery status as determined by the original random draw was later over-written with enrollment status. We fixed most such mistakes via a detailed student-by-student review of applicant histories in consultation with school staff. We believe this review succeeded in reconstructing the original lottery assignment data. Encouragingly, regression estimates with pre-treatment characteristics on the left-hand side show only one significant difference between winners and losers (proportion Asian), and the F-statistic from a joint test of balance on all observable characteristics gives little cause for concern. Differences in baseline scores shrink further when adjusted for the demographic variables in the upper half of Table 1, as can be seen in column (5).[7]

While Table 1 shows that lottery winners and losers are similar at the time of the lottery,

[6]KIPP Lynn lotteries sequenced all applicants. For example, if there were 150 applicants for 80 non-sibling seats, all 150 were randomly sequenced, and the first 80 immediately offered a seat. The remaining applicants constituted that cohort's wait list. Our lottery status variable indicates children who were offered a seat at KIPP either on the night of the relevant admissions lottery or any time after, up until the first day of school that year (early or mid-August). Late offers - i.e., those made after lottery day and before the first day of school - were made to students based on the original sequence as determined on lottery day.

[7]The estimates in column (4) control for year and grade of application (except for the row for application grade). The column (5) models add all demographic controls with the exception of LEP status to the regressions for LEP, baseline scores, and 4th grade application status.

subsequent attrition can produce important differences in the follow-up sample if the attrition process is non-random. Of particular concern is the possibility that lottery losers might be more likely to be missing. The MCAS is a statewide test (and we have all scores in the state) but any KIPP applicants who enroll in private school or leave Massachusetts are lost to follow-up. Selective attrition of this sort has the potential to effectively undo the apples-to-apples nature of comparisons based on the original randomization (for example, losers who are lost to follow-up might come from stronger family backgrounds than winners lost to follow-up).

Table 3 probes for evidence of differential attrition between lottery winners and losers. Taking account of the years in which our applicant cohorts should have been tested, we expect to observe 971 test scores after the lottery (Appendix Table A3 lists the number of test scores expected and observed for each applicant cohort). Among those who lose the lottery, about 85% of expected scores were found. This rate is about five percentage points higher for those who win the lottery, a difference that shrinks and become insignificant with the inclusion of demographic controls (these regressions include the same controls as the first stage equations discussed in the next section). Overall, the small follow-up differentials documented in Table 3 seem unlikely to impart substantial selection bias in our impact analysis.[8]

B. 2SLS Strategy

The causal effect of KIPP Lynn attendance on test scores is modeled as a function of time spent attending KIPP Lynn using the equation

$$y_{igt} = \alpha_t + \beta_g + \sum_j \delta_j d_{ij} + \gamma' X_i + \rho s_{igt} + \epsilon_{igt}, \tag{1}$$

where y_{igt} denotes the scores of student i tested in year t in grade g. The variable s_{igt} records calendar years spent at KIPP Lynn as of the test date, counting any repeated grades. The (average) causal effect of interest 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 captures random fluctuation in test scores. The dummies d_{ij} indicate three of the four

[8]This conclusion is reinforced by the fact that among lottery losers in our sample, those lost to follow-up tend to have *lower* test scores than those for whom we have data. Hence, perhaps surprisingly, attrition pulls losers' scores up, while among winners, those lost to follow-up have the same average score as those for whom we have data.

KIPP Lynn application cohorts, indexed by j. Note that application cohort is an important control variable because the probability of winning a seat at KIPP varies from year to year.[9]

We use randomly assigned lottery offers as an instrument for s_{igt}. The first stage equation can be written:

$$s_{igt} = \lambda_t + \kappa_g + \sum_j \mu_j d_{ij} + \Gamma' X_i + \pi Z_i + \eta_{igt}, \tag{2}$$

where λ_t and κ_g are year-of-test and grade effects. The excluded instrument is the lottery offer dummy Z_i, with first stage effect π. Specifically, Z_i indicates students offered a seat at KIPP Lynn sometime between the lottery date for the relevant application cohort (usually in March) and the start of the following school year. These offers were determined by randomly assigned lottery sequence numbers.

The reduced form generated by this two-equation system comes from substituting (2) for s_{igt} in (1). The reduced form effect is the coefficient on Z_i in a regression of y_{igt} on Z_i with the same controls and data structure as for equations (1) and (2). Because the model is just-identified, 2SLS estimates of ρ are given by the ratio of reduced form to first stage coefficients.

IV. Results

The lottery first stage is a little over 1.2 years, as can be seen in the first column of Table 4. This estimate indicates that, by the time they were tested, lottery winners had spent an average of 1.2 years more at KIPP than lottery losers. The addition of demographic variables and baseline scores has almost no effect on the first stage estimates.[10] In a world with perfect lottery compliance, no late entry or grade repetition, and no loss to follow-up, the first stage in our sample would be 1.75, but this is reduced by the fact that some winners never enroll in KIPP (or leave before finishing) and some losers end up in KIPP later.[11] On the other hand,

[9] All specifications include a dummy indicating whether an applicant's sibling is in the lottery, as well as the interaction of this dummy with year of application. Siblings who apply together are more likely to get in, since having a winning sibling improves the losing sibling's position on the wait list. Note that applicants with siblings *already* enrolled in KIPP are excluded from the analysis sample, since such applicants are guaranteed admission.

[10] We report separate first stages for math and ELA because samples differ slightly by subject.

[11] The 2005 cohort contributes one score after one year (in 5^{th}grade), one after two years (in 6^{th}grade), one after three years (in 7^{th}grade), and one after four years (in 8^{th}grade) for an average of 2.5 years in KIPP across grades. A similar calculation for the other cohorts, who are seen in fewer grades, produces 2.0 potential years

although KIPP schools are sometimes said to encourage weaker students to leave (see, e.g., the discussion in chapter 46 of Mathews, 2009), we found that lottery winners were about as likely as losers to change schools in grades 6-8, a result discussed further, below.

Lottery winners score about 0.4 standard deviations higher than losers in math. This reduced-form estimate is reported in column (2) in the top half of Table 4. This result is robust to the inclusion of demographic controls and baseline scores. The reduced-form estimates for ELA, reported in the bottom half of the table, are more variable across specifications, ranging from $0.12-0.18\sigma$ as the set of controls varies. This variation probably reflects the modest imbalance between winners and losers in the proportion LEP documented in Table 1. The estimated effect on ELA is marginally significant in models with demographic and baseline score controls.

Because the first-stage coefficients are over one, the 2SLS estimates are smaller than the reduced-form estimates, though they also have a different interpretation. The 2SLS estimates imply that math scores increase by about 0.35σ for each year at KIPP Lynn. The more modest 2SLS estimates for ELA show per-year gains on the order of $0.1 - 0.15\sigma$. The most precise of these is 0.12σ, estimated in models with demographic and baseline score controls (s.e.=0.058). These effects are remarkably similar to the middle school results in Abdulkadiroğlu et al. (2009), which come from a larger sample of charter schools in Boston. Measured against Lynn's Hispanic-White score gaps of about 0.5σ in math and 0.6σ in ELA, both the math and ELA effects are substantial. Perhaps surprisingly, the OLS estimates of math effects reported in column (4) of Table 4 are close to the corresponding 2SLS estimates, though the OLS estimates of ELA effects are a little larger. The similarity of OLS and 2SLS estimates (and the fact that the OLS estimates are insensitive to controls) suggests that in the sample of KIPP Lynn applicants selection bias is minor.[12]

We noted above that the match rate from KIPP Lynn lottery records to SIMS data was almost perfect among students who attended a Lynn public school at baseline. Also significant

in KIPP for the 2006 cohort, 1.5 potential years in KIPP for the 2007 cohort, and one potential year in KIPP for the 2008 cohort. The average of these is 1.75.

[12]We also experimented with an alternative IV model where the instrument is the grade- and cohort-specific potential time in KIPP for winners. This is the first-stage specification used by Hoxby and Murarka (2009). The first stage in this case indicates that each potential year in KIPP causes about 0.7 actual years in KIPP, as shown in column (1) of Appendix Table A4. The corresponding 2SLS results, reported in column (6) of Table A4, are similar to the 2SLS estimates reported in Table 4.

for our analysis, baseline covariates are almost perfectly balanced across lottery status in this subsample. Therefore, as a robustness check, Table 4 reports results for the sample of applicants who attended an LPS school at the time they applied. These estimates, shown in columns (5) and (6) of Table 4, are nearly identical to the estimates obtained from the larger sample. This set of results bolsters our confidence that the full-sample results are not driven by differences in match rates or the small differences in covariates between winners and losers that were detected in Table 1.

To provide a sense of whether the KIPP Lynn treatment effect has been increasing over time, Figure 1 plots reduced-form estimates by cohort and grade. The plots start in fourth grade in order to document any baseline differences.[13] Not surprisingly, treatment effects estimated at this level of disaggregation are fairly noisy and few are individually significant. On the other hand, the math results appear to have increased somewhat for more recent applicant cohorts, while the evolution through grades suggests a cumulative effect. Consistent with the smaller pooled estimates for ELA, the ELA estimates in the plot are mostly smaller than the math estimates and take longer to emerge. The math and ELA results both show an (insignificant) negative effect in 8th grade but this result comes from a single cohort - KIPP Lynn's second, admitted in 2005 - for which the first stage is also relatively small. Overall, these figures suggest that KIPP Lynn's achievement effects are cumulative through grades and increasing across cohorts since the school first opened in 2004.

A. Subgroups, Interactions, and Distribution Effects

KIPP Lynn serves more LEP and special education students than the typical charter school in the Northeast, and therefore offers a unique opportunity to look at the effects of charter attendance on these important subgroups. The first four columns of Table 5 show that math gains are somewhat larger for LEP and special education students than among other students. By contrast, reading gains are markedly larger for students in these groups. In fact, the moderate reading gains found earlier in Table 4 (0.12σ) appear to be driven almost entirely by very large gains among LEP students (roughly 0.4σ) and special education students (roughly

[13]The sample used to construct Figure 1 includes 4th grade applicants only. The reduced-form estimates plotted in the figure come from models that include demographic controls. The estimates are reported in Table A5.

0.3σ).[14]

Table 5 also shows separate results for boys and girls. Nationwide, boys lag behind girls on standardized tests, especially in reading and particularly among Blacks and Hispanics.[15] Boys are also more likely to be classified as needing special education services; among Massachusetts 4th graders, 23 percent of boys receive special education services, compared to 13 percent of girls. These gender differences are similar for students at KIPP Lynn.

Estimates by gender show math effects that are about the same for boys and girls, as can be seen in columns (5) and (6) of Table 5. On the other hand, consistent with the higher proportion of male special needs students and the results reported here showing larger reading effects in the special education subgroup, the effect of KIPP Lynn attendance on boys' ELA scores is considerably larger than the effect for girls (0.15σ vs 0.06σ in models with baseline scores).

Charter skeptics have argued that even if relatively motivated and able students benefit from charter school attendance, weaker students lose out. For instance, Rothstein (2004, p. 82) writes about KIPP: "They select from the top of the ability distribution those lower-class children with innate intelligence, well-motivated parents, or their own personal drives, and give these children educations they can use to succeed in life." If this depiction is accurate, then we might expect KIPP to have benefits that are larger or perhaps even exclusively apparent for those low-income children who are at the upper end of the distribution of academic ability. We explored this type of treatment effect heterogeneity by adding the interaction of baseline (4th grade) scores with years spent at KIPP Lynn to our 2SLS model. The interaction terms are identified in these models by adding an interaction between baseline scores and the lottery offer dummy variable (Z_i in equation 2) to the list of excluded instruments. The interaction terms in these models are normalized so that the main effect of years spent at KIPP reflects the impact of time in KIPP evaluated at the mean of the baseline score distribution.

The results of estimation with interaction terms, reported in columns (7-8) of Table 5, suggest that KIPP Lynn raises achievement *more* for weaker students. Specifically, the reading results indicate that children with baseline reading scores half a standard deviation below the KIPP applicant mean get an additional reading boost of about 0.08σ (=-0.5*-0.157) from each

[14]Surprisingly, given the difference in effects by LEP status, we find that results are similar for Hispanics and non-Hispanics (see Angrist, *et al.* 2010 for details).

[15]See Lee, Grigg and Donahue (2007) and Figure 3 in Mead (2006).

year spent at KIPP, compared to a student whose baseline scores are at the mean. This translates into annual reading gains of 0.14σ per year for the average child at KIPP Lynn (who enters with reading scores 0.4σ below the Massachusetts mean) and annual gains of 0.22σ for a child entering half a standard deviation behind her KIPP classmates (that is, 0.9σ below the Massachusetts mean). Students who enter with the weakest math scores also see a larger math achievement gain from their time at KIPP. The typical KIPP Lynn student experiences math gains of 0.37σ per year, while a student who starts out half a standard deviation behind her KIPP peers realizes annual gains of 0.42σ.

We also looked at the impact of KIPP attendance on the distribution of students across the MCAS proficiency categories that Massachusetts uses to determine whether schools are meeting the AYP standards laid out in federal NCLB legislation. Massachusetts classifies raw MCAS scores in four mutually exclusive categories: Advanced, Proficient, Needs Improvement, and Warning. Under current NCLB provisions, a school is designated as meeting the AYP standard if the school's average score, as well as the average score among various subgroups, falls into the Proficient or Advanced categories. MCAS categories therefore give a simple and policy-relevant picture of the effects of KIPP attendance on the distribution of MCAS scores.

The first row of Table 6 shows the effect of a year spent at KIPP Lynn on the probability a student's score lands him or her in one of the four MCAS categories for math scores. Each year at KIPP is estimated to reduce the probability of falling into the Warning category by ten percentage points, while the probability of performing at an Advanced level rises by ten percentage points. There are no detectable effects in the middle categories. This pattern most likely reflects an across-the-board rightward shift in the distribution of math achievement (since few students are likely to jump all the way from the lowest to the highest category). In contrast, the ELA results show about a 6 point movement away from the Warning group into the Needs Improvement category, with no other change.

While the score gains generated by KIPP Lynn are clearly broader for math than for ELA, it's noteworthy that achievement gains in both subjects reflect a shift out of the lowest group. This conclusion is reinforced by Panel B of Table 6, which reports effects estimated separately for students from each quartile of the baseline (fourth-grade) score distribution in our sample. The estimates in Panel B show positive and significant effects in all baseline score quartiles for math. In contrast, the only significant ELA effect is for students with the lowest baseline scores.

It seems fair to say that these results, which show KIPP Lynn raises the achievement of minority students starting from a very low baseline (The fourth grade scores of KIPP applicants in the lowest baseline quartile are roughly one standard deviation below the Massachusetts average), should be seen as encouraging by policy-makers interested in reducing achievement gaps.

B. School Switching

Can the positive effects reported here be explained by high rates of exit from KIPP? This question is motivated in part by evidence that KIPP schools in the San Francisco area experienced high rates of exit, though it is not clear whether these rates were out of line with those in the host public school districts (Woodworth, et al., 2008). It's worth noting in this context that our lottery-based estimation procedure focuses on score differences between winners and losers (i.e., the reduced form effect of winning the lottery) without regard to whether the winners remain in KIPP. Movements out of KIPP by lottery winners therefore reduce the lottery first stage, while leaving the causal interpretation of the IV estimates uncompromised even if the weakest or least motivated KIPP students switch out. On the other hand, if the score gains generated by KIPP come in part from a small but highly motivated group that remains in the school after winning the lottery, while weaker or less-motivated students wash out, selective switching may have a beneficial peer effect that augments effects for those who remain. High exit rates from KIPP might also limit the external validity of our estimates for a broader and perhaps less motivated population.

Our school switching analysis uses the same empirical framework as that used to investigate attrition in Table 3, but the dependent variable in this case indicates whether a KIPP applicant changed schools between grades five and eight. These results, reported in the first column of Table 7, show that KIPP Lynn lottery winners were much *less likely* to change schools than those who lost the lottery. As can be seen in column (2), this difference (and the overall high mean switch rate) can be explained by the fact that KIPP Lynn students stay at KIPP in the transition from 5th to 6th grade, when LPS students move from elementary to middle school. Excluding the transition from 5th to 6th grade, the results show no difference in switching between lottery winners and losers, as can be seen in the last column of Table 7. This implies that KIPP students are no more likely to change middle schools than their LPS peers (in this case, the mean switch rate for both groups is only about 0.08).

V. Concluding Comments

KIPP is a large and growing charter management organization that exemplifies the No Excuses approach to public education. The lottery-based estimates reported here suggest that New England's only KIPP school - KIPP Lynn - has generated substantial score gains for KIPP students, with effects on the order of 0.35σ in math and 0.12σ for ELA. Score gains are largest for special education students and students with limited English proficiency. Reading gains are realized almost exclusively by the students in these groups. Male and female students appear to have experienced similar achievement gains in math, but reading gains are much stronger for boys. A specification that interacts KIPP attendance with baseline scores indicates that effects are larger for those who start out lagging their peers than for more advanced students. Finally, an analysis of effects on MCAS performance categories shows that KIPP lifts students out of the lowest performance category for both math and ELA.

A recent charter study concludes that newly opened charter schools do worse than traditional public schools, with effects that tend to improve as schools age (Zimmer, Gill, Booker, Lavertu, Sass and Witte 2009). It's therefore worth noting that the results reported here are from the first few cohorts to attend KIPP Lynn.

A natural question in this context is how general the lessons learned from a single school can be. Each school has unique features and charter schools are especially diverse, ranging from the highly structured and in many ways quite traditional school environment typical of KIPP to Ohio's "virtual charter schools" that use technology to deliver education to students in their homes. We see the experiences of KIPP Lynn not as definitive for charters as a whole, but rather as revealing important possibilities, just as the smaller-scale but highly influential 1962 Perry Preschool experiment provided a kind of pilot for national Head Start. While our results are for a single school, the KIPP organization runs many similar schools across the country. Key elements of the KIPP program also feature in other No Excuses charter schools, such as those in Boston and New York. Our findings suggest the major elements of this replicable model produce substantial achievement gains overall, and especially large gains for relatively weak students and those needing special services.

Table 1: Descriptive Statistics and Covariate Balance

	Means			Balance regressions	
	Lynn Public 5th graders	KIPP Lynn 5th graders	KIPP Lynn lottery applicants	No controls	Demographic controls
	(1)	(2)	(3)	(4)	(5)
Hispanic	0.418	0.565	0.538	-0.052	-
				(0.053)	
Black	0.173	0.235	0.254	0.027	-
				(0.044)	
White	0.296	0.168	0.182	-0.010	-
				(0.040)	
Asian	0.108	0.021	0.022	0.026*	-
				(0.015)	
Female	0.480	0.474	0.484	-0.010	-
				(0.054)	
Free/reduced price lunch	0.770	0.842	0.825	-0.030	-
				(0.041)	
Special Education	0.185	0.189	0.197	-0.013	-
				(0.042)	
Limited English Proficiency	0.221	0.172	0.206	-0.075	
				(0.047)	
Baseline Math Score	-0.307	-0.336	-0.390	0.097	0.033
				(0.114)	(0.106)
Baseline Verbal Score	-0.356	-0.399	-0.438	0.054	-0.021
				(0.118)	(0.104)
F-value from joint test				0.756	0.082
p-value from F-test				0.671	0.921
N for demographics	3964	285	457	457	457
N for baseline Math	3808	284	446	446	446
N for baseline ELA	3805	284	447	447	447

Notes: Columns (1), (2), and (3) report means of the variable indicated in each row. Column (1) reports 4th grade means for students that attended 5th grade in Lynn public schools in Fall 2005-2008. Column (2) reports 4th grade means for all students who attended KIPP Academy Lynn in these years, and column (3) reports 4th grade means for lottery applicants to KIPP Academy Lynn over the same period. The sample for columns (3)-(5) is restricted to randomized applicants with baseline demographics and excludes students who had completed 6th or 7th grade prior to applying. Column (4) reports coefficients from regressions of the variable indicated in each row on an indicator variable equal to one if the student won the lottery. These regressions include dummies for year of application and application grade and exclude students with sibling priority and those without baseline demographics. Column (5) adds all of the demographic controls to the regressions for baseline scores. F-tests are for the null hypothesis that the coefficients on winning the lottery in all regressions are equal to zero. These tests statistics are calculated for the subsample that has non-missing values for all variables tested.

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

Table 2: KIPP Academy Lynn Lotteries

Lottery Cohort	Calendar years observed	Grades observed	Number of applicants	Number of applicants in lottery sample	Percent offered	Percent attended	Average years at KAL (winners)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2005-2006	2006-2009	5-8	138	107	0.925	0.673	2.56
2006-2007	2007-2009	5-7	117	86	0.674	0.535	2.29
2007-2008	2008-2009	5-6	167	127	0.654	0.567	1.71
2008-2009	2009	5	207	137	0.540	0.401	0.70
All cohorts	2006-2009	5-8	629	457	0.687	0.536	1.85

Notes: This table reports characteristics of the four lotteries conducted at KIPP Academy Lynn from 2005 to 2008. Column (2) reports the calendar years (Spring) in which test scores are observed for applicants in each lottery cohort, and column (3) reports the corresponding outcome grades. Column (4) gives the total number of applicants in each year, and column (5) gives the number of applicants in the lottery sample, which excludes sibling applicants, late applicants, repeat applicants, applicants without baseline demographics, applicants who could not be matched to the MCAS data, and applicants who had completed 6th or 7th grade prior to the lottery. Columns (6)-(8) gi summary statistics for the lottery sample.

Table 3: Attrition

		Differential follow-up (winner - loser)		
	Proportion of non-offered with MCAS scores	Basic controls	Demographics	Demographics and baseline scores
Subject	(1)	(2)	(3)	(4)
Math	0.851	0.055*	0.041	0.044
		(0.032)	(0.031)	(0.030)
		971	971	957
ELA	0.855	0.046	0.031	0.041
		(0.033)	(0.032)	(0.031)
		971	971	958

Notes: Column (1) reports the fraction of test scores found for non-offered students. Columns (2)-(4) report coefficients from regressions of an indicator variable equal to one if the outcome test score is non-missing on an indicator variable equal to one if the student won the lottery. Grades are pooled, and all regressions include grade dummies. The regression in column (2) includes dummies for outcome grade, year of baseline, application grade, and year of application interacted with a contemporaneous sibling applicant dummy. Column (3) adds demographic variables, and column (4) adds baseline test scores. Samples are restricted to cohorts for which we should observe follow-up scores. Robust standard errors (clustered at the student level) are reported in parentheses.

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

Table 4: Lottery Results

		all applicants				Lynn public schools at baseline	
		First Stage	Reduced Form	2SLS	OLS	2SLS	OLS
Subject	Controls	(1)	(2)	(3)	(4)	(5)	(6)
Math	Basic	1.222***	0.431***	0.353***	0.304***	0.342***	0.308***
		(0.063)	(0.116)	(0.095)	(0.048)	(0.109)	(0.054)
		865	865	865	865	704	704
	Demographics	1.232***	0.392***	0.318***	0.316***	0.309***	0.336***
		(0.065)	(0.105)	(0.084)	(0.041)	(0.098)	(0.046)
		865	865	865	865	704	704
	Demographics &	1.228***	0.425***	0.346***	0.317***	0.341***	0.346***
	Baseline Scores	(0.066)	(0.066)	(0.052)	(0.032)	(0.064)	(0.038)
		856	856	856	856	696	696
ELA	Basic	1.223***	0.183	0.150	0.170***	0.217*	0.168***
		(0.063)	(0.117)	(0.094)	(0.049)	(0.115)	(0.057)
		866	866	866	866	705	705
	Demographics	1.235***	0.118	0.095	0.172***	0.150	0.180***
		(0.066)	(0.097)	(0.077)	(0.041)	(0.092)	(0.047)
		866	866	866	866	705	705
	Demographics &	1.234***	0.149**	0.120**	0.172***	0.132*	0.182***
	Baseline Scores	(0.066)	(0.073)	(0.058)	(0.031)	(0.068)	(0.036)
		856	856	856	856	698	698

Notes: This table reports the coefficients from regressions of test scores on years spent at KIPP Academy Lynn. The sample uses students who applied to KIPP Lynn between 2005 and 2008. It is restricted to students with baseline demographic characteristics and excludes applicants with sibling priority. Grades are pooled, and all regressions include grade dummies. All regressions also include year of test dummies, year of application dummies interacted with a contemporaneous sibling applicant dummy, and grade of application dummies. Some regressions 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. Columns (1)-(3) report the first stage, reduced form, and 2SLS coefficients from instrumenting years in KIPP Lynn using the lottery win/loss dummy. Column (4) reports the coefficients from OLS regressions of test scores on years in KIPP Lynn and controls. Columns (5) and (6) report 2SLS and OLS results using only students that indicated Lynn Public School attendance prior to the lottery on their KIPP Lynn applications. Robust standard errors (clustered at the student level) are reported in parentheses.

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

Table 5: Subgroups and Interactions

		Effects by Subgroup						Baseline Score Interaction	
		LEP	Non-LEP	SPED	Non-SPED	Male	Female	Main effect	Interaction term
Subject	Controls	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	Demographics	0.628***	0.254***	0.527**	0.271***	0.323***	0.290**	-	-
		(0.197)	(0.093)	(0.215)	(0.087)	(0.111)	(0.126)		
		132	733	175	690	444	421		
	Demographics and Baseline Scores	0.451***	0.312***	0.441***	0.325***	0.322***	0.385***	0.367***	-0.106***
		(0.155)	(0.056)	(0.146)	(0.053)	(0.071)	(0.079)	(0.054)	(0.041)
		131	725	174	682	439	417	856	
ELA	Demographics	0.416**	0.019	0.220	0.038	0.140	0.010	-	-
		(0.183)	(0.084)	(0.216)	(0.079)	(0.104)	(0.116)		
		131	735	176	690	442	424		
	Demographics and Baseline Scores	0.384***	0.051	0.298*	0.049	0.152*	0.061	0.139**	-0.157***
		(0.140)	(0.062)	(0.162)	(0.058)	(0.079)	(0.086)	(0.057)	(0.045)
		130	726	174	682	436	420	856	

Notes: Columns (1)-(4) report 2SLS estimates in subsets of the lottery sample. The sample for each regression is restricted to individuals who were classified as limited english proficient (LEP), special education (SPED), or male in columns (1), (3), and (5), compared to those who were not in columns (2), (4), and (6) respectively. The LEP estimation sample includes 79 students, while the non-LEP sample includes 319. The SPED estimation sample includes 78 students, while the non-SPED sample includes 320. The male estimation sample includes 205 students, while the female sample includes 196. Columns (7) and (8) report results from models interacting baseline test score with years at KIPP Academy Lynn. Main effects are at the mean. The interaction models are estimated by including the offer dummy interacted with baseline score as a second instrument. Robust standard errors (clustered at the student level) are reported in parentheses.

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

Table 6: Distribution Effects

Subject	Lowest Group (1)	Second Lowest (2)	Second Highest (3)	Highest Group (4)
Panel A. Effects on MCAS Performance Categories				
Math	-0.100***	-0.019	0.016	0.103***
	(0.028)	(0.038)	(0.039)	(0.026)
Fraction in category	0.183	0.319	0.335	0.162
N		856		
ELA	-0.055***	0.068*	-0.005	-0.003
	(0.020)	(0.037)	(0.037)	(0.017)
Fraction in category	0.084	0.350	0.500	0.065
N		856		
Panel B. Effects by Baseline Score Quartile				
Math effect	0.508***	0.464***	0.456***	0.198***
	(0.158)	(0.108)	(0.101)	(0.056)
Mean score by quartile	-0.961	-0.139	0.225	0.753
N	215	225	202	214
ELA effect	0.477***	0.046	0.012	-0.076
	(0.143)	(0.121)	(0.096)	(0.075)
Mean score by quartile	-1.129	-0.381	-0.005	0.474
N	218	211	214	213

Notes: Panel A reports coefficients from 2SLS regressions of indicator variables for each of the 4 MCAS performance levels on years in KIPP Lynn instrumented by the lottery offer dummy. Panel B reports 2SLS estimates of test score effects by baseline score quartile (defined by the distribution of 4th grade scores in our sample). Regressions are run separately for each quartile. Grades are stacked. Controls include demographics and baseline scores. Robust standard errors (clustered at the student level) are reported in parentheses.

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

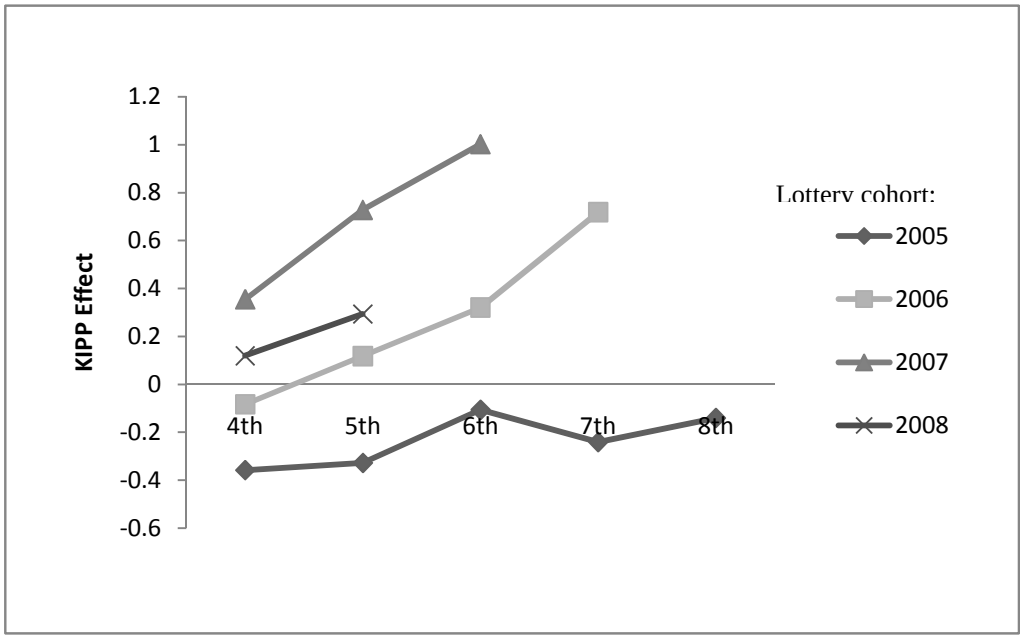
Table 7: School Switching Regressions

	Differential (winner - loser)		
	Any switch	6th grade school is different from 5th	Any switch excluding 5th-6th transition
Controls	(1)	(2)	(3)
Demographics	-0.291***	-0.503***	-0.006
	(0.044)	(0.060)	(0.033)
Mean loser switch rate	0.504	0.855	0.081
	419	294	419
Demographics and Baseline Scores	-0.294***	-0.509***	-0.004
	(0.045)	(0.059)	(0.034)
Mean loser switch rate	0.513	0.853	0.084
	412	291	412

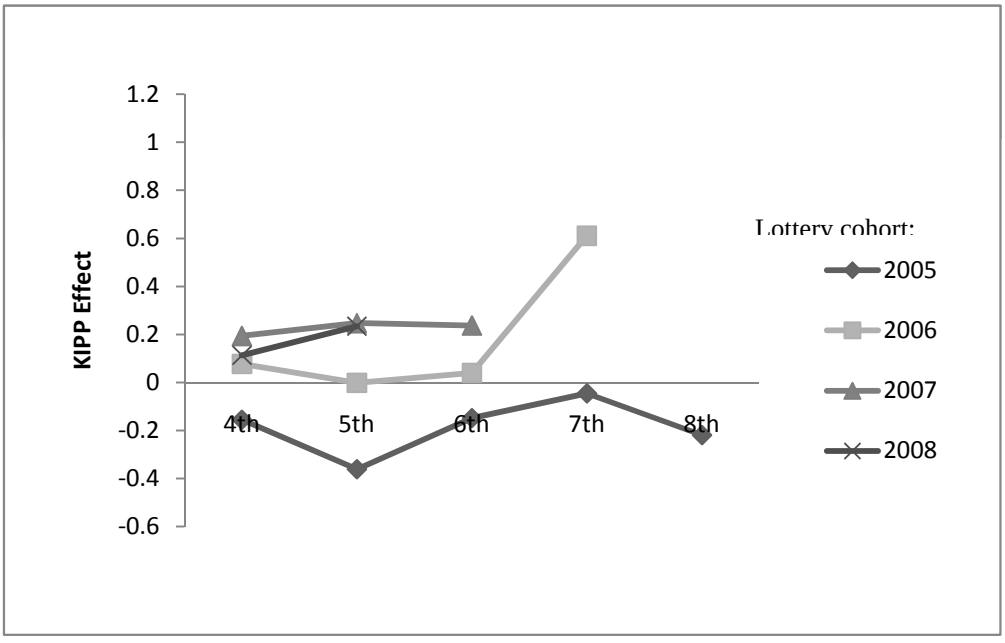
Notes: This table reports coefficients from regressions of an indicator variable equal to one if a student switched schools on an indicator variable equal to one if the student won the KIPP Academy Lynn lottery. The dependent variable in column (1) is 1 if a student ever moves from one observed school to another from 5th to 8th grade, either within a school year or between school years. The dependent variable in column (2) is 1 if a student switches schools between 5th and 6th grade; only observations where both schools are observed are used. The dependent variable in column (3) is 1 if a student switches schools at any time besides the transition from 5th to 6th grade. The sample is restricted to cohorts for which we should observe follow-up test scores and excludes applicants with sibling priority. Robust standard errors are reported in parentheses.

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

Figure 1



A. Math Reduced Form



B. ELA Reduced Form

Notes: This figure plots the coefficients from a regression of test scores on the lottery offer dummy interacted with dummies for grade of test*application year. Basic and demographic controls are included. Sample restricted to 4th grade applicants.

Data Appendix

Data for this project come from KIPP Academy Lynn lottery records, student demographic and school attendance information in the Massachusetts Student Information Management System (SIMS), and test scores from the Massachusetts Comprehensive Assessment System (MCAS) state database. This appendix describes these data sets and details the procedures used to clean and match them.

A. Data Sets

KIPP Academy Lynn Lottery Data

Data description and sample restrictions

Our sample of applicants is drawn from records of the four lotteries that took place at KIPP Academy Lynn from 2005 through 2008. These records include applicants' names, date of birth, previous school and grade, and contact information for a parent, guardian, or the name of a sponsoring organization such as the Lynn Boys Club. The first five rows of Table A2 summarize the raw lottery data and sample restrictions used here. A few students who repeated grades were listed in the lottery data to remind school staff to reserve an appropriate number of slots. These records are not included in the analysis sample. We also excluded duplicate records. If a student applied to KIPP Academy Lynn more than once, only the first application is included. Late applicants (after lotteries) were excluded as were siblings and students who went directly onto the waiting list (these are mostly 6th grade applicants in early cohorts). Imposing these restrictions reduces the number of lottery records from 629 to 542.

Coding the offer variable

Lottery records were used to reconstruct an indicator for whether applicants won the chance to attend KIPP Lynn through the lottery process. We coded this from information on whether each student attended KIPP Lynn in the year after the lottery, attempts to contact lottery winners, and offers that were declined. Attempts to contact winners and declined offers were not always recorded; we filled this in by reviewing each applicant record with school staff. Of the 542 randomized applicants in our lottery sample, 350 were coded as receiving offers.

Student Information Management System Data

Data description

This project uses SIMS data from the 2001-2002 school year through the 2008-2009 school year. Each year of data includes an October file and an end-of-year file. The SIMS records demographic and attendance information for all Massachusetts public school students. SIMS records refer to a student in a school in a year, though there are some student-school-year duplicates for students that switch grades or programs within a school and year.

Coding of demographics and attendance

The SIMS variables of interest include grade, year, name, town of residence, date of birth, sex, race, special education (SPED) and limited English proficiency (LEP) status, free/reduced price lunch, and school attended. We constructed a wide-format data set that captures each student's demographic information for each grade in which he or she is present in the SIMS data. This file uses the demographic information from the longest-attended school in the first calendar year encountered for each grade. Attendance ties were broken at random (this affects only 0.014 percent of records). If a student is classified as SPED, LEP, or qualified for free/reduced price lunch in any record within a school-year-grade, then he or she is coded that way for the entire school-year-grade record.

KIPP Lynn attendance is measured in calendar years. A student was coded as attending KIPP Lynn when there is any SIMS record for KIPP attendance in that year. Our analysis uses grade of application as determined by the SIMS (as some parents record this incorrectly on lottery applications).

Massachusetts Comprehensive Assessment System Data

Data description and sample restrictions

This project uses MCAS data from the 2001-2002 school year through the 2008-2009 school year. Each record in the MCAS data corresponds to a student's test results for a given grade and year. We use Math and English Language Arts (ELA) tests from grades 4-8. Our outcome grades are 5-8, so only tests taken in 2006-2007 or later are used for these grades; prior years give baseline (4th grade) scores. We standardized scores to have mean zero and standard deviation one within a subject-grade-year in Massachusetts. Repetitions of the same test subject and

grade were dropped. In one case with multiple records within a year and grade, scores were chosen at random.

B. Matching Data Sets

Match from the MCAS to the SIMS

The cleaned MCAS and SIMS files were merged by grade, year and a state student identifier known as the SASID. In grades 4-8, 99.3 percent of MCAS scores were matched to a student in the SIMS. Scores that could not be matched to the SIMS were dropped.

Match from the KIPP Academy Lynn lotteries to the SIMS/MCAS

Match procedure

Students in the lottery sample were matched to the SIMS data by name, pre-lottery grade, and year. In some cases, this did not produce a unique match, most often in cases where the lottery data were incomplete. We accepted some matches based on fewer criteria where the information on grade, year, and town of residence seemed to make sense.

Match success rates

Table A1 reports match rates from lottery records to the SIMS/MCAS file. The overall match rate is 91.3 percent (495 students out of 542). The match rate for offered students is 95.4 percent, while it is 83.9 percent for students who did not receive an offer. The differential is much lower for lottery applicants coming from an LPS school; the match rates for the offered and non-offered students in this subgroup are 96.4 percent and 93.4 percent, respectively. The differentials quoted in the text come from regressions of a match dummy on application year and LPS status (or just application year in the sample coming from LPS).

Construction of the Outcome Data Set

The lottery/SIMS/MCAS matched sample includes 495 lottery applicants with demographic and test score information. Of these, we use only students with baseline (4th grade) demographics in the SIMS. We also exclude 10 applicants who had completed 6th or 7th grade prior

to the lottery, leaving a sample of 457 students. This is the sample of students used for the calculations reported in Table 2. Rows 6-8 of Table A2 summarize the impact of these restrictions on sample size.

Stacking grades

Outcome regressions stack grades and include multiple test scores for individual students. The follow-up window closes in Spring 2009, generating differences in the number of outcomes observed across lottery cohorts. For example, a 4th grade applicant for the 2005-2006 school year contributes 5th grade through 8th grade scores, whereas we see 5th grade only for 2008 applicants. Years in KIPP Lynn is defined as the number of school years spent at KIPP up to and including the outcome year.

Outcomes excluded from the sample

KIPP Lynn typically asked 5th grade applicants to repeat. These applicants might be expected to do better on 5th grade MCAS tests just by virtue of repeating. We therefore assume that all 5th grade applicants repeat and look only at their 6th grade and higher scores. We also drop a few 3rd grade applicants. These restrictions reduce the sample to 419, eliminating 38 2008-9 applicants from 5th grade (and a handful from 3rd).

Final set of outcomes and students

Table A3 summarizes the stacked analysis file. Of the 971 post-lottery outcomes we could hope to observe for each subject, we found 865 for Math and 866 for ELA; 401 of our 419 remaining students have at least one test score. These outcomes and students were used to produce the estimates in Table 4. For specifications that control for baseline test scores, the sample sizes are further reduced to 856 outcomes for both Math and ELA; 4 students out of 401 lack baseline Math and ELA scores.

Table A.1: KIPP Academy Lynn Lottery Records

	Lottery cohort				
	2005-2006	2006-2007	2007-2008	2008-2009	All lotteries
	(1)	(2)	(3)	(4)	(5)
Total number of records	138	117	167	207	629
Excluding KIPP students and duplicates within year	138	117	162	205	622
Excluding repeat applicants	138	115	158	196	607
Excluding late/non-randomized applicants	127	110	155	194	586
Excluding siblings	122	102	143	175	542
Excluding students not matched to the SIMS	115	95	131	154	495
Excluding students without baseline demographics	110	86	127	144	467
Excluding 6th and 7th grade applicants	107	86	127	137	457
Excluding applicants who should not have a test score	107	86	127	99	419

Notes: This table summarizes the raw KIPP Academy Lynn lottery data. The top row gives the total number of records, and each successive row adds sample restrictions. The second row eliminates KIPP Lynn students who repeat grades and are listed in the lottery data as placeholders, as well as duplicate student records within a lottery year. The third row keeps only the first lottery year in which a given student applies, and the fourth row excludes late (post-lottery) applicants as well as other non-randomized applicants. The fifth row eliminates students with sibling priority. The sixth row eliminates students who cannot be matched to the SIMS database. The seventh row excludes students without baseline (4th grade) demographics. The seventh row excludes students who had completed 6th or 7th grade prior to the lottery. The eighth row excludes students who should not have a non-repeat test score based on application grade and cohort (which eliminates 3rd and 5th grade applicants in 2008).

Table A.2: Match from KIPP Academy Lynn lottery data to SIMS

	All applicants				Applicants from Lynn Public Schools			
		Fraction with SIMS match				Fraction with SIMS match		
Lottery cohort	Number of students (1)	Total (2)	Offered (3)	Not offered (4)	Number of students (5)	Total (6)	Offered (7)	Not offered (8)
2005-2006	122	0.943	0.955	0.833	103	0.971	0.968	1.000
2006-2007	102	0.931	0.971	0.848	76	0.934	0.963	0.864
2007-2008	143	0.916	0.955	0.852	100	0.960	0.971	0.935
2008-2009	175	0.880	0.939	0.828	117	0.949	0.947	0.950
All cohorts	542	0.913	0.954	0.839	396	0.955	0.964	0.934

Notes: This table summarizes the match from the KIPP Academy Lynn lottery data to the SIMS. The sample excludes repeat applicants, late applicants, and siblings. Columns (1)-(4) report statistics for all other applicants, and columns (5)-(8) report statistics for students whose previous schools in the KAL lottery data are part of the Lynn Public School system.

Table A.3: Outcome data for KIPP Academy Lynn Applicants

Lottery cohort	Number of students (1)	Number with an observed test score (2)	Number of test scores expected (3)	Math test scores observed (4)	ELA test scores observed (5)
2005-2006	107	105	405	357	361
2006-2007	86	84	238	212	211
2007-2008	127	122	229	206	204
2008-2009	99	90	99	90	90
All cohorts	419	401	971	865	866

Notes: This table summarizes observed test score outcomes for KIPP Academy Lynn applicants. The sample is restricted to randomized applicants who are matched to baseline (4th grade) SIMS demographics and who should have at least one test score. 6th and 7th grade applicants are excluded. Column (2) reports the number of students for whom at least one outcome is observed. Column (3) gives the number of test scores that should be observed (for both Math and ELA) given each applicant's lottery cohort and application grade. Columns (4) and (5) report the numbers of Math and ELA outcomes that are observed in the data.

Table A.4: Alternative Instruments

		Offer instrument			Alternative instrument		
		First Stage	Reduced Form	2SLS	First Stage	Reduced Form	2SLS
Subject	Controls	(1)	(2)	(3)	(4)	(5)	(6)
Math	Basic	1.222***	0.431***	0.353***	0.684***	0.206***	0.301***
		(0.063)	(0.116)	(0.095)	(0.039)	(0.063)	(0.089)
		865	865	865	865	865	865
	Demographics	1.232***	0.392***	0.318***	0.687***	0.185***	0.269***
		(0.065)	(0.105)	(0.084)	(0.040)	(0.057)	(0.077)
		865	865	865	865	865	865
	Demographics & Baseline Scores	1.228***	0.425***	0.346***	0.688***	0.232***	0.337***
		(0.066)	(0.066)	(0.052)	(0.040)	(0.038)	(0.051)
		856	856	856	856	856	856
ELA	Basic	1.223***	0.183	0.150	0.685***	0.081	0.118
		(0.063)	(0.117)	(0.094)	(0.039)	(0.060)	(0.086)
		866	866	866	866	866	866
	Demographics	1.235***	0.118	0.095	0.689***	0.050	0.072
		(0.066)	(0.097)	(0.077)	(0.039)	(0.048)	(0.067)
		866	866	866	866	866	866
	Demographics & Baseline Scores	1.234***	0.149**	0.120**	0.690***	0.091**	0.131**
		(0.066)	(0.073)	(0.058)	(0.039)	(0.038)	(0.054)
		856	856	856	856	856	856

Notes: This table reports instrumental variables results similar to those in Table 4. It is restricted to students with baseline demographic characteristics and excludes applicants with sibling priority. Grades are pooled, and all regressions include grade dummies. Columns (1)-(3) report the first stage, reduced form, and 2SLS coefficients from instrumenting years in KIPP Lynn with the lottery offer dummy as in Table 4. Columns (4)-(6) report results using potential years in KIPP Lynn interacted with the offer dummy as the instrument. Potential years in KIPP Lynn is calculated as the number of years a student would accumulate by attending KIPP Lynn in each post-lottery year until the outcome grade without repeating (except for 5th grade applicants, who are assumed to repeat one grade). Robust standard errors (clustered at the student level) are reported in parentheses.

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

Table A.5: Reduced Forms by Grade and Cohort

Subject	Test Grade	2005 Cohort	2006 Cohort	2007 Cohort	2008 Cohort	All Cohorts
Math	4th	-0.358	-0.083	0.355	0.119	0.137
		(0.241)	(0.256)	(0.227)	(0.187)	(0.120)
		83	65	101	97	346
	5th	-0.328	0.118	0.728***	0.293	0.336***
		(0.242)	(0.228)	(0.229)	(0.182)	(0.119)
		83	65	95	90	333
	6th	-0.106	0.320	1.002***	-	0.607***
		(0.219)	(0.259)	(0.218)		(0.153)
		79	61	86		226
	7th	-0.241*	0.718***	-	-	0.653***
		(0.146)	(0.231)			(0.200)
		77	54			131
	8th	-0.140	-	-	-	-0.124
		(0.160)				(0.151)
		62				62
ELA	4th	-0.155	0.077	0.194	0.113	0.125
		(0.224)	(0.202)	(0.212)	(0.198)	(0.112)
		83	65	101	98	347
	5th	-0.361	-0.002	0.247	0.234	0.136
		(0.173)	(0.213)	(0.242)	(0.193)	(0.119)
		82	64	93	90	329
	6th	-0.148	0.040	0.237	-	0.128
		(0.300)	(0.237)	(0.223)		(0.142)
		80	61	86		227
	7th	-0.046	0.611***	-	-	0.517***
		(0.156)	(0.206)			(0.177)
		78	54			132
	8th	-0.219	-	-	-	-0.209
		(0.223)				(0.227)
		63				63

Notes: This table reports coefficients from regressing test scores on a full set of application cohort*test grade dummies interacted with the lottery offer dummy. The "all cohorts" coefficients are produced from regressions interacting grade dummies with the lottery offer dummy. Basic and demographic controls are included in all regressions. Sample is restricted to 4th grade applicants. Standard errors are clustered at the student level.
* significant at 10%; ** significant at 5%; *** significant at 1%

References

[1] Abdulkadiroğlu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2009. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." NBER Working Paper 15549.

[2] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005a. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling." *Journal of Human Resources* 40:4, 791-821.

[3] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005b. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113, 1514.

[4] Angrist, Joshua D. Susan M. Dynarski, Thomas J. Kane, Parag A. Pathak, and Christopher Walters. 2010. "Inputs and Impacts in Charter Schools: KIPP Lynn." Forthcoming, in *American Economic Review Papers and Proceedings* 100(2).

[5] Carnoy, Martin, Rebecca Jacobsen, Lawrence Mishel, and Richard Rothstein. 2005. *The Charter School Dust-Up: Examining Evidence on Student Achievement.* Washington, DC: Economic Policy Institute Press.

[6] Carter, Samuel Casey. 2000. "No Excuses: Lessons from 21 High-Performing, High-Poverty Schools." Washington, DC: Heritage Foundation.

[7] Childress, Stacey and Maura Marino. 2008. "KIPP 2007: Implementing a Smart Growth Strategy." Harvard Business School Case 9-308-073.

[8] Coleman, James S. 1966. *Equality of Educational Opportunity.* United States Department of Health, Education and Welfare.

[9] CREDO. 2009. *Multiple Choice: Charter Performance in 16 states.* Stanford University: Center for Research on Education Outcomes.

[10] Dobbie, Will and Roland G. Fryer. 2009. "Are High Quality Schools Enough to Close the Achievement Gap? Evidence from a Social Experiment in Harlem." NBER Working Paper 15473.

[11] Fryer, Roland G. 2010. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." NBER Working Paper 15898.

[12] Hoxby, Caroline M. and Sonali Murarka. 2009. "Charter Schools in New York City: Who Enrolls and How They Affect Student Achievement." NBER Working Paper 14852.

[13] KIPP Academy Lynn Charter School. 2003. "Final Application Submitted to the Massachusetts Department of Education."

[14] Lee, J., Grigg, W., and Donahue, P. 2007. "The National Report Card: Reading 2007. "˙U.S. Department of Education, Washington, D.C.

[15] Leschly, Stig. 2003. "KIPP National, 2002 (B): Managing a School Network." Harvard Business School Case 9-804-049.

[16] Leschly, Stig. 2008. "KIPP National (A) (Abridged)." Harvard Business School Case 9-805-068.

[17] Massachusetts Department of Elementary and Secondary Education. 2008. "KIPP Academy Lynn Charter School: Summary of Review. "

[18] Mathews, Jay. 2009. *Work Hard. Be Nice.* Chapel Hill, NC: Algonquin Books.

[19] Mead, Sara. 2006. "The Truth about Boys and Girls." Education Sector, Washington, D.C.

[20] Rothstein, Richard. 2004. *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap.* New York: Teachers College Press.

[21] Thernstrom, Abigail and Stephen Thernstrom. 2003. *No Excuses: Closing the Racial Gap in Learning.* New York: Simon & Schuster.

[22] United Federation of Teachers. 2010. "Separate and Unequal: The Failure of New York City Charters Schools to Serve the City's Neediest Students." Accessed January 28, 2010 at http://www.uft.org.

[23] Whitman, David. 2008. *Sweating the Small Stuff: Inner-City Schools and the New Paternalism.* Washington, DC: The Fordham Institute.

[24] Woodworth, Katrina R., Jane L. David, Roneeta Guha, Haiwen Wang, and Alejandra Lopez-Torkos. 2008. *San Francisco Bay Area KIPP Schools: A Study of Early Implementation and Achievement, Final Report.* Menlo Park, CA: SRI International.

[25] Zimmer, Ron, Brian Gill, Kevin Booker, Stephane Lavertu, Tim R. Sass and John Witte. 2009. "Charter Schools in Eight States: Effects on Achievement, Attainment, Integration and Competition." Santa Monica, CA: RAND Corporation.