NBER WORKING PAPER SERIES

CHARTERS WITHOUT LOTTERIES:
TESTING TAKEOVERS IN NEW ORLEANS AND BOSTON

Atila Abdulkadiroğlu
Joshua D. Angrist
Peter D. Hull
Parag A. Pathak

Working Paper 20792
http://www.nber.org/papers/w20792

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2014

Our thanks to Raymond Cwiertniewicz, Alvin David, Gabriela Fighetti, and Jill Zimmerman from the Recovery School District; to Kamal Chavda and the Boston Public Schools; and to Scott Given, Ryan Knight and the staff at Unlocking Potential for graciously sharing data and answering our many questions. We're grateful to Alonso Bucarey, Stephanie Cheng, Olivia Kim, Elizabeth Setren, Mayara Silva, Daisy Sun, and Danielle Wedde for exceptional research assistance and to MIT SEII program manager Annice Correia for invaluable administrative support. Data from the Recovery School District were made available to us through the Institute for Innovation in Public School Choice. We gratefully acknowledge financial support from the Institute for Education Sciences (under Award R305A120269), from the National Science Foundation (under award SES-1426541), and from the the Laura and John Arnold Foundation. Thanks also go to seminar participants at the Federal Reserve Bank of New York, the Stanford Graduate School of Business, and the Fall 2014 NBER Economics of Education Meeting for helpful comments. Joshua Angrist's daughter teaches at UP Academy Charter School of Boston. The views expressed here are those of the authors alone and do not necessarily reflect the views of the National Bureau of Economic Research or those of personnel at the schools, school districts, or government agencies connected with this work.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w20792.ack

Charters Without Lotteries: Testing Takeovers in New Orleans and Boston
Atila Abdulkadiroğlu, Joshua D. Angrist, Peter D. Hull, and Parag A. Pathak
NBER Working Paper No. 20792
December 2014
JEL No. C21,C26,C31,I21,I22,I28,J24

ABSTRACT

Lottery estimates suggest oversubscribed urban charter schools boost student achievement markedly. But these estimates needn't capture treatment effects for students who haven't applied to charter schools or for students attending charters for which demand is weak. This paper reports estimates of the effect of charter school attendance on middle-schoolers in charter takeovers in New Orleans and Boston. Takeovers are traditional public schools that close and then re-open as charter schools. Students enrolled in the schools designated for closure are eligible for "grandfathering" into the new schools; that is, they are guaranteed seats. We use this fact to construct instrumental variables estimates of the effects of passive charter attendance: the grandfathering instrument compares students at schools designated for takeover with students who appear similar at baseline and who were attending similar schools not yet closed, while adjusting for possible violations of the exclusion restriction in such comparisons. Estimates for a large sample of takeover schools in the New Orleans Recovery School District show substantial gains from takeover enrollment. In Boston, where we can compare grandfathering and lottery estimates for a middle school, grandfathered students see achievement gains at least as large as the gains for students assigned seats in lotteries. Larger reading gains for grandfathering compliers are explained by a worse non-charter fallback.

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

Joshua D. Angrist
Department of Economics, E17-226
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
angrist@mit.edu

Peter D. Hull
Department of Economics,
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
hull@mit.edu

Parag A. Pathak
Department of Economics, E17-240
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
ppathak@mit.edu

No child's chance in life should be determined by the luck of a lottery

– President Obama (quoted in *The Boston Globe*, March 13, 2011)

1 Introduction

The question of how best to improve large urban school districts remains a touchstone in the debate over American school reform. The role of charter schools – publicly funded schools operated outside the public sector – is especially controversial. Nationwide, charter school enrollment grew from under one percent in 2000 to over four percent in 2011. Charter expansion has since continued apace: the National Alliance for Public Charter schools reports a net increase of 381 charter schools operating between Fall 2011 and Fall 2012, with a charter enrollment gain of 13.5 percent. Growth has been especially strong in large urban districts such as Boston, Los Angeles, Oakland, Newark, New York, and Philadelphia, where many students are poor and most are nonwhite. The schools in these districts are often described as low-performing, with low standardized test scores, high truancy rates, and high dropout rates.[1]

In the 2014-15 school year, the New Orleans Recovery School District (RSD) became America's first all-charter public school district. RSD emerged from a 2003 effort to improve underperforming public schools in New Orleans, home to some of the worst schools in the country. State legislation known as Act 9 allowed the Louisiana Department of Education (LDE) to take control of, manage, and outsource the operation of schools deemed low-performing based on measures related to achievement, attendance, and graduation rates. As a result of Act 9, New Orleans public schools that came under state control became part of RSD, while other schools remained under the authority of the Orleans Parish School Board (OPSB).[2]

Hurricane Katrina, which decimated New Orleans public schools in August 2005 along with the rest of the city's infrastructure, was a watershed event in the history of RSD. Among other disruptions caused by the hurricane, many public schools closed and enrollment plummeted.[3] The scramble to reopen New Orleans schools prompted further legislative action in November of 2005. Louisiana's Act 35 allowed RSD to assume control of 114 low-performing New Orleans schools, leaving OPSB with authority over only 17 of the schools it ran before Katrina.

[1] See NCES (2013) for national enrollment statistics by school type and NACPS (2013a) for statistics on charter growth. The latter report notes 531 new charters schools opened and 150 charter schools closed. CREDO (2013a) compares the demographic characteristics of traditional public and charter school students; NACPS (2013b) gives statistics on charter shares by district.

[2] Cowen (2011c) gives a history of RSD.

[3] See Vigdor (2007), Groen and Polivka (2008), Sacerdote (2012), Imberman et al. (2012), and Deryugina et al. (2014) for analyses of the effects of Katrina on outcomes related to education and the labor market.

In the following years, as enrollment grew from the immediate post-Katrina trough, both RSD and OPSB converted increasing numbers of low-performing schools to charters. By Fall 2008, when combined RSD and OPSB enrollment had reached 36,000 (just over half of pre-Katrina OPSB enrollment), the much-reduced OPSB district had 73% of its students in charters, while the RSD charter share hit 49%. Since 2008, RSD charter growth has accelerated, and September 2014 saw the closure of the few remaining direct-run traditional public schools in RSD (OPSB continues to operate a mix of traditional and charter schools). The 2008 school year also marked the beginning of a period of relative stability in RSD enrollment, leadership, and finances, along with district-wide improvements in achievement. RSD achievement gains – in both direct-run and charter schools – can be seen in Figure 1, which compares post-2008 achievement trends in RSD, OPSB, and the rest of Louisiana. Average achievement for traditional and RSD charter students runs mostly below the statewide and OPSB averages, but the gap was much reduced by 2014.

An important and distinctive feature of New Orleans' charter expansion is the fact that most of the RSD charter schools that have opened since 2008 are *takeovers*. A charter takeover occurs when an existing public school, including its facilities and staff, come under charter management. Importantly, takeovers guarantee seats for incumbent students, "grandfathering" these students into the new school. By contrast, most charter schools in other districts open as *startups*, that is, new schools (sometimes in existing school buildings), with no seats guaranteed by virtue of previous enrollment.

Boston's experiment with charter takeovers has unfolded with less urgency than New Orleans', but the forces behind it are similar. At the end of the 2010-2011 school year, nine schools in the Boston Public School (BPS) district were closed for persistently low performance. In an effort to turn two of these schools around, the UP Academy Charter School of Boston replaced the former Gavin middle school, while Boston Green Academy (BGA) replaced the former Odyssey high school, both in Fall 2011. These *in-district charter schools*, known in the state bureaucracy as Type-III Horace Mann schools, mark a new approach to charter authorization and school autonomy in Massachusetts. The Boston School Committee authorizes in-district charter schools and funds them through the BPS general budget like their predecessors. In-district charter teachers are also members of the Boston Teachers Union. Outside of pay and benefits, however, terms of the relevant collective bargaining agreements are waived and these schools are free to operate according to their charters. Boston's in-district charters opened with new school leaders and new teaching staff, employed on an essentially at-will basis, while guaranteeing seats to students formerly at Gavin and Odyssey.[4]

[4]The charter schools studied in our earlier work using lotteries (Abdulkadiroğlu et al., 2011) are known as "Commonwealth charters." Commonwealth Charters are authorized by the state as startups and operate as independent school districts.

This paper evaluates the causal effect of RSD and Boston takeover schools on student achievement with an instrumental variables (IV) strategy that exploits the grandfathering provisions used to fill takeover seats. Grandfathering offers the opportunity to answer new questions about urban school reform. The growing set of estimates exploiting charter school admissions lotteries, while consistently showing large gains for students at urban charters, necessarily captures causal effects only for charter applicants, a self-selected population that may be especially likely to see gains from the charter treatment.[5] By contrast, grandfathered enrollment in charter takeovers is essentially passive: an existing population is guaranteed seats in the new school. Takeover experiments therefore identify causal effects on students who haven't actively sought a charter seat. Grandfathering into takeover charter also identifies charter attendance effects for schools with lotteries that aren't over-subscribed. The grandfathering experiment also allows us to study the effect of charters with unusable or undersubscribed lotteries if these schools were once takeovers.

Our econometric strategy uses grandfathering eligibility indicators to instrument takeover attendance in samples of public middle school students. The grandfathering identification strategy, while appealing on substantive grounds, raises two implementation issues. First, attendance at schools slated for closure (we refer to these as *legacy schools*) may have a direct effect on student achievement independent of subsequent matriculation at the takeover school. The trend in student achievement at closing schools in RSD suggests such violations of the exclusion restriction are indeed a concern. We therefore develop an estimation scheme that allows for legacy school enrollment effects in grandfathering-based IV strategies. A second issue is the heterogeneous nature of the takeover counterfactual, which mixes students at traditional public schools with students who enroll at non-takeover charters. This is of particular concern in RSD, which saw an increasing share of non-takeover charter enrollment over the study period. A simple two-stage least squares (2SLS) procedure addresses the problem of a mixed non-takeover counterfactual.

The empirical results reported here should be of immediate policy interest. The proliferation of traditional public schools that have been closed and reconstituted as charter schools reflects a federal push to encourage states to "...require significant changes in schools that are chronically underperforming and aren't getting better" (Duncan, 2010). The FY2011 federal budget operationalized this goal by adding three billion dollars to around $500 million previously appropriated for School Improvement

[5]Lottery estimates are reported in, e.g., Abdulkadiroğlu et al. (2011), Angrist et al. (2012), Angrist et al. (2013), Dobbie and Fryer (2011), Dobbie and Fryer (2013), Hoxby et al. (2009), and Tuttle et al. (2013). Ravitch (2010) (pp. 141-144) and Rothstein (2011) challenge the external validity of charter treatment effects estimated using lotteries. See also Rothstein's account of high scores at 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." (Rothstein, 2004, p. 82.)

Grants (SIGs). Federal SIGs, which offer up to two million dollars annually per qualifying school, support three restructuring models, one of which is the restart model, described as follows in (USDOE, 2009):

> A restart model is one in which [a local education agency] converts a school or closes and reopens a school under a charter school operator, a charter management organization (CMO), or an education management organization (EMO) that has been selected through a rigorous review process. A restart model must enroll, within the grades it serves, any former student who wishes to attend the school.

RSD and BPS takeover charters qualify for federal support under this heading.

The rest of the paper is organized as follows. Following a brief background discussion in Section 2, Section 3 explains the grandfathering identification strategy and shows how to accommodate possible violations of the exclusion restriction in our instrumental variables framework. Section 4 presents a detailed econometric analysis of takeovers in New Orleans RSD, and interprets these results against the backdrop of rising charter enrollment. Section 5 deploys the grandfathering research design in Boston, comparing grandfathering and lottery estimates of achievement effects at UP Academy. We show here that the unusually large reading gains seen for UP's grandfathering compliers are explained by the weak non-charter outcomes experienced by these students. The final section summarizes our findings and briefly considers implications for policy.

2 Background

2.1 Takeovers in New Orleans RSD

The RSD takeover explosion is documented in Figure 2. Of the RSD charters that have opened since Fall 2008 and were operating in Spring 2014 (excluding alternative schools that serve special populations), 21 are takeovers while 13 are startups. Even by the standards of the heated debate over school reform, the proliferation of charter takeovers in New Orleans and elsewhere has proven to be especially controversial.[6] At the same time, the perception that RSD's takeover policy has been fruitful has prompted ongoing explorations of similar approaches in Michigan and Tennessee.

[6]See, for example, Darling-Hammond (2012), who writes "In the new vision for ESEA, these schools, once identified, will be subjected to school 'turnaround' models that require the schools to be closed, turned into charters, reconstituted ... These approaches have a dubious track record. Many reconstitutions–where staff are fired and replaced–have resulted in a less qualified teaching staff and lower achievement after the reform." A February 2014 civil rights complaint lodged with the US Department of Education's Office of Civil Rights alleges closure of the final five traditional public schools in RSD has hurt the mostly African American students who attended these schools, a complaint that is now under review (Drellinger, 2014).

Appendix Table A1 lists the 18 New Orleans RSD schools that experienced what the district calls a *full charter takeover* between the Fall 2008 and Fall 2013. Full takeovers convert all grades in the legacy school in a single academic year; the takeover school grandfathers legacy students in the relevant grades, and typically opens in the legacy school building.[7] Alternatives to the full takeover model include principal-led conversions and school mergers. We focus on full takeovers because this intervention is broad and well-defined, with a clearly identified grandfathering cohort at the relevant legacy school. Five of the full takeovers we identify are conversions of existing charter schools and are therefore omitted from our study, which focuses on charter effects for passive students who would typically not seek charter seats. The two high schools in the table are also omitted; our analysis focuses on schools with middle school grades (in RSD, these are almost all K-8 schools) because this is where takeovers are most common and because the legacy school scores used here to adjust for violations of the exclusion restriction are readily available for middle schoolers.[8] The decision to effect a takeover at low-performing RSD schools was driven in part by test scores and in part by the availability of a charter operator who met RSD qualifications for school management.[9] Table A1 shows that the 11 legacy schools in our study were taken over by six charter management organizations (CMOs), with the Crescent City and ReNEW CMOs each operating multiple schools. In two cases in 2013, two legacy schools were merged into a single takeover school. Table A1 also shows that 7 out of 9 study takeover schools were operated by CMOs that describe themselves as embracing No Excuses pedagogy.

The No Excuses model for urban education is characterized by extensive use of tutoring and targeted remedial support, reliance on data and teacher feedback, a curriculum focused on basic skills, high expectations from students and staff, and an emphasis on discipline and comportment.[10]

RSD's charter schools function outside the collective bargaining agreement between OPSB and the United Teachers of New Orleans union that represents teachers at non-charter OPSB schools and a few OPSB charters (Cowen, 2011c). Appendix Table A2 compares teacher characteristics at RSD direct-run and charter schools. Teachers at RSD charters tend to be younger, have less experience, and earn lower

[7] With the advent of OneApp in 2012, grandfathering-eligible students who wanted a takeover seat needed only indicate their desire to return to their current school on RSD's common application.

[8] Louisiana allows five types of charters, classified according to whether the school is authorized by the local school board or the LDE, whether the school is new or a conversion, and whether the school is in RSD. RSD's Type 5 charter schools, the focus of our study, are authorized and overseen by the LDE. The takeover/startup distinction is less clear in OPSB than in RSD. Of the 14 charters operating in OPSB in Fall 2013, three are startups and 11 were created in the immediate Katrina aftermath. Although these "Katrina takeovers" were tied to the closure of particular traditional schools, and admissions policies at the new schools reference preferences for those who attended the schools they replaced, for the most part they do not appear to have guaranteed seats to these students. In contrast with RSD charters, four OPSB charters have selective admissions policies.

[9] Operators of the schools in our study typically applied for their Type 5 charters early in the legacy year, with some indicating a preference for specific school buildings. RSD decisions to implement a takeover were typically announced in December of the legacy year, with the charter operator announced between January and May.

[10] An online appendix table lists our sources for this classification.

base salaries than those at direct-run schools. Class sizes at takeover and legacy schools are similar and close to those seen at other charter and direct-run public schools. Per pupil expenditure is somewhat lower at RSD charter schools, though this may reflect compositional differences in the student body and the experience distribution among teaching staff. The PPE contrast between takeover and legacy schools shows only a small gap.[11]

2.2 UP from Gavin Middle School

Our Boston analysis focuses on the UP Academy Charter School of Boston, the middle school in the pair of original Boston in-district charters. The Unlocking Potential charter network is rapidly expanding, with two schools recently chartered in Boston's Dorchester neighborhood (one elementary and one K-8), and two (non-charter) middle schools recently opened in Lawrence, Massachusetts. Our middle school focus necessarily excludes BGA, Boston's in-district charter high school. In this context, it's worth noting that BGA is more of an in-district conversion than a takeover, since it was initially staffed by BPS teachers and administrators from elsewhere in the district.[12]

Boston's in-district model arose in the wake of a 2010 Massachusetts law that allowed BPS to open up to four charter schools without union approval. The in-district model was meant to quickly improve schools with persistently low performance.[13] As in RSD, the birth of an in-district charter reflects both the district's desire to address low school performance and the presence of a willing operator: Unlocking Potential was selected as a in-district operator partly because it was ready to grandfather all Gavin students (Toness, 2010). Gavin students were automatically admitted to UP Boston, though a simple application was required (UP staff visited Gavin students' homes to encourage application).[14]

[11]RSD schools, both direct-run and charter, are funded from local and state taxes using a formula that allocates funding by enrollment. Federal grants (such as Title I funds) flow to direct-run schools through RSD and to charters through their CMOs (Cowen, 2011a). Unlike direct-run schools, charter schools can save unspent funds from one year for use in another (Cowen, 2011b).

[12]BGA's founding headmaster and chairman of the board came from Boston Fenway High School, a Pilot School (see Abdulkadiroğlu et al. (2011) for an evaluation of pilot schools). Concerns about poor record-keeping and continued low performance at the school recently prompted the state commissioner of education to recommend that BGA be put on probation for the remainder of its charter (Vaznis, 2014).

[13]Gavin and Odyssey were ranked at the bottom of a list of low-performing schools in 2010, by some measures the worst on the list. Yet for various reasons, these schools didn't qualify for state designation as "Level 4," a decision that would have made these schools candidates for a number of alternative interventions being used elsewhere in the district. In response to our queries, BPS administrators emphasized that in-district charter conversion was one of the strategies available to the district for schools designated Level 3.

[14]Special arrangements were made for some special education students at Gavin. Of 67 high-needs special education students, 19 stayed at UP and 13 enrolled at another BPS school. Most of the rest likely attended private day schools at district expense (as do many high-needs special education students in the district), as they disappear from Massachusetts' public schools' enrollment database in the takeover year. These cases notwithstanding, the overall UP enrollment take-up rate for grandfathered special education students is close to that for other grandfathered students. Our IV strategy treats all grandfathered students the same, and estimates conditioning on baseline special education status are qualitatively similar to those for the full samples, though often imprecise.

Unlike other charter schools in Boston, which operate as independent districts and are funded by inter-district transfers, UP spending appears in the BPS budget. Former Gavin teachers were free to apply for positions at UP, and a handful did so, but their positions were not grandfathered and none were ultimately hired to work at UP (Knight, 2013). UP administrators and staff are part of the collective bargaining units representing other BPS workers, but the school functions in a looser framework established in memoranda between UP and the district. UP is required to pay collectively bargained wage rates (or more), but school leaders and UP administrators make personnel decisions freely, as in a non-union workplace.

UP's teachers in our sample period are much younger than was the Gavin staff: 60 percent of UP's teachers are no older than 28, as can be seen in column 8 of Table A2. This is unusually youthful even by the standards of Boston's other charter schools. UP class sizes are smaller and per-pupil expenditure is somewhat lower than was the case at Gavin school. Like most of our RSD schools, UP's charter aligns itself with the No Excuses model.[15] The UP school day is two hours longer than the Gavin day had been and UP teachers are expected to report for work each year on August 1.

2.3 Related Research

Dee (2012) uses the test proficiency cutoffs that determine qualification for federal SIG funding to frame a regression discontinuity study of the causal effects of SIG awards. Dee's estimates suggest that SIG-funded interventions improve performance for students at treated schools. His companion difference-in-differences analysis points to the intermediate federal turnaround model as the most effective, while estimates for the remaining two SIG strategies, including restarts, are not significantly different from zero. It's worth noting, however, that very few California schools opted for the more radical restart intervention, and Dee's estimates for the restart treatment are correspondingly imprecise.

Houston's pioneering Apollo 20 program revamped educational practices along No Excuses' lines in 20 of Houston's lowest performing schools, while replacing most school leaders and half of the teaching staff in these schools; a similar effort was undertaken on a smaller scale in Denver. The insertion of charter school best practices in existing public schools provides a natural alternative to the takeover model studied here, and qualifies for the same sort of federal support. Fryer (2014)'s cluster-randomized trial and quasi-experimental analyses of the Apollo makeovers show statistically significant gains in math of between one-fifth and one-sixth of a standard deviation, with little effect on reading. Fryer (2014)'s quasi-experimental analysis uses baseline enrollment zones to construct instruments

[15] UP's charter application states "all stakeholders should not make or accept excuses for anything less than excellence," and describes key No Excuses practices as part of their educational programming (UP Academy, 2010).

for exposure to treatment. Our grandfathering strategy similarly uses a baseline condition to isolate exogenous variation, but the approach here incorporates matching on baseline school characteristics to eliminate covariate differences associated with the grandfathering instrument. Our IV identification strategy also allows for violations of the exclusion restriction that may compromise naive grandfathering estimates.

In a recent report, CREDO (2013b) uses a variety of comparison methods to evaluate the effects of attending three RSD takeover charters. The CREDO study presents a fine-grained analysis that distinguishes many types of students based on their baseline and post-takeover enrollment status, comparing, for example, students who move into and who exit from schools slated for charter conversion. This analysis generates a complicated picture of mixed positive and negative effects, though these sorts of comparisons do not appear to fit into a causal framework except under stronger conditional independence assumptions than those invoked here.

Somewhat further afield, Epple et al. (2013) outline a structural model of school choice in a large urban district with schools slated for closure, estimating the model using lagged endogenous variables as instruments. Finally, as noted in the introduction, our work is closely related to the growing body of research using charter lotteries to identify causal effects of charter school attendance.

3 Grandfathering Identification

3.1 The RSD Comparison Group

Our grandfathering research design uses a combination of matching and regression to mitigate omitted variables bias in comparisons of grandfathering-eligible and ineligible students. To see how the matched comparison group is constructed, consider the set of 6th graders enrolling at an RSD school slated for takeover at year's end: 6th grade legacy school enrollment entitles this group to 7th grade seats in the takeover charter. Since legacy and takeover schools in RSD typically enroll grades K-8, there are few non-legacy 6th graders who share a 5th grade school with the grandfathering-eligible group. We therefore look for a comparison group in the population of 6th graders not enrolled at the legacy school, but who attended schools similar to that attended by legacy school students in 5th grade (we refer to these 5th grade schools as *baseline schools*). Specifically, baseline schools are matched if they have school performance scores (SPS) in the same five-point bin.[16] In addition to baseline schools, we

[16]SPS scores are used for accountability purposes within RSD. Until academic year 2011-2012 (the last baseline year for our sample of RSD takeovers), SPS scores ranged from 0 to 200, and have since transitioned to a 0-150 scale. Matches are stable when smaller bins are used, but bins wider than about 10 points generate a coarse match with many low-scoring schools grouped together.

construct the RSD comparison sample by matching grandfathering-eligible and ineligible students on race, sex, baseline year, baseline special education status, and baseline subsidized lunch eligibility.

In practice, the RSD grandfathering experiment involves multiple grades, schools, and years. The relationship between legacy grades, baseline grades, and takeover grades in each RSD grandfathering scenario is described in Table 1. Because the earliest baseline information available is from 3rd grade, our RSD sample covers legacy school enrollment in grades 4-7 and takeover charter enrollment in grades 5-8. Potential takeover exposure thus ranges from one year (for students in 7th grade in the legacy year) to four years (for students in 4th grade in the legacy year), or more if grades are repeated. A given matching cell may contain students who were eligible for grandfathering into multiple takeover charters. The grandfathering instrument indicates eligibility at any of the 9 takeover schools we study. When pooling across grades, we retain students in the first year they become or are matched to a grandfathering-eligible student. The number of grandfathering-eligible students enrolled in a legacy school in the fall of the year prior to takeover averages roughly 70 students per school and is about one-third the size of the matched comparison group (Appendix Table A3 reports the number of observations contributed by each RSD legacy school).

Our primary analysis outcomes in RSD are Math and English Language Arts (ELA) standardized test scores, measured by the Louisiana Educational Assessment Program (LEAP) in 4th and 8th grade and the Integrated Louisiana Educational Assessment Program (iLEAP) in grades 5-7, from Spring 2011 (the first exposure Spring of the first takeovers in our sample) through 2014.[17] The Data Appendix details the construction of our analysis files from raw student enrollment, demographic, and outcome data. For the purposes of statistical analyses, scores are standardized to the population of RSD test-takers in the relevant subject, grade, and year (excluding students in alternative programs).

Table 2 reports descriptive statistics for the RSD analysis sample and for broader samples of RSD students with the same distribution of baseline grades and years. As can be seen in the first two columns of the table, almost all RSD and RSD charter-bound students (those enrolled in an RSD charter school in the grades following baseline) are black, and most are poor enough to qualify for a subsidized school lunch. RSD charter-bound students have baseline scores near the overall district mean (which is zero by construction). By contrast, students who enroll in takeover charters and those eligible for grandfathering have much lower baseline test scores. For example, the average baseline math score of grandfathering-eligible students in our analysis sample is around 0.27σ below the corresponding

[17]LEAP and iLEAP include multiple-choice and open-answer questions. LEAP scores are used for determining grade-progression according to Louisiana state guidelines. The iLEAP test combines a test of academic standards and a norm-referenced component from the Iowa Test of Basic Skills (ITBS) through 2012-2013. The 2013-2014 iLEAP tests no longer contain the ITBS portion.

RSD population average. This marks an important contrast with baseline achievement in samples of lottery applicants at many oversubscribed charter schools, a group that tends to be positively selected on baseline characteristics.[18]

The RSD comparison group appears to be well-matched to the RSD grandfathering cohort. This is documented in column 5 of Table 2, which reports regression-adjusted differences in variables not used for matching between grandfathering-eligible students and the matched comparison group in our analysis sample. The balance coefficients come from a model that includes a full set of matching-cell fixed effects, with no further controls. These estimates show no statistically significant differences in the limited English proficiency rates or in baseline scores (balance here is mostly achieved by controlling for baseline school SPS bins).

Appendix Table A4 reports follow-up rates and gauges differential attrition in the RSD analysis sample. Follow-up scores are available for almost three-quarters of students in the first two years following a takeover. The follow-up rate declines markedly in years three and four, reflecting RSD's highly mobile low-income population, a pattern seen in other urban high-poverty districts. Importantly, however, the likelihood an RSD student contributes an outcome score to the analysis sample is unrelated to his or her grandfathering eligibility status, and, as shown in column 6 of Table 2, baseline covariates remain balanced in the analysis subsample for which we can measure outcomes.

3.2 RSD Grandfathering Graphics

We motivate the grandfathering identification strategy for RSD with a graphical comparison of achievement trends in the grandfathering-eligible and matched comparison samples. Provided that scores in the eligible cohort and the comparison group move in parallel in the pre-takeover period, differences in score growth between eligible and ineligible students in the post-takeover period offer compelling evidence of a takeover treatment effect. As in Table 2, we regression-adjust these trends by matching-cell fixed effects, with no further controls.

The upper panels of Figures 3 and 4 show remarkably similar pre-takeover trajectories for the math and ELA scores of grandfathering-eligible and ineligible students. The data plotted here are standardized to samples of students at RSD's direct-run schools, so that achievement trends are cast relative to this group. Consistent with RSD's focus on low-performing schools when assigning takeovers, relative achievement at legacy schools declines in the grade before takeover, though the broader comparison group trend is essentially flat (for ELA) or generally increasing (for math). Importantly, the pre-

[18] In the middle school sample analyzed in Abdulkadiroğlu et al. (2011), for example, the baseline math gap between charter applicants and Boston students is around 0.36σ.

takeover dip (reminiscent of the pre-treatment earnings dip documented by Ashenfelter (1978) for applicants to training programs) is mirrored in the matched comparison group.

Matching effectively eliminates baseline differences by grandfathering status, so that simple post-treatment comparisons seem likely to reveal causal effects. We nevertheless present difference-in-differences (DD) style comparisons of achievement growth, a natural econometric starting point. These comparisons appear in the lower panels of Figures 3 and 4, which plot achievement growth in the grandfathering-eligible and ineligible subsamples relative to the baseline grade. Pre-baseline growth differences by grandfathering status are centered around zero, while achievement contrasts after the legacy year strongly favor the grandfathered cohort. Since around 78% of grandfathering-eligible students matriculated at a takeover charter (a figure that appears in Table 4, below), this pattern suggests takeover enrollment significantly boosted achievement.

Figures 3b and 4b show remarkable parallelism in pre-takeover score trends up to, but not including, the year of legacy enrollment. The negative and significant (for math) DD contrast in the legacy year signals a possible causal effect of legacy enrollment *per se*, regardless of whether legacy attendance leads to subsequent enrollment in the takeover charter. This is an unsurprising but potentially important finding: legacy schools were slated for closure in part because of extraordinarily low and even declining achievement, a fact that may have had lasting consequences for their students. Our grandfathering instrumental variables strategy therefore allows for direct effects of legacy school attendance when using legacy enrollment to instrument takeover attendance.

3.3 Econometric Framework

Our grandfathering IV estimator contrasts legacy-to-post-takeover achievement growth by grandfathering status. To interpret this procedure, consider a group of legacy school students and their matched comparison counterparts with covariate values falling in a single matching stratum. Achievement for each student is observed in two grades: at the end of the legacy grade, immediately prior to the takeover (grade l) and after the takeover (grade g). The grandfathering-eligible group is mostly enrolled in the takeover school in grade g, while few in the comparison group are. A dummy variable denoted by Z – the grandfathering instrument – indicates legacy school enrollment in grade l (observed at the start of the school year) while the variable D indicates takeover school enrollment at any time in grade g. Achievement in the two grades is denoted Y^l and Y^g, observed at the conclusion of the school year.

Legacy school enrollment in grade l potentially affects grade g achievement through two causal channels: by increasing the likelihood of takeover attendance in grade g and by adding a year's exposure to the legacy school in grade l, an event that may have lasting consequences if learning is cumulative. Potential outcomes in grade g are therefore double-indexed. Specifically, we write Y^g_{zd} to indicate the grade g outcome that would be observed when $Z = z$ and $D = d$. Potential outcomes in grade l, written Y^l_z, are indexed against Z alone, since grade l predates takeover exposure. Using the potential treatments notation introduced by Imbens and Angrist (1994), legacy enrollment shifts takeover exposure from D_0 to D_1. In this setup, observed outcomes are determined by potential outcomes and by the instrument as follows:

$$
\begin{aligned}
Y^l &= Y^l_0 + Z(Y^l_1 - Y^l_0), \\
D &= D_0 + Z(D_1 - D_0), \\
Y^g &= Y^g_{00} + Z(Y^g_{10} - Y^g_{00}) + D(Y^g_{01} - Y^g_{00} + Z(Y^g_{11} - Y^g_{10} - (Y^g_{01} - Y^g_{00}))) \\
&= Y^g_{00} + Z(Y^g_{10} - Y^g_{00}) + (D_0 + Z(D_1 - D_0))(Y^g_{01} - Y^g_{00} + Z(Y^g_{11} - Y^g_{10} - (Y^g_{01} - Y^g_{00}))),
\end{aligned}
$$

where the last line uses the expression for D to obtain a representation for observed Y^g as a function of potential outcomes, potential treatments, and the instrument, Z.

Potential outcomes and treatments are assumed to satisfy the following assumptions:

Assumption 1 *(Independence)* $\{Y^l_0, Y^l_1, Y^g_{00}, Y^g_{01}, Y^g_{10}, Y^g_{11}, D_0, D_1\} \perp\!\!\!\perp Z$.

Assumption 2 *(Monotonicity)* $Pr(D_1 \geq D_0) = 1$.

Assumption 3 *(First-stage)* $E[D_1 - D_0] > 0$.

Assumption 1 – Independence – asserts that the grandfathering instrument is as good as randomly assigned with respect to potential outcomes and treatment take-up (implicitly, within matching strata). Table 2 and Figures 3 and 4, which show that matching eliminates covariate and baseline score differences in our RSD analysis sample, support this. Monotonicity says that legacy enrollment either induces takeover enrollment or has no effect for all individuals in the analysis sample. Assumption 3 requires legacy enrollment to induce takeover enrollment, at least for some.

As in the Angrist et al. (1996) framework for identification of local average treatment effects (LATE) with possible violations of the exclusion restriction, Assumptions 1-3 allow for possible direct effects of legacy exposure on grade g outcomes. Such effects arise if

$$Y^g_{1d} \neq Y^g_{0d},$$

when D is fixed at d. In other words, maintaining the assumption that legacy enrollment is as good as randomly assigned, we've allowed for violations of the exclusion restriction associated with use of Z as an instrument for D. In view of the low achievement seen at the legacy school, and the close link between legacy attendance and the grandfathering instrument, the possibility of such violations seems inherent in the grandfathering research design.

Rather than defend a conventional exclusion restriction in this setting, we replace it with a closely related but weaker restriction on potential achievement *gains*. This allows for direct additive effects of legacy enrollment that are free to vary within the LATE subpopulations of always-takers, never-takers, and compliers:

Assumption 4 *(Gains Exclusion)* $E[Y^g_{1d} - Y^l_1|T] = E[Y^g_{0d} - Y^l_0|T]$, where $T = aD_0 + n(1 - D_1) + c(D_1 - D_0)$ identifies always-takers (a), never-takers (n), and compliers (c).

Assumption 4 requires that expected potential achievement gains be the same for those who do and don't attend the legacy school in grade l, once takeover enrollment is fixed. This allows $Y^g_{1d} \neq Y^g_{0d}$, while also weakening the canonical exclusion restriction applied to gains, which says that $Y^g_{1d} - Y^l_1 = Y^g_{0d} - Y^l_0$ for everyone, rather than just on average.

We can interpret Assumption 4 as implied by an additive structure for expected potential outcomes in each grade:

$$E[Y^l_z|T = s] = \alpha_{1s} + z\gamma_s \tag{1}$$

$$E[Y^g_{zd}|T = s] = \alpha_{2s} + z\gamma_s + d\beta_s. \tag{2}$$

The parameters α_{1s} and α_{2s} in these expressions are subgroup-specific potential outcome means with both the legacy- and takeover-enrollment indicators switched off; γ_s is an additive legacy school enrollment effect, common to grades l and g; and β_s is the causal effect of takeover attendance for LATE subgroup s. This additive model rules out interactions between legacy and takeover attendance, while allowing legacy effects to be persistent across grades.

The appendix shows that under Assumptions 1-4, a Wald-type IV estimand applied to achievement gains captures the average causal effects of takeover attendance on compliers' grade g achievement as follows:

Theorem 1 Under Assumptions 1-4,

$$\frac{E[Y^g - Y^l|Z = 1] - E[Y^g - Y^l|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]} = E[Y^g_{11} - Y^g_{10}|D_1 > D_0] = E[Y^g_{01} - Y^g_{00}|D_1 > D_0].$$

Proof: See appendix.

This theorem identifies average causal effects of takeover exposure on test score *levels*, notwithstanding violations of the exclusion restriction due to legacy school enrollment (in the notation of equations (1) and (2), the theorem identifies β_c).

We use Theorem 1 in two ways: to capture causal effects of takeover enrollment in the year following a takeover and to capture causal effects of an ordered treatment that counts years of takeover exposure. The latter use is supported by a corollary detailed in the appendix, which shows how our IV estimand for an ordered treatment can be interpreted as a weighted average of incremental average causal effects. To explore the robustness of our conclusions to Assumption 4, the econometric appendix also discusses models and estimates that allow legacy school effects in the takeover grade to only be partially carried over to the legacy grade.

Motivated by these theoretical results, we estimate the causal effects of takeover attendance with a second-stage estimating equation that can be written

$$Y_{it}^{g} - Y_{i}^{l} = \alpha' X_{it} + \sum_{j} \kappa_j d_{ij} + \beta D_{it} + \eta_{it}, \tag{3}$$

where Y_{it}^{g} is student i's score in year t in grade g and Y_{i}^{l} is i's score in the last grade in which he or she was potentially enrolled in the legacy school. The treatment variable here, D_{it}, counts the number of years student i spent at the takeover school as of year t, up to and including the grade enrolled in that year (D_{it} is Bernoulli for tests taken in the first year of takeover operation).

The first stage equation that accompanies this second stage is

$$D_{it} = \delta' X_{it} + \sum_{j} \mu_j d_{ij} + \pi Z_i + \nu_{it}, \tag{4}$$

where Z_i is the excluded instrument, an indicator of legacy enrollment in the fall of the legacy school's final year in operation, and π is the associated first stage coefficient. As with the models used to investigate covariate balance, equations (3) and (4) control for matching cell fixed effects. In particular, because the comparison sample consists of an exact match on race, sex, baseline special education status, baseline subsidized lunch eligibility, baseline school, baseline year, and the legacy grade, equations (3) and (4) include dummies for each of these cells, denoted d_{ij} for cell j. The empirical first- and second-stage models also include dummies for English proficiency and year-of-test (denoted by the vector X_{it}, with coefficients α and δ). Finally, although baseline score controls appear to be uncorrelated with grandfathering exposure in RSD, X_{it} includes these score controls to boost precision.

4 Charters without Lotteries in New Orleans RSD

4.1 Grandfathering Results

Attendance at RSD takeover charters is estimated to increase math and ELA scores by an average of 0.21σ and 0.14σ, respectively, per year enrolled. These IV estimates, reported in the last column of

Table 3, are generated by a first stage of about 1.1 years of takeover exposure (first stage estimates are reported in column 3). The associated standard errors are on the order of 0.04.[19] Analyses that disaggregate by outcome grade and by years of potential takeover exposure show that takeover effects are larger in 7th and 8th grade than earlier, and are larger in the first two years of takeover exposure than later. The first stage effect of grandfathering eligibility on first-exposure-year enrollment, reported at the top of panel B, reveals that grandfathering boosted initial takeover enrollment rates by about 66 percentage points.

The IV estimates generated by the grandfathering design exceed (and, in many cases, are significantly different from) the corresponding OLS estimates reported in column 2 of Table 3. This suggests that uninstrumented comparisons by takeover enrollment status, such as those reported in CREDO (2013a), reflect substantial negative selection bias. It's also worth noting that IV estimates that fail to adjust for legacy enrollment, such as those reported in Fryer (2014), would appear to be biased downwards. Fitting versions of equations equations (3) and (4) to post-treatment levels rather than gains generates math and ELA effects of 0.16σ and 0.11σ, respectively. Differences between these estimates and those for gains are consistent with the negative legacy-year treatment effects suggested by Figures 3 and 4. Appendix Table A7 reports legacy year treatment effects and estimates of models that weaken Assumption A4 to allow for partial pass-through of legacy effects. See the econometric appendix for details.

4.2 Interpreting RSD Takeover Effects

The RSD grandfathering identification strategy compares students that mostly attend takeover charters with a grandfathering-ineligible comparison group that went to various sorts of schools. Most students in the comparison group began middle school at one of RSD's direct-run public schools. But the distribution of takeover alternatives evolved as RSD closed its direct-run schools and as students changed schools for reasons other than closure, entering charters through lotteries instead of by grandfathering. Estimates of RSD takeover effects therefore reflect a growing share of charter-to-charter comparisons. If *all* RSD charters boost student achievement, such comparisons mask a higher overall charter treatment effect.

Table 4 describes the grandfathering attendance counterfactual in detail, focusing on the distinction between the charters that define the takeover treatment for purposes of Table 3 (defined as "study takeovers"), other takeover schools (including charter-to-charter conversions), non-takeover RSD charters, and direct-run RSD schools. Specifically, the first two columns show the distribution of school

[19]Estimates of effects on science and social science are similar, and are reported in the online appendix.

types by grandfathering status, while column 3 describes the types of schools attended by untreated compliers. These complier attendance counterfactuals were constructed by estimating causal effects of the takeover enrollment dummy, D, on school sector indicators, W. Associated with each W are potential attendance outcomes, W_0 and W_1, describing school choices in non-treated and treated states (that is, potential school type when $D = 0$ and $D = 1$). Column 3 of Table 4 reports estimates of $E[W_0|D_1 > D_0]$, the distribution of school types among compliers when they do not enroll in a takeover (these were estimated using the weighting procedure derived by Abadie (2003)).[20] By definition, treated compliers enroll in a takeover school when they're grandfathering-eligible; column 4 in the table is included as a reminder of this fact.

The first stage for enrollment in a study takeover contrasts a 78 percent first-year takeover enrollment rate for those grandfathered (reported in column 2 of Table 4) with an 8.9 percent comparison group enrollment rate (reported in column 1 of Table 4).[21] The first-year increase in (study) takeover enrollment reflects a substantial reduction in rates of attendance at non-takeover charters (compare 33 with 15) and, especially, a sharp reduction in attendance at direct-run schools (compare 51 with 3.4). The counterfactual attendance distribution in column 3 shows that 32 percent of untreated compliers enrolled initially in a non-takeover charter school, while 60 percent attended a direct-run school.

Not surprisingly, both the takeover first stage and the proportion of the comparison group in direct-run schools falls over time. The (study) takeover first stage in the third year of exposure is around 0.48 (0.754-0.277), while the counterfactual direct-run enrollment share falls to about 0.27. The balance of third-year non-treated complier enrollment was in other RSD charter schools. Reflecting RSD's accelerating charter transformation, the other-charter enrollment rate for compliers in our sample exceeded 86 percent after four years of exposure.

The growing share of the RSD comparison sample enrolled in charter schools dilutes estimated takeover effects if other charter schools generate similar gains. This motivates a 2SLS model with two endogenous variables, one tracking study takeover attendance and one tracking attendance at other sorts of charters. The model with two types of treatments can be written

$$Y_{it}^g - Y_i^l = \alpha' X_{it} + \sum_j \kappa_j d_{ij} + \beta_D D_{it} + \beta_C C_{it} + \eta_{it}, \tag{5}$$

where C_{it} counts the number of years of attendance in charters other than those covered by D_{it} prior to testing. Equation (5) is identified by adding interactions between the grandfathering instrument

[20] The weights in this case are given by $\kappa_0(X) = (1-D)\frac{E[Z|X]-Z}{E[Z|X](1-E[Z|X])}$. We estimate $E[Z|X]$ using probit with the same covariate specification used to construct the estimates in Table 3. See Table 1 in Abdulkadiroğlu et al. (2014a) for a similar analysis of school choice in a sample of applicants to Boston and New York exam schools.

[21] The discrepancy between these estimates and the first stage in Table 3 is due to the omission (from estimates in Table 4) of controls for matching strata and other covariates.

and covariates to the instrument list (specifically, interactions with baseline year, special education status, and SPS five-point bins). These interactions generate a first stage for C_{it} because students with differing characteristics (covariate values) are more or less likely to wind up in non-takeover charters in the event they aren't grandfathered.

Removing other charters from the counterfactual outcome distribution with the aid of equation (5) nearly doubles the estimated takeover effect on math scores. This can be seen in the contrast between the estimates in columns 1 and 3 of Table 5. Column 1 repeats the takeover effect for the all-grades sample shown in Table 3, while column 3 reports 2SLS estimates of β_D and β_C. The takeover estimate for math in the latter specification rises to 0.36σ, while the other RSD charter effect is 0.32σ. These results are remarkably similar to the estimates of middle school math effects for Boston charter lottery applicants reported in Abdulkadiroğlu et al. (2011). On the other hand, the other-charter ELA effect in column 3 is close to zero. Consequently, the takeover effect on ELA scores remains near 0.14σ with or without a second endogenous variable to capture other-charter attendance effects.

The estimates in column 3 of Table 5 suggest takeover and other charters have similar effects on math scores. We can therefore construct more precise estimates of this common charter effect by estimating a version of equation (5) that replaces $\beta_D D_{it} + \beta_C C_{it}$ with $\beta_A A_{it}$, where the variable $A_{it} = D_{it} + C_{it}$ counts years of attendance at any RSD charter. The resulting estimates of β_A, reported in columns 4 and 5 of Table 5 for just-identified and over-identified specifications (that is, without and with covariate interactions in the instrument list), indeed show a precision gain, with standard errors falling from 0.073 and 0.152 in column 3 to 0.059 in column 5. The pooled specification for ELA generates a similar reduction in standard errors. It should be noted, however, that the divergence in estimated takeover and other-charter effects in column 3 make the pooled ELA results harder to interpret.[22]

5 Measuring UP in Boston

Estimates from RSD suggest charter takeover attendance increased middle school achievement sharply. At the same time, RSD's transformation to an all-charter district complicates the interpretation of RSD takeover effects. The 2011 takeover of Gavin middle school affords another opportunity to measure charter takeover effects with the grandfathering research design, in this case against a more homogeneous and stable backdrop. The availability of over-subscribed admission lotteries at UP also allows a

[22]The common-effects model produces a weighted average of β_D and β_C, but the weighting scheme in this case need not be convex. The fact that the estimates in columns 4 and 5 exceed those in columns 1 and 3 reflect the negative weight this scheme assigns to the other-charter effect.

direct comparison of results from lottery and grandfathering research designs.

5.1 The UP Comparison Group

As in the analysis of RSD, we use a combination of regression and matching to reduce omitted variables bias in grandfathering comparisons. Middle schoolers grandfathered into UP were enrolled at Gavin in 6th or 7th grade in the fall of 2010. Because both Gavin and UP serve grades 6-8, we match each grandfathered student to non-Gavin students who attended the same school in 5th grade. The Gavin comparison group consists of non-Gavin students matched on 5th grade school, and on race, sex, 5th grade special education status, and 5th grade subsidized lunch eligibility (the Boston column of Table 1 describes the timing of the grandfathering research design for UP). Each grandfathered student is again matched to one or more comparison students. The resulting analysis sample contains 290 grandfathering-eligible Gavin students, with 913 students in the comparison group; a total of 1,147 of these have baseline score data.[23]

On-track 6th and 7th graders at Gavin transitioned to 7th and 8th grade when UP opened in Fall 2011. Achievement outcomes come from 7th and 8th grade Massachusetts Comprehensive Assessment System (MCAS) tests given in Spring 2012-2014. For the purposes of statistical analyses, MCAS scores were standardized to the population of BPS and Boston charter students from the relevant subject and year, excluding students in alternative schools.

Most BPS 5th graders are black or Hispanic, a fact documented in the first two columns of Table 6, which describes the population of Boston 5th graders in years covered by the UP analysis sample along with the subsample of Boston students headed for a charter middle school in grades 6-8. Like other charter students, those at UP or who were grandfathering-eligible are even more likely to be black, while Hispanics are under-represented in the charter-bound and grandfather-eligible groups. Almost all UP and grandfathering-eligible students qualify for a subsidized lunch. In contrast with the positive selection seen in the wider sample of charter-bound students in Table 6, UP students and those eligible for grandfathering into UP in the analysis sample have baseline scores well below those of students in the general BPS population.

The extent to which matching on baseline characteristics produces balanced grandfathering comparisons is explored in the last three columns of Table 6. The estimates in column 5 are from models that control only for matching cells; these show significant grandfathering gaps in baseline scores. The differences in column 5 suggest the comparison group here is not as well-matched as for RSD. Impor-

[23] As can be seen in Table 1, baseline information in the Boston sample comes from 5th grade for students whose legacy grade was 7th as well as for those whose legacy grade was 6th.

tantly, however, the difference in baseline scores can be eliminated by conditioning on a further lagged score. The power of lagged score controls to produce balanced comparisons is illustrated in column 6 of the table, which shows the results of including fourth grade (pre-baseline) scores in the model used to construct the balance estimates. The addition of these controls eliminates the grandfathering gap in 5th grade scores. In other words, lagged score controls neutralize differences in measured achievement in a subsequent pre-takeover grade. All estimates of UP takeover effects thus control linearly for baseline achievement.

Follow-up scores are available for 80-90 percent of our grandfathering-eligible and matched comparison groups, a somewhat higher follow-up rate than for RSD over the same horizon. One year out, differences in follow-up between the grandfathered and comparison groups are small and not significantly different from zero when estimated with lagged baseline score controls, a result shown in Appendix Table A4. Follow-up differences are somewhat more pronounced for the cohort seen two years out, though these differences are still only marginally significant. This modest difference in follow-up rates seems unlikely to account for the large score advantage our analysis uncovers for the grandfathered group.

5.2 UP Grandfathering DD

Achievement in the Gavin grandfathering cohort and the matched comparison group move largely in parallel in pre-takeover grades, diverging thereafter. This is apparent in Figures 5 and 6, which plot achievement paths in the same format used for RSD in Figures 3 and 4. 95% confidence bands for difference-in-differences comparisons are plotted with dotted lines in the bottom panel of these figures. The solid lines compare score growth in the grandfathered and comparison groups, relative to scores from the year preceding the last year of legacy enrollment. These DD estimates show marked and statistically significant differences in score growth in post-treatment years, with no significant differences earlier.

Interestingly, and in contrast with RSD, the Gavin experiment generates a positive DD estimate for legacy-year math scores (of about one tenth of a standard deviation). This result is marginally significant at best and may therefore be a chance finding. Taken at face value, however, this modest gain may reflect an effort by Gavin staff to improve outcomes in advance of–and perhaps in response to–the threat of school closure.

5.3 UP Estimates

The UP enrollment change induced by grandfathering Gavin students boosted middle school math and ELA scores by an average of $0.3\sigma - 0.4\sigma$ per year. This can be seen in the pooled IV estimates of equation (3) reported at the top of Table 7. The first stage that generates these results is just over one, meaning grandfathering eligibility generated an additional year at UP, on average, an estimate shown in column 3 of the table.

The first stage estimate for the cohort that took the 7th grade MCAS reveals the proportion of grandfathered 6th graders who remained at UP; this estimate, shown in the set of results by grade of test reported in panel A of Table 7, is around 80 percent. Math estimates by grade tested are similar across grades, but the ELA estimate for 7th graders, indicating a score gain of almost two-thirds of a standard deviation, is more than double that for 8th graders. Such large gains in reading skills are rarely seen in research on school reform.

Panel B of Table 7 reports results disaggregated by potential years of exposure to UP, contrasting estimates for the cohort that attended for at most one year, a Bernoulli treatment, with estimates for those exposed for up to two years. The average causal effect of a year at UP on students' ELA scores falls after the second year of exposure, from 0.5σ to 0.27σ. Given the exceptionally large first-year ELA impact this seems unsurprising, and is consistent with Figure 6's difference-in-differences evidence for ELA, which shows a post-takeover achievement jump, followed by a plateau.[24]

5.4 UP Lottery Estimates

Since Fall 2012, UP Academy, like other Boston charters, has filled its 6th grade seats through open lotteries, with priority going to current BPS students. Earlier, UP used lotteries to allocate seats not taken by grandfathering-eligible students. A natural benchmark for the Gavin grandfathering strategy is the causal effect of charter attendance on UP students who participated in the lotteries used to fill the 7th grade seats not taken by former Gavin students in Fall 2011, and to fill all 6th grade seats then and since (few students apply for 8th grade seats at UP).

Our UP lottery sample includes applicants who applied for 6th grade seats in the school years beginning in Fall 2011 and Fall 2012, the first two years of UP operation, when we also measure outcomes for grandfathered cohorts. We also look at smaller number of lottery applicants for 7th grade seats in 2011. Lotteries for other entry grades through Fall 2013 were not oversubscribed by first-round, non-sibling BPS applicants. Outcome data are from 6th-8th grade tests, taken in Spring

[24] UP results without differencing post-takeover and legacy-grade scores are similar to those reported in Table 7, at 0.43σ for math and 0.24σ for ELA.

2012-2014. Baseline scores for the lottery sample are from 5th grade for applicants for 6th grade seats and from 6th grade for applicants for 7th grade seats. As with the grandfathering estimates, the lottery sample is limited to students who attended a BPS elementary school in the baseline grade. Appendix Table A5 gives an account of our UP applicant data processing procedures.

Appendix Table A6 describes the UP lottery sample and documents baseline covariate balance by win/loss status. We construct two indicators for lotteried admission: an *initial offer* dummy indicates that a student was given an offer to attend on lottery night, while a *waitlist offer* indicator is switched on for students with the best randomly-ordered waitlist ranks (specifically, the waitlist instrument indicates applicants with lottery numbers below the highest number offered a seat in the relevant application cohort through September). This table also compares lottery applicants with the sample eligible for grandfathering into UP and with a general Boston sample that includes students in the same baseline grades and years attending BPS schools plus Boston charters (excluding those at alternative schools). In comparison to the Boston population, Black students are somewhat over-represented and Hispanic students under-represented among UP lottery applicants. Poverty rates, special education status, and limited English proficiency rates for lottery applicants are much like those seen elsewhere in Boston. Consistent with random assignment of lottery offers, the balance coefficients in columns 5-6 of Table A6 show UP lottery winners and losers are similar. Likewise, we see little evidence of excess loss to follow-up in the loser group.

Importantly, while UP lottery applicants' share many characteristics with the majority of Boston students in the same grade, and their baseline scores are not very different from the overall Boston mean, lottery applicants' baseline achievement far exceeds that in the grandfathering sample, which has baseline scores roughly a quarter of standard deviation below those for Boston. UP lottery applicants are also less likely than the grandfathered cohort to be poor enough to qualify for a subsidized lunch.

The lottery IV framework looks much like that described by equations (3) and (4) for grandfathering instruments, with three modifications. First, there's no matched comparison sample. Rather, the estimation sample consists of all lottery applicants, while the empirical models adjust for year and grade of application (that is, lottery "risk sets") instead of matching cell fixed effects. Second, the dependent variable is the level of Y_{it}^g and not the gain relative to a legacy year. Finally, we use two lottery instruments. The lottery estimating equations can therefore be written

$$Y_{it}^g = \alpha' X_{it} + \sum_j \kappa_j d_{ij} + \beta D_{it} + \eta_{it} \tag{6}$$

$$D_{it} = \delta' X_{it} + \sum_j \mu_j d_{ij} + \pi_1 Z_{1i} + \pi_2 Z_{2i} + \nu_{it}, \tag{7}$$

where Z_{1i} and Z_{2i} are the initial and waitlist offer instruments, dummies d_{ij} indicate lottery risk sets, and X_{it} is a vector of additional controls. The endogenous variable in this context counts years enrolled at UP between lottery application and outcome test date.[25]

The first stage effect of an immediate offer, close to 0.8 for the full sample, exceeds the first stage for waitlist offers, which is just under 0.6. These estimates appear at the top of columns 3 and 4 in Table 8. Looking at first stage effects in the first year of possible exposure to UP, we see that immediate and waitlist offers boost UP enrollment rates by 0.52 and 0.4, respectively. These estimates appear in the first row of panel B in Table 8. UP lottery applicants offered a seat in 6th and 7th grade admissions lotteries earned higher math and ELA scores as a result. Pooled 6th-8th grade 2SLS estimates, reported at the top of the last column of Table 8, show statistically significant average per-year score gains of 0.27σ in math and 0.12σ in ELA. Estimates by grade tested, reported below the pooled estimates, suggest these gains are largest in 6th grade, where they reach 0.35σ in math and 0.21σ in ELA. Likewise, disaggregation by exposure time generates larger average effects after one year than after year two.

The results in Tables 7 and 8 suggest that the benefits of UP enrollment for those enrolled there by virtue of grandfathering are at least as large as for UP students who won their seats in a lottery. The contrast in ELA estimates also favors grandfathering when we compare students who had equal exposure to UP: after one year, gains for the lottery cohort are 0.37σ in math and 0.22σ in ELA, while gains after one year for those grandfathered into UP come to 0.33σ in math and 0.5σ in ELA in 7th and 8th grade. Gains for the grandfathered cohort after two years of potential exposure are estimated to be 0.32σ in math and 0.27σ in ELA. This can be compared with estimated gains of 0.24σ in math and 0.08σ in ELA for similarly-exposed lottery cohorts.

As in our analysis of RSD takeover effects, an important consideration in this context is the type of school attended by compliers not enrolling in UP. Differences in counterfactual school selection might account for the somewhat smaller achievement gains seen for lottery compliers. Perhaps an especially large fraction of those not offered seats in UP lotteries wound up at other high-performing Boston charters, thereby diluting lottery-generated treatment effects.

Roughly 86 percent of untreated compliers in the grandfathering research design enrolled in a traditional BPS school, with 7 percent winding up in another Boston charter. This can be seen in Table 9, which details UP first stages and counterfactual school choices in the same format as Table 4

[25] In addition to lottery risk sets (application year and grade), our lottery analysis controls for student race, sex, special education status, limited English proficiency, subsidized lunch status, baseline test scores, and outcome year and grade effects. As with the grandfathering specification, additional baseline controls yield more precise, though qualitatively similar, estimates.

(likewise computed using the weighting formula in Abadie (2003)). By way of comparison, the lottery design leaves 94 percent of untreated compliers in a traditional BPS school, with only 6 percent in other charters. Counterfactual enrollment rates for both designs appear in column 3 of the table. The low proportion attending other charters, and the even smaller proportion at other charters among lottery compliers, imply that the excess of grandfathering over lottery estimates of UP attendance effects is not explained by a diluted counterfactual in the lottery control group.

5.5 Explaining the Grandfathering Advantage at UP

What accounts for the large ELA gains experienced by UP's grandfathering compliers relative to those experienced by UP's lottery compliers? Differences in average causal effects are necessarily generated by differences in average potential outcomes in treated and untreated states. We explain the grandfathering/lottery contrast in takeover impact by comparing marginal means of counterfactual outcomes in the relevant compliant subpopulations.

As a preliminary for this comparison, note that the 2SLS procedure producing the lottery estimates reported in Table 8 is founded on the presumption that lottery instruments satisfy a traditional exclusion restriction. This restriction can be expressed formally by writing

$$Y^g_{1d} = Y^g_{0d} \equiv Y^g_d,$$

where Y_{zd} is a potential outcome indexed against lottery offers (Z) and (Bernoulli) enrollment at UP (D). Abadie (2002) shows that in a LATE framework with potential outcomes Y^g_1 and Y^g_0 satisfying independence, first-stage, monotonicity, and exclusion restrictions, a simple 2SLS procedure identifies $E[Y^g_1|D_1 > D_0]$ and $E[Y^g_0|D_1 > D_0]$. We implement this procedure for lottery instruments by using 2SLS to estimate equations of the form:

$$Y^g_{it}(1 - D_i) = \alpha_0' X_{it} + \sum_j \kappa_{0j} d_{ij} + \beta_0(1 - D_i) + \eta_{0it} \tag{8}$$

$$Y^g_{it} D_i = \alpha_1' X_{it} + \sum_j \kappa_{1j} d_{ij} + \beta_1 D_i + \eta_{1it}, \tag{9}$$

where X_{it} is the vector of covariates used to construct the estimates in Table 8. The parameters β_1 and β_0 describe outcome means for treated and untreated compliers, that is, $E[Y^g_1|D_1 > D_0]$ and $E[Y^g_0|D_1 > D_0]$ (in this case, averages weighted over matching strata).[26]

[26] The 2SLS procedure used here offers an alternative to the semiparametric Abadie (2003) weighting scheme used to construct the complier counterfactual enrollment outcomes reported in Tables 4 and 9. We prefer the 2SLS procedure in this case so that the estimates of (8) and (9) will match the over-identified 2SLS estimates in Table 8 that mark our starting point. 2SLS estimation also simplifies the computation of standard errors. In practice, weighting and 2SLS generate very similar results.

In contrast with lottery potential outcomes Y_d^g, the grandfathering counterpart of marginal mean potential outcomes for lottery compliers involves doubly-indexed Y_{zd}^g. Because lottery compliers have no experience of the legacy school that induces enrollment for grandfathering compliers, we focus here on identification of the average of Y_{01}^g and Y_{00}^g, that is, the potential outcomes experienced when students go to a takeover school without also attending a legacy school.

Identification of marginal means for grandfathering compliers is further complicated by use of a differenced dependent variable in the grandfathering models, that is, $Y_{it}^g - Y_i^l$ instead of Y_{it}^g on the left-hand side. The principal econometric challenge in this context is the move from counterfactual average gains, $E[Y_{01}^g - Y_0^l|D_1 > D_0]$ and $E[Y_{00}^g - Y_0^l|D_1 > D_0]$ for grandfathering compliers to counterfactual average levels, $E[Y_{01}^g|D_1 > D_0]$ and $E[Y_{00}^g|D_1 > D_0]$, so that the latter can be compared with $E[Y_1^g|D_1 > D_0]$ and $E[Y_0^g|D_1 > D_0]$ for lottery compliers. A simple solution for this problem is outlined in the econometric appendix.

The fruits of our analysis of marginal mean potential outcomes appear in Table 10, starting with treatment effects for the effects of UP attendance in the first year of exposure in columns 1 and 2. These estimates, which match those reported in panel B of Tables 7 and 8, show roughly equal grandfathering and lottery estimates for math (of about 0.35σ), along with grandfathering estimates for ELA that are markedly larger than the corresponding lottery estimates. The grandfathering-lottery gap here is a statistically significant 0.28σ.

An important difference between lottery and grandfathering compliers is apparent in columns 7 and 8 of Table 10: in both subjects, lottery compliers have average untreated outcome means slightly above those for the reference population, while the untreated outcomes means for grandfathering compliers are about -0.15σ below the reference average. The UP treatment effect in math pushes both groups of compliers in parallel: the common gain of 0.35σ leaves $E[Y_{01}^g|D_1 > D_0]$ for grandfathering compliers and $E[Y_1^g|D_1 > D_0]$ for lottery compliers about as far apart as in the untreated state.

In a clear contrast to the math results, the bottom of Table 10 reveals roughly equal treated-state ELA scores for grandfathering and lottery compliers: the difference in estimates of $E[Y_{01}^g|D_1 > D_0]$ for grandfathering compliers and $E[Y_1^g|D_1 > D_0]$ for lottery compliers is only 0.05σ and well short of statistical significance. At the same time, UP grandfathering compliers start almost a quarter of standard deviation below their lottery counterparts. Specifically, the difference in estimates of $E[Y_{00}^g|D_1 > D_0]$ for grandfathering compliers and $E[Y_0^g|D_1 > D_0]$ for lottery compliers is around -0.23. Thus, the UP ELA treatment effect of nearly half a standard deviation for grandfathering compliers reflects the fact that UP pushes this group's reading level from well below that of lottery compliers to effective parity. In this sense, "low Y_0" explains large ELA effects for the grandfathered group. At the same time, the absence of a differential impact for math (with an equally low starting point for grandfathering compliers) shows that a poor non-charter counterfactual is not a sufficient condition for exceptionally large charter achievement gains.

6 Summary and Conclusions

Charter school takeovers in the New Orleans Recovery School District appear to have generated substantial achievement gains for a highly disadvantaged student population that enrolled in these schools passively. Our analysis uncovers even larger effects for students grandfathered in to Boston's first in-district charter middle school, UP Academy. These results contribute to a growing body of evidence showing large positive effects of No Excuses charter schools on a range of outcomes.

Our econometric framework addresses important methodological issues that arise in the grandfathering research design. First, while legacy school enrollment provides a valuable source of exogenous variation in charter exposure, the grandfathering IV strategy should adjust for violations of the exclusion restriction due to legacy exposure itself. Second, in an environment with schools of many types, charter treatment effects may be diluted by charter attendance in the control group. A simple 2SLS procedure allowing for multiple treatment channels generates an easier-to-interpret counterfactual. In practice, cleaning up the non-charter counterfactual substantially boosts our estimates of RSD takeover effects on math, from about 0.21σ to 0.36σ, while leaving the smaller ELA estimates largely unchanged. Finally, we show how an unusually large reading gain for UP's grandfathering compliers can be explained by exceptionally low achievement in the counterfactual state. The grandfathering research design can be applied to other incumbency-based interventions, not necessarily involving charter schools, so these solutions should be useful elsewhere as well.

The strong results for RSD and the comparison of estimates from grandfathering and lottery-based research designs for Boston's UP Academy weigh against the view that urban charter lottery applicants enjoy an unusually large and therefore unrepresentative benefit from charter attendance because they're highly motivated or uniquely primed to benefit from the education these schools offer. Boston and RSD takeovers generate gains for their passively enrolled students that are broadly similar to, and sometimes even larger than, the lottery estimates reported in Angrist et al. (2013) for a sample of Massachusetts urban charters. The achievement gains generated by takeover enrollment also appear to exceed those seen in the Apollo experiment, which introduced No Excuses practices into traditional public schools, without implementing a full takeover (these results are detailed in Fryer (2014)).

Our findings echo those of Walters (2014) in highlighting the importance of charter access. In a pioneering effort to streamline charter admissions and broaden school choice, RSD introduced a district-wide school assignment scheme known as OneApp in 2012. OneApp combined all RSD schools, direct-run and charter, using the tools of market design to increase access and improve school-to-student matching (Abdulkadiroğlu et al., 2014b). Denver, the District of Columbia, and Newark use similarly unified enrollment systems (Ash, 2013). Other districts, however, including OPSB, have resisted attempts to centralize school assignment in general and to integrate charter and direct-run assignment in particular. The results reported here suggest the possibility of substantial gains from centralized school assignment mechanisms, like OneApp, that promote charter attendance among students who might not otherwise choose to apply.

Figure 1a: Math scores in RSD and elsewhere

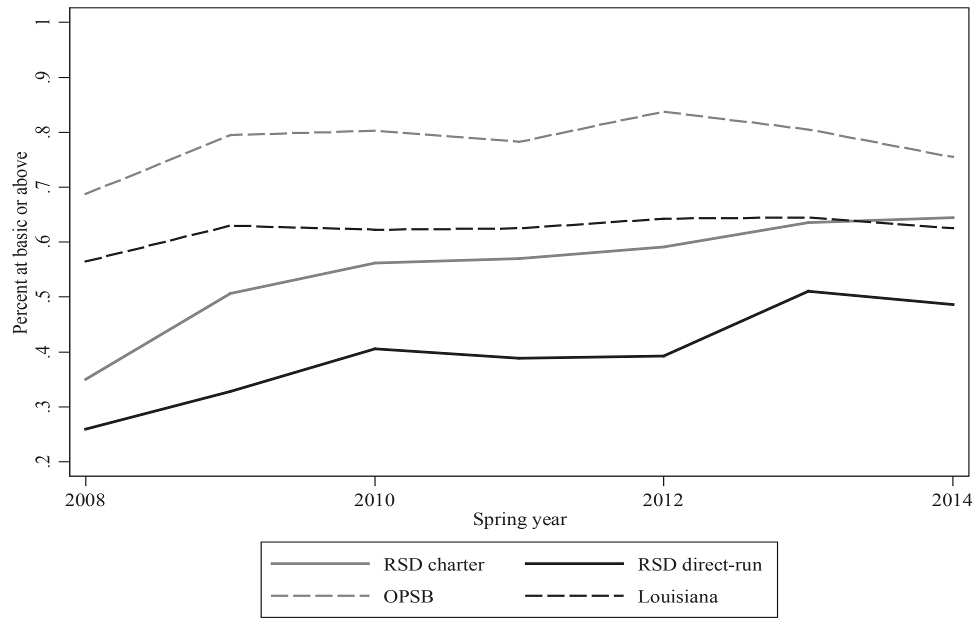
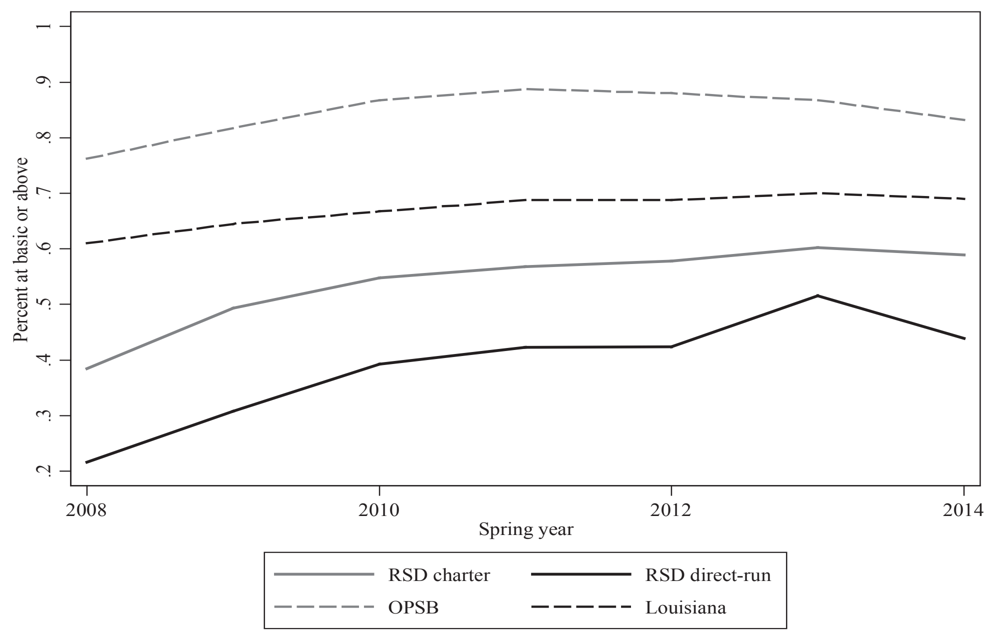
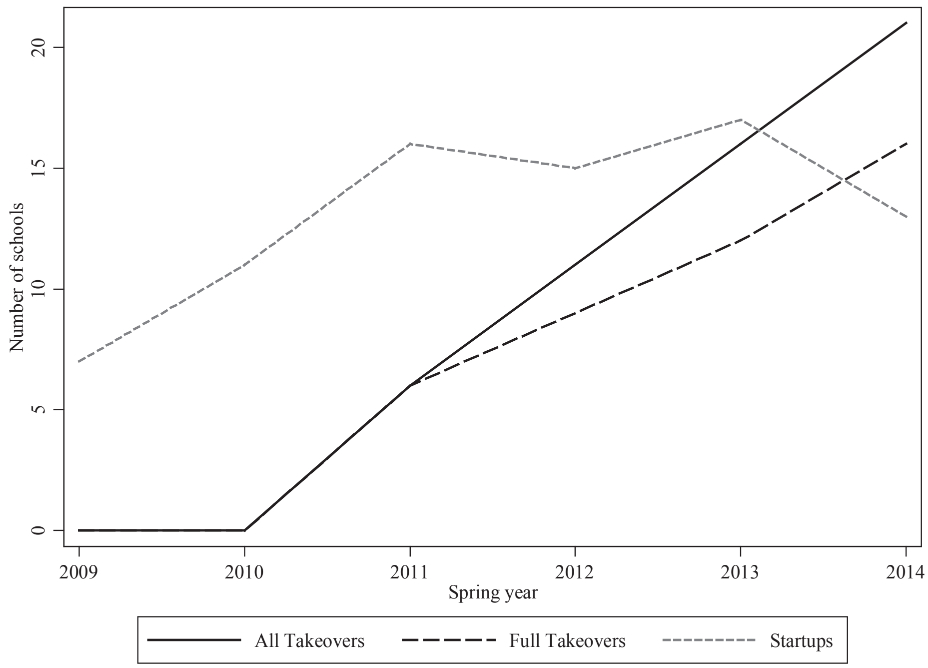


Figure 1b: ELA scores in RSD and elsewhere



Notes: These figures plot the average percentage of RSD, OPSB, and Louisiana students that achieve basic or above status on LEAP/iLEAP math (Figure 1a) or ELA (Figure 1b) exams in 5th-8th grades. Scores for OPSB and Louisiana are from https://www.louisianabelieves.com/resources/library/test-results. Statistics plotted are unweighted averages across grades for each year, and are computed separately for students enrolled in RSD charter and RSD direct-run schools.

Figure 2: Charter school expansion in RSD



Notes: This figure plots the number of New Orleans Recover School District charter schools (serving any grades) created between academic years 2008-09 and 2013-14 (excluding alternative schools). Takeovers are charter schools tied to closure of a legacy school, with seats reserved in the new school for legacy school students. Full takeovers are takeover schools (excluding charter mergers and principal-led conversions) that grandfather all grades at the legacy school in a single academic year. Startup schools are those not directly tied to a legacy school, with all seats filled in the first year through open enrollment.

Figure 3a: Math scores in the RSD grandfathering sample

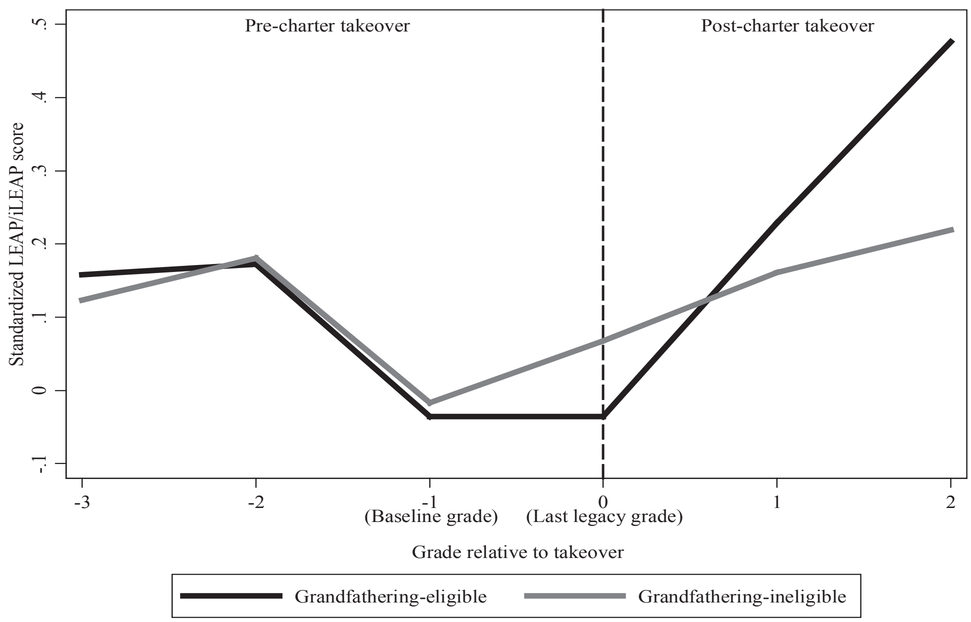
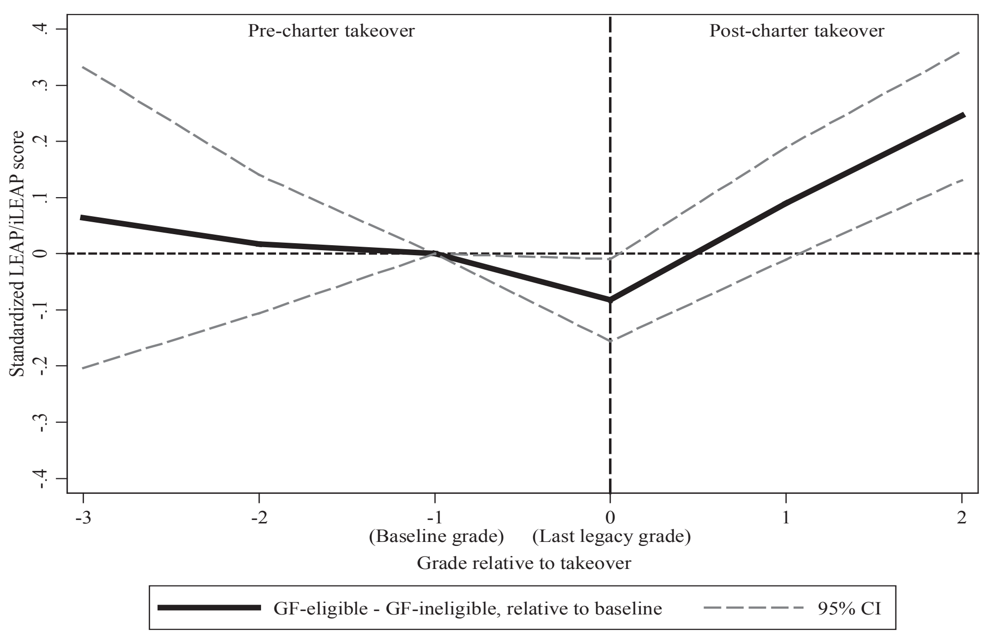


Figure 3b: Grandfathering DD for math



Notes: Figure 3a plots average LEAP/iLEAP math scores of students in the RSD legacy middle school matched sample. Figure 3b plots achievement growth relative to the baseline grade. Estimates in both figures control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in the set of direct-run schools in New Orleans RSD.

Figure 4a: ELA scores in the RSD grandfathering sample

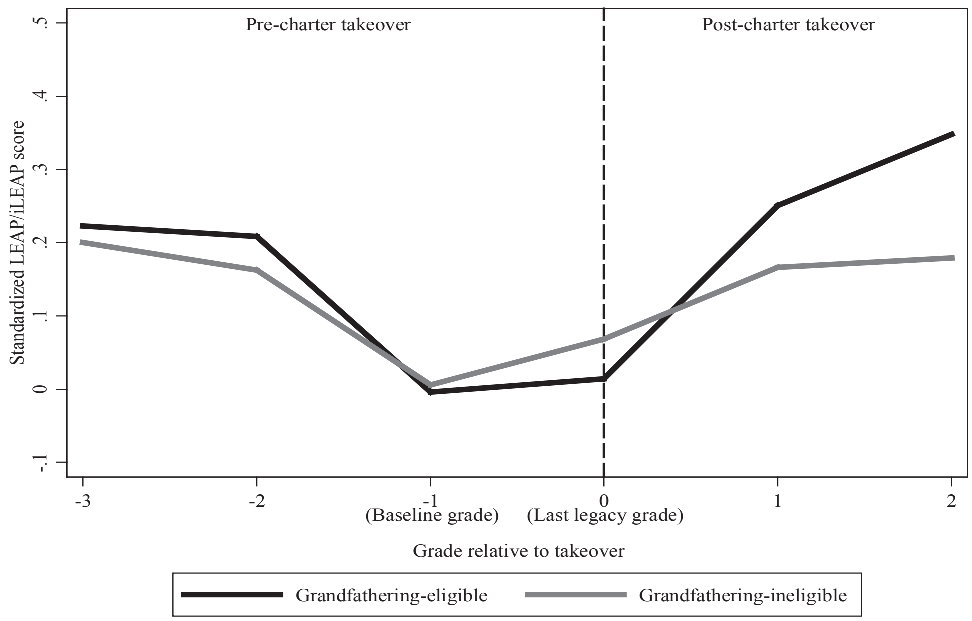
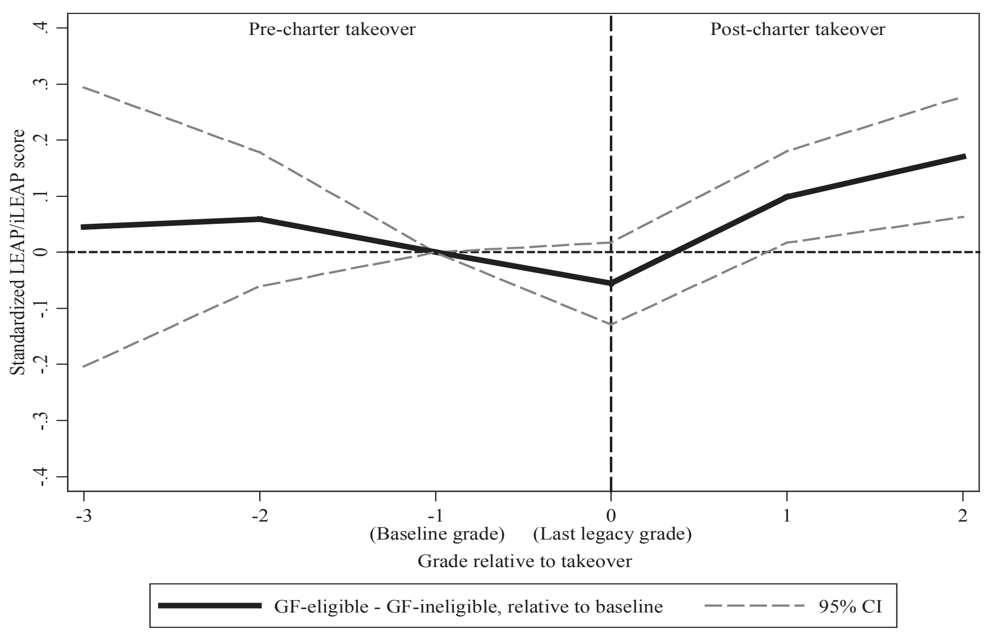


Figure 4b: Grandfathering DD for ELA



Notes: Figure 4a plots average LEAP/iLEAP ELA scores of students in the RSD legacy middle school matched sample. Figure 4b plots achievement growth relative to the baseline grade. Estimates in both figures control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in the set of direct-run schools in New Orleans RSD.

Figure 5a: Math scores for the UP grandfathering sample

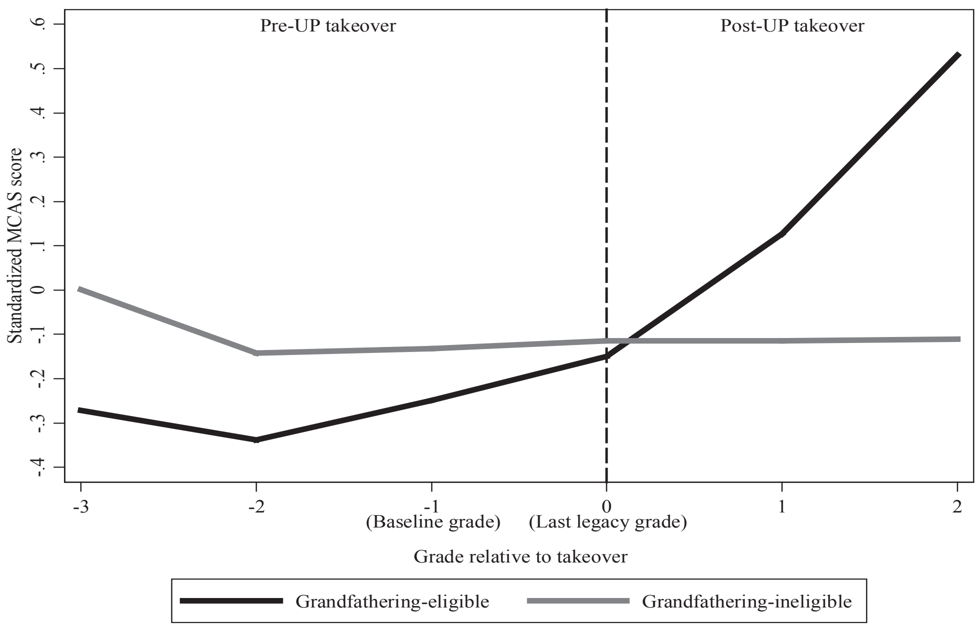
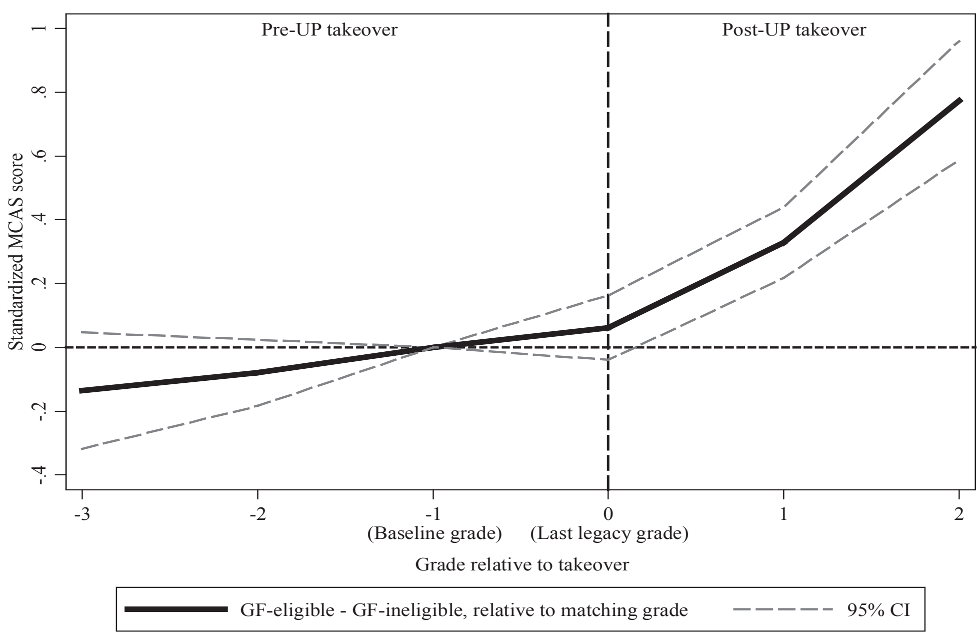


Figure 5b: Grandfathering DD for math



Notes: Figure 5a plots average MCAS math scores of students in the Gavin Middle School matched sample. Figure 5b plots achievement growth relative to the baseline grade. Estimates in both figures control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in BPS. For 7th grade legacy students, baseline grade scores are from 5th grade and last legacy grade scores are from 7th grade.

Figure 6a: ELA Scores for the UP grandfathering sample

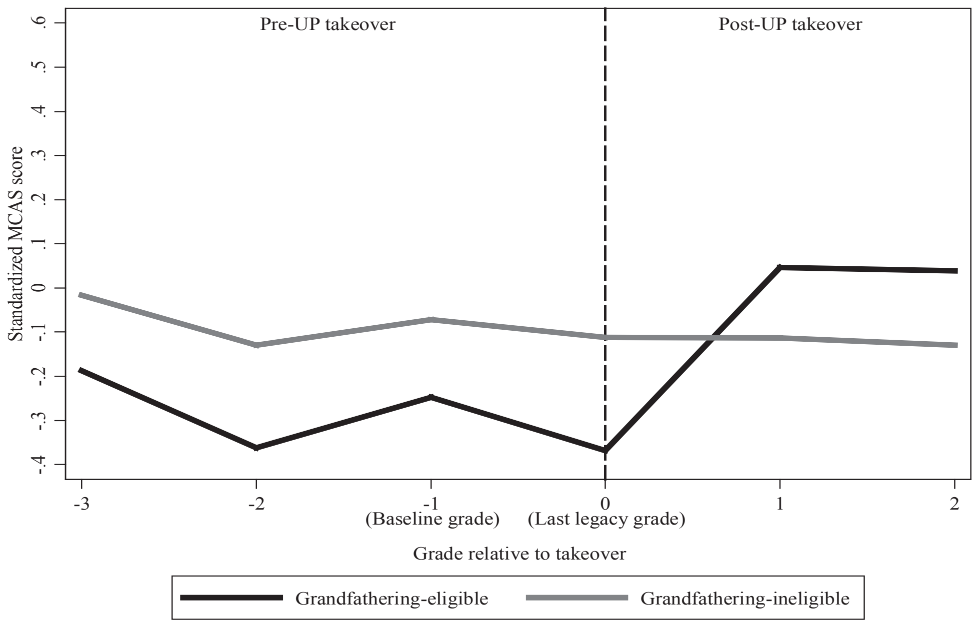
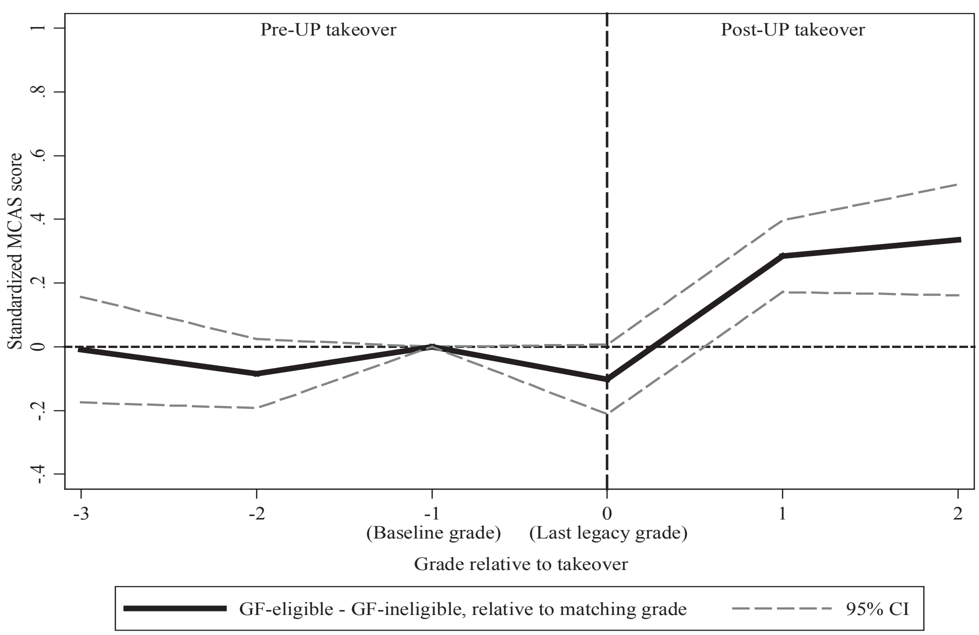


Figure 6b: Grandfathering DD for ELA



Notes: Figure 6a plots average MCAS ELA scores of students in the Gavin Middle School matched sample. Figure 6b plots achievement growth relative to the baseline grade. Estimates in both figures control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in BPS. For 7th grade legacy students, baseline grade scores are from 5th grade and last legacy grade scores are from 7th grade.

Table 1: Timing in the grandfathering research design

	RSD				Boston	
Baseline (matching) grade	3	4	5	6	5	5
Legacy enrollment grade	**4**	**5**	**6**	**7**	**6**	**7**
First takeover grade	5	6	7	8	7	8
Second takeover grade	6	7	8		8	
Third takeover grade	7	8				
Fourth takeover grade	8					
Legacy enrollment years (schools)	2009-10 (5) 2010-11 (1) 2011-12 (1) 2012-13 (4)				2010-11 (1)	

Notes: This table summarizes grade-based timing for matching, grandfathering eligibility, and takeover outcomes in the RSD and Boston analysis samples. Grandfathering eligibility is determined by enrollment in the fall of the legacy enrollment year, while matching uses information from the baseline grade. Outcomes are from the spring of the corresponding school year of each takeover grade. The number of grandfathered schools in each academic year are in parentheses.

Table 2: RSD descriptive statistics and grandfathering balance

	Sample means				Balance coefficients	
	RSD		Analysis sample			
	RSD students	Charter-bound RSD students	Takeover charter students	Grandfathering-eligible students	Analysis sample	First exposure year sample
	(1)	(2)	(3)	(4)	(5)	(6)
Hispanic	0.026	0.024	0.018	0.029	--	--
Black	0.964	0.971	0.994	0.982	--	--
White	0.019	0.017	0.008	0.016	--	--
Asian	0.008	0.008	0.001	0.009	--	--
Female	0.477	0.475	0.489	0.501	--	--
Special education	0.069	0.066	0.071	0.093	--	--
Free/reduced price lunch	0.913	0.926	0.955	0.919	--	--
Limited English proficient	0.017	0.016	0.013	0.020	0.000	-0.001
					(0.001)	(0.001)
N	14,554	11,358	1,040	763	3,503	2,572
Baseline math score	0.001	0.023	-0.320	-0.266	-0.019	-0.042
					(0.048)	(0.052)
N	12,945	10,557	1,038	760	3,500	2,570
Baseline ELA score	0.004	0.027	-0.303	-0.261	-0.009	-0.032
					(0.048)	(0.055)
N	12,947	10,557	1,040	762	3,502	2,572
Baseline science score	0.008	0.020	-0.218	-0.208	-0.039	-0.071
					(0.050)	(0.059)
N	12,910	10,534	1,033	758	3,487	2,562
Baseline social science score	0.003	0.022	-0.236	-0.219	-0.027	-0.016
					(0.047)	(0.056)
N	12,905	10,530	1,032	758	3,480	2,561

Notes: This table reports sample means and coefficients from regressions of the variable in each row on a grandfathering eligibility dummy indicating enrollment in an RSD takeover legacy school in the fall of the academic year prior to takeover. All regressions include matching cell fixed effects (cells are defined by race, sex, special education status, subsidized lunch eligibility, baseline grade and year, and baseline school SPS scores in five-point bins). The sample in columns 3-6 is restricted to students enrolled in an RSD direct-run school at baseline. Column 1 reports means for a sample of RSD students in the same baseline years as the analysis sample, while column 2 is restricted to those students that enroll in an RSD charter school in grades following the baseline. Column 3 reports means for students that enroll in a takeover charter in potential takeover grades, while column 4 describes students enrolled in a legacy school. Robust standard errors are reported in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 3: Grandfathering IV estimates of RSD takeover attendance effects

		Comparison group mean	OLS	2SLS estimates: First stage	2SLS estimates: Attendance effect
		(1)	(2)	(3)	(4)
All grades	Math	-0.089	0.123***	1.073***	0.212***
	(N: 5,625)		(0.020)	(0.052)	(0.038)
	ELA	-0.092	0.082***	1.075***	0.143***
	(N: 5,621)		(0.018)	(0.052)	(0.039)
			A. By grade		
5th & 6th grades	Math	-0.091	0.099***	0.738***	0.165**
	(N: 2,579)		(0.035)	(0.041)	(0.068)
	ELA	-0.116	0.023	0.745***	0.101
	(N: 2,579)		(0.033)	(0.042)	(0.070)
7th & 8th grades	Math	-0.086	0.133***	1.355***	0.231***
	(N: 3,046)		(0.020)	(0.070)	(0.037)
	ELA	-0.071	0.104***	1.352***	0.171***
	(N: 3,042)		(0.019)	(0.070)	(0.036)
			B. By potential exposure		
First exposure year	Math	-0.105	0.200***	0.659***	0.230***
(5th-8th grades)	(N: 2,553)		(0.044)	(0.023)	(0.069)
	ELA	-0.103	0.099**	0.659***	0.197***
	(N: 2,553)		(0.043)	(0.023)	(0.068)
Second exposure year	Math	-0.151	0.168***	1.148***	0.332***
(6th-8th grades)	(N: 1,664)		(0.031)	(0.061)	(0.058)
	ELA	-0.124	0.101***	1.158***	0.158***
	(N: 1,664)		(0.028)	(0.061)	(0.051)
Third & fourth exposure year	Math	0.015	0.097***	1.698***	0.117***
(7th & 8th grades)	(N: 1,408)		(0.022)	(0.131)	(0.042)
	ELA	-0.033	0.077***	1.698***	0.094**
	(N: 1,404)		(0.020)	(0.132)	(0.043)

Notes: This table reports OLS and 2SLS estimates of the effects of RSD takeover charter enrollment on 5th-8th grade LEAP/iLEAP math and ELA test scores using the grandfathering eligibility instrument. The sample in columns 2-4 includes RSD direct-run school students matched to a pre-takeover year legacy school student. The endogenous regressor counts the number of years enrolled at a takeover charter prior to testing. All models control for matching strata, limited English proficiency, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Means in column 1 are outcome grade scores for grandfathering-ineligible matched students.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 4: School choice in the RSD analysis sample

	All students		Compliers	
	Z=0	Z=1	Z=0	Z=1
	(1)	(2)	(3)	(4)
		A. First exposure year		
Enrolled in study takeover	0.089	0.780	--	1.000
… in other RSD takeover	0.072	0.032	0.087	--
… in non-takeover RSD charter	0.331	0.154	0.316	--
… in RSD direct-run	0.508	0.034	0.597	--
N	2,027	531		
		B. Second exposure year		
Enrolled in study takeover	0.206	0.714	--	1.000
… in other RSD takeover	0.105	0.016	0.142	--
… in non-takeover RSD charter	0.395	0.223	0.393	--
… in RSD direct-run	0.294	0.047	0.465	--
N	1,349	318		
		C. Third exposure year		
Enrolled in study takeover	0.277	0.754	--	1.000
… in other RSD takeover	0.112	0.031	0.218	--
… in non-takeover RSD charter	0.450	0.188	0.517	--
… in RSD direct-run	0.161	0.026	0.265	--
N	795	191		
		D. Fourth exposure year		
Enrolled in study takeover	0.316	0.646	--	1.000
… in other RSD takeover	0.167	0.051	0.030	--
… in non-takeover RSD charter	0.485	0.291	0.864	--
… in RSD direct-run	0.032	0.013	0.106	--
N	342	79		

Notes: This table describes school enrollment in the RSD analysis sample. Columns 1-2 characterize enrollment for grandfathering eligible (Z=1) and ineligible (Z=0) students, while columns 3-4 show the same for grandfathering compliers. Other RSD takeover charters include charter-to-charter conversions, principal-led conversions, and mergers. Non-takeover RSD charters include startup charters created since the 2008-2009 academic year, and charters operating in the RSD as of 2007-2008. Complier means are estimated by the method outlined in Abadie (2003), using a probit specification for E[Z|X] and the same controls as were used to construct the estimates in Table 3.

Table 5: Grandfathering IV estimates of RSD charter attendance effects

	(1)	(2)	(3)	(4)	(5)
			A. Math (N:5,625)		
Takeover charter	0.212***	0.217***	0.357***		
	(0.038)	(0.037)	(0.073)		
	[425.7]	[3323.0]	[632.7]		
Other RSD charter			0.324**		
			(0.152)		
			[47.1]		
Any RSD charter				0.385***	0.371***
				(0.071)	(0.059)
				[223.4]	[464.9]
Instruments	1	31	31	1	31
			B. ELA (N:5,621)		
Takeover charter	0.143***	0.154***	0.125		
	(0.039)	(0.037)	(0.081)		
	[424.7]	[3319.8]	[620.6]		
Other RSD charter			-0.069		
			(0.174)		
			[46.3]		
Any RSD charter				0.257***	0.214***
				(0.072)	(0.059)
				[228.0]	[486.5]
Instruments	1	31	31	1	31

Notes: This table reports 2SLS estimates of the effects of study takeover and other RSD charter enrollment on 5th-8th grade LEAP/iLEAP math and ELA test scores. The sample includes RSD direct-run school students matched to a pre-takeover year legacy school student as described in Table 2. The endogenous regressors count the number of years enrolled in RSD charters prior to testing. The instrument for columns 1 and 4 is grandfathering eligibility. For columns 2, 3, and 4, grandfathering eligibility was interacted with baseline year, special education status, and SPS bin cells. All models control for matching strata, limited English proficiency, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Angrist-Pischke multivariate first-stage F statistics are reported in brackets.

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

Table 6: UP descriptive statistics and grandfathering balance

	Sample means				Balance coefficients		
	Boston		Analysis sample		Analysis sample		First exposure year sample
	Boston students	Charter-bound Boston students	UP students	Grandfathering-eligible students	No score controls	Pre-baseline score controls	Pre-baseline score controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Hispanic	0.346	0.273	0.240	0.241	--	--	--
Black	0.407	0.519	0.511	0.469	--	--	--
White	0.135	0.152	0.124	0.152	--	--	--
Asian	0.072	0.023	0.089	0.100	--	--	--
Female	0.483	0.502	0.489	0.483	--	--	--
Special education	0.226	0.188	0.267	0.317	--	--	--
Free/reduced price lunch	0.804	0.751	0.951	0.928	--	--	--
Limited English proficient	0.233	0.130	0.342	0.307	0.045	0.034	0.026
					(0.032)	(0.033)	(0.033)
N	8,506	1,552	225	290	1,203	1,060	998
Baseline math score	0.008	0.173	-0.286	-0.253	-0.117*	-0.040	-0.032
					(0.070)	(0.053)	(0.055)
N	8,049	1,518	210	258	1,142	1,037	983
Baseline ELA score	0.012	0.175	-0.273	-0.235	-0.177***	-0.012	-0.014
					(0.065)	(0.053)	(0.054)
N	7,923	1,515	208	254	1,105	1,036	982

Notes: This table reports sample means and coefficients from regressions of the variable in each row on a grandfathering eligibility dummy indicating enrollment in Gavin Middle School in 6th or 7th grade in the fall of 2010. All regressions include matching cell fixed effects (cells are defined by race, sex, special education status, subsidized lunch eligibility, and 5th grade school and year). Regressions in columns 6 and 7 also control for 4th grade MCAS scores. The sample in columns 3-7 is restricted to students enrolled at a BPS school at baseline. Column 1 reports means for a sample of Boston students in the same baseline years as the analysis sample, while column 2 is restricted to students from the Boston sample who enroll in a Boston charter school in grades 6-8. Column 3 reports means for students in the analysis sample who enroll at UP in grades 7 and 8, while column 4 describes students enrolled at Gavin Middle School in the fall of 2010. Robust standard errors are reported in parentheses.

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

Table 7: Grandfathering IV estimates of UP attendance effects

		Comparison group mean	OLS	2SLS estimates: First stage	2SLS estimates: Attendance effect
		(1)	(2)	(3)	(4)
All grades	Math	-0.233	0.400***	1.051***	0.321***
	(N: 1,543)		(0.032)	(0.040)	(0.039)
	ELA	-0.214	0.296***	1.040***	0.394***
	(N: 1,539)		(0.035)	(0.041)	(0.044)
			A. By grade		
7th grade	Math	-0.247	0.399***	0.807***	0.317***
	(N: 531)		(0.063)	(0.036)	(0.068)
	ELA	-0.260	0.581***	0.803***	0.646***
	(N: 528)		(0.073)	(0.037)	(0.082)
8th grade	Math	-0.226	0.397***	1.196***	0.327***
	(N: 1,012)		(0.033)	(0.052)	(0.038)
	ELA	-0.190	0.244***	1.184***	0.287***
	(N: 1,011)		(0.037)	(0.052)	(0.044)
			B. By potential exposure		
First exposure year	Math	-0.214	0.365***	0.822***	0.325***
(7th & 8th grades)	(N: 1,028)		(0.047)	(0.025)	(0.048)
	ELA	-0.195	0.475***	0.809***	0.495***
	(N: 1,025)		(0.055)	(0.026)	(0.060)
Second exposure year	Math	-0.272	0.408***	1.541***	0.324***
(8th grade)	(N: 515)		(0.038)	(0.087)	(0.044)
	ELA	-0.252	0.221***	1.543***	0.271***
	(N: 514)		(0.042)	(0.087)	(0.049)

Notes: This table reports OLS and 2SLS estimates of the effects of UP enrollment on 7th and 8th grade MCAS math and ELA test scores using the grandfathering eligibility instrument. The sample in columns 2-4 includes BPS students matched to a 2010-11 6th or 7th grade Gavin Middle School student. The endogenous regressor counts the number of years enrolled at UP prior to testing. All models control for matching strata, limited English proficiency, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Means in column 1 are outcome grade scores for grandfathering-ineligible matched students.

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

Table 8: Lottery IV estimates of UP attendance effects

		Comparison group mean	OLS	2SLS estimates: First stage: Immediate offer	Waitlist offer	Attendance effect
		(1)	(2)	(3)	(4)	(5)
All grades	Math	0.059	0.301***	0.760***	0.562***	0.270***
	(N: 2,202)		(0.022)	(0.063)	(0.067)	(0.056)
	ELA	0.103	0.148***	0.759***	0.562***	0.118**
	(N: 2,205)		(0.020)	(0.063)	(0.067)	(0.051)
A. By grade						
6th grade	Math	0.064	0.334***	0.505***	0.407***	0.350***
	(N: 838)		(0.045)	(0.035)	(0.040)	(0.090)
	ELA	0.062	0.234***	0.505***	0.407***	0.214**
	(N: 839)		(0.045)	(0.035)	(0.040)	(0.091)
7th & 8th grades	Math	0.057	0.296***	0.919***	0.653***	0.248***
	(N: 1,364)		(0.021)	(0.085)	(0.088)	(0.054)
	ELA	0.126	0.133***	0.918***	0.653***	0.088*
	(N: 1,366)		(0.020)	(0.085)	(0.088)	(0.047)
B. By potential exposure						
First exposure year	Math	0.056	0.347***	0.519***	0.397***	0.365***
(6th & 7th grades)	(N: 881)		(0.044)	(0.034)	(0.038)	(0.086)
	ELA	0.058	0.239***	0.521***	0.394***	0.220**
	(N: 882)		(0.044)	(0.034)	(0.038)	(0.088)
Second & third exposure year	Math	0.061	0.294***	0.921***	0.665***	0.242***
(7th & 8th grades)	(N: 1,321)		(0.021)	(0.088)	(0.091)	(0.054)
	ELA	0.129	0.131***	0.918***	0.668***	0.083*
	(N: 1,323)		(0.020)	(0.088)	(0.091)	(0.047)

Notes: This table reports OLS and 2SLS estimates of the effects of UP enrollment on 6th-8th grade MCAS test scores using 6th and 7th grade lottery offer instruments. The sample in columns 2-4 includes Boston students entering 6th grade in the 2011-12 and 2012-13 academic years and 7th grade in the 2011-12 academic year with baseline demographic information. The endogenous regressor counts the number of years enrolled at UP prior to testing. The instruments are immediate and waitlist offer dummies. Immediate offer is equal to one when a student is offered a seat immediately following the lottery in March, while waitlist offer is equal to one for students offered seats later, up through the end of September. All models control for cohort dummies and student race, sex, special education status, limited English proficiency, subsidized lunch status, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Means in column 1 are outcome grade scores for applicants not given an immediate or waitlist offer.

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

Table 9: School choice in the UP analysis samples

	All students		Compliers	
	Z=0	Z=1	Z=0	Z=1
	(1)	(2)	(3)	(4)
	A. Grandfathering			
Enrolled in UP	0.012	0.794	--	1.000
…in other Boston charter	0.063	0.009	0.070	--
…in BPS	0.837	0.137	0.856	--
…in other Massachusetts	0.087	0.060	0.075	--
N	804	233		
	B. Lottery			
Enrolled in UP	0.040	0.514	--	1.000
…in other Boston charter	0.179	0.157	0.057	--
…in BPS	0.748	0.307	0.937	--
…in other Massachusetts	0.033	0.022	0.006	--
N	425	453		

Notes: This table describes school enrollment in the first exposure year for students in the UP grandfathering and lottery analysis samples. Columns 1-2 in panel A characterize enrollment for grandfathering-eligible (Z=1) and grandfathering-ineligible (Z=0) students, while columns 3-4 show the same for grandfathering compliers. Columns 1-2 in panel B characterize enrollment for "ever offered" (Z=1) and not offered (Z=0) students, while columns 3-4 show the same for lottery offer compliers. Ever offered lottery applicants are those who received either an immediate or a waitlist offer. Complier means in panels A and B are estimated by the method outlined in Abadie (2003), using a probit specification for E[Z|X] and the same controls as were used to construct the estimates in Tables 7 and 8.

Table 10: UP grandfathering and lottery comparison

	Treatment Effect		Treated outcome			Untreated outcome		
	GF	Lottery	GF G_1	GF Y_1	Lottery Y_1	GF G_0	GF Y_0	Lottery Y_0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	0.325***	0.365***	0.347***	0.176***	0.423***	0.021	-0.149***	0.057
	(0.048)	(0.086)	(0.037)	(0.050)	(0.050)	(0.033)	(0.045)	(0.077)
N	1,028	881						
GF-Lottery gap		-0.040			-0.247***			-0.207**
		(0.099)			(0.070)			(0.089)
ELA	0.495***	0.220**	0.494***	0.343***	0.293***	-0.001	-0.153***	0.073
	(0.060)	(0.088)	(0.046)	(0.057)	(0.053)	(0.042)	(0.049)	(0.076)
N	1,025	882						
GF-Lottery gap		0.275***			0.050			-0.225**
		(0.106)			(0.077)			(0.090)

Notes: This table reports 2SLS estimates of LATEs and average potential outcomes for grandfathering and lottery compliers in the UP analysis samples. The estimates in columns 1 and 2 match the first exposure year results in Tables 7 and 8. Complier treatment effects and average potential outcomes reported in columns 1, 2, 3, 5, 6, and 8 are estimated by 2SLS. Column 3 shows the estimated $E[Y^g_{01} - Y^l_0 \mid D_1 > D_0]$ for grandfathering compliers, while column 4 adds an estimate of $E[Y^l_0 \mid D_0=0]$ (-0.17 for math and -0.15 for ELA) to generate an estimate of $E[Y^g_{01} \mid D_1 > D_0]$. Similarly, column 6 reports an estimate of $E[Y^g_{00} - Y^l_0 \mid D_1 > D_0]$ for grandfathering compliers, and column 7 adds the estimate of $E[Y^l_0 \mid D_0=0]$ to generate an estimate of $E[Y^g_{00} \mid D_1 > D_0]$. Robust standard errors, clustered by student, are reported in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

References

Abadie, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variables Models," *Journal of the American Statistical Association*, 97.

——— (2003): "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 113.

Abdulkadiroğlu, A., J. Angrist, S. Dynarski, T. J. Kane, and P. Pathak (2011): "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots," *Quarterly Journal of Economics*, 126, 699–748.

Abdulkadiroğlu, A., J. Angrist, and P. Pathak (2014a): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 82(1), 137–196.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2014b): "Designing the New Orleans School Match: TTC vs. DA," Working Paper, MIT.

Angrist, J. D. (1998): "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 66(2), 249–288.

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

Angrist, J. D., G. Imbens, and D. Rubin (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91(434), 444–455.

Angrist, J. D. and G. W. Imbens (1995): "Two-stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of the American Statistical Association*, 90, 431–442.

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

Ash, K. (2013): "Charters Adopt Common Application Systems." Available at `http://www.edweek.org/ew/articles/2013/09/25/05charter_ep.h33.html`. Last accessed: December 17, 2014.

Ashenfelter, O. (1978): "Estimating the effect of training programs on earnings," *Review of Economics and Statistics*, 60, 47–57.

Chamberlain, G. (1984): "Panel data," in *Handbook of Econometrics*, ed. by Z. Griliches and M. D. Intriligator, Elsevier, vol. 2, chap. 22, 1247–1318, 1 ed.

Cowen (2011a): "Public School Funding in Louisiana," Available at `http://www.coweninstitute.com/wp-content/uploads/2010/03/SPELA-Funding-and-ARRA.pdf`. Last Accessed November 11, 2014.

——— (2011b): "The State of Public Education in New Orleans: School Finances," Available at `http://www.coweninstitute.com/wp-content/uploads/2011/04/SPENO-Finances-Appendix-Final-5April11.pdf`. Last Accessed November 11, 2014.

——— (2011c): "Transforming Public Education in New Orleans: The Recovery School District 2003-2011," Available at `http://www.coweninstitute.com/wp-content/uploads/2011/12/History-of-the-RSD-Report-2011.pdf`. Last Accessed: October 10, 2014.

CREDO (2013a): "National Charter School Study," Center for Research on Education Outcomes. Available at `https://credo.stanford.edu/documents/NCSS%202013%20Final%20Draft.pdf`. Last Accessed: October 10, 2014.

——— (2013b): "New Schools for New Orleans, Year 2 Report," Center for Research on Education Outcomes. Available at `http://credo.stanford.edu/pdfs/NSNOYear2Report.pdf`. Last Accessed: October 10, 2014.

Darling-Hammond, L. (2012): "Darling-Hammond: Why is Congress redlining our schools?" Available at `http://www.thenation.com/article/165575/why-congress-redlining-our-schools`. Last accessed: December 17, 2014.

Dee, T. (2012): "School Turnarounds: Evidence from the 2009 Stimulus," NBER Working Paper, No. 17990.

Deryugina, T., L. Kawano, and S. Levitt (2014): "The Economic Impact of Hurricane Katrina on its Victims: Evidence from Individual Tax Returns," NBER Working Paper No. 20713.

Dobbie, W. and R. G. Fryer (2011): "Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 3, 158–187.

——— (2013): "Getting Beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 5, 58–75.

Drellinger, D. (2014): "U.S. Education Department Opens Civil Rights Investigation of New Orleans Public School Closures," The Times-Picayune, September 24. Available at http://www.nola.com/education/index.ssf/2014/09/us_education_department_opens.html. Last Accessed: October 10, 2014.

Duncan, A. (2010): "Secretary Duncan Message to School Boards." Available at https://www.youtube.com/watch?v=DlmyBh3dTUYl. Last accessed: November 30, 2014.

Epple, D., A. Jha, and H. Sieg (2013): "The Superintendent's Dilemma: Managing School District Capacity as Parents Vote with Their Feet," Working Paper. Available at http://web.stanford.edu/~akshayaj/school_choice_20_2_13.pdf, Last Accessed: October 10, 2014.

Fryer, R. G. (2014): "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments," *Quarterly Journal of Economics*, 129, 1355–1407.

Groen, J. A. and A. E. Polivka (2008): "The Effect of Hurricane Katrina on the Labor Market Outcomes of Evacuees," *The American Economic Review*, 98, 43–48.

Hoxby, C. M., S. Murarka, and J. Kang (2009): "How New York City's charter schools affect achievement," New York City Charter Schools Evaluation Project, Cambridge, MA.

Hull, P. D. (2014): "IsoLATEing: Identifying Heterogeneous Effects of Multiple Treatments," Working Paper. Available at http://www.mit.edu/~hull/isoLATE_12162014.pdf.

Imbens, G. W. and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–475.

Imberman, S. A., A. D. Kugler, and B. I. Sacerdote (2012): "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees," *The American Economic Review*, 102, 2048–2082.

Knight, R. (2013): Interview at UP Academy Boston, November.

NACPS (2013a): "Back to School Tallies: Estimated Number of Public Charter Schools and Students, 2012-2013." Available at http://www.publiccharters.org/wp-content/uploads/2014/01/NAPCS-2012-13-New-and-Closed-Charter-Schools_20130114T161322.pdf. Last accessed: October 10, 2014.

——— (2013b): "A Growing Movement: America's Largest Charter School Communities." Available at http://www.publiccharters.org/wp-content/uploads/2013/12/Market-Share-Report-2013.pdf. Last accessed: October 10, 2014.

NCES (2013): "Digest of Education Statistics, Table 216.20: Number and enrollment of public elementary and secondary schools, by school level, type, and charter and magnet status: Selected years, 1990-91 through 2011-12." Available at http://nces.ed.gov/programs/digest/d13/tables/dt13_216.20.asp. Last accessed: December 17, 2014.

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

Rothstein, R. (2004): *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*, New York: Teachers College Press.

——— (2011): "Grading the Education Reformers," Slate. Available at http://www.slate.com/articles/arts/books/2011/08/grading_the_education_reformers.html. Last Accessed: November 30, 2014.

Sacerdote, B. (2012): "When the Saints go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita," *American Economic Journal: Applied Economics*, 4, 109–135.

Toness, B. V. (2010): "Charter School Will Do Things the Hard Way," 90.9WBUR. Available at http://www.wbur.org/2010/12/09/charter-challenges. Last accessed: October 17, 2014.

Tuttle, C. C., B. Gill, P. Gleason, V. Knechtel, I. Nichols-Barrer, and A. Resch (2013): "KIPP Middle Schools: Impacts on Achievement and Other Outcomes," Princeton: Mathematica Policy Research.

UP Academy (2010): "Application for a Massachusetts Horace Mann III Public Charter School: UP Academy Charter School of Boston," Submitted November 8, 2010. Available at http://www.doe.mass.edu/charter/finalists/10/UPApp.doc. Last accessed April 16, 2014.

USDOE (2009): "Guidance on School Improvement Grants Under Section 1003(g) of the Elementary and Secondary Education Act of 1965." Available at http://www2.ed.gov/programs/sif/guidance-20091218.doc. Last Accessed: October 10, 2014.

Vaznis, J. (2014): “Education chief backs probation for charter school,” The Boston Globe, October 15. Available at http://www.bostonglobe.com/metro/2014/10/15/state-education-chief-backs-probation-for-brighton-indistrict-charter-school/b7fCiu9xcBVJjoMR8oGIhK/story.html. Last Accessed: November 9, 2014.

Vigdor, J. L. (2007): “The Katrina Effect: Was There a Bright Side to the Evacuation of Greater New Orleans?” *The BE Journal of Economic Analysis & Policy*, 7.

Walters, C. R. (2014): “The Demand for Effective Charter Schools,” UC-Berkeley Mimeo.

A Econometric Appendix

Theorem 1. Under Assumptions 1-4,

$$\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} = E[Y^g_{11} - Y^g_{10}|D_1 > D_0] = E[Y^g_{01} - Y^g_{00}|D_1 > D_0].$$

Proof: Using monotonicity (Assumption 2) to partition the $Z = 1$ and $Z = 0$ populations into second-period subpopulations of always-takers, never-takers, and compliers, we have:

$$\begin{aligned}
E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0] =& E[Y^g - Y^l|D_0 = 1, Z=1]P(D_0 = 1|Z=1) \\
&+ E[Y^g - Y^l|D_1 = 0, Z=1]P(D_1 = 0|Z=1) \\
&+ E[Y^g - Y^l|D_1 > D_0, Z=1]P(D_1 > D_0|Z=1) \\
&- E[Y^g - Y^l|D_0 = 1, Z=0]P(D_0 = 1|Z=0) \\
&- E[Y^g - Y^l|D_1 = 0, Z=0]P(D_1 = 0|Z=0) \\
&- E[Y^g - Y^l|D_1 > D_0, Z=0]P(D_1 > D_0|Z=0), \\
=& \left(E[Y^g_{11} - Y^l_1|D_0 = 1] - E[Y^g_{01} - Y^1_0|D_0 = 1]\right) P(D_0 = 1) \qquad (10) \\
&+ \left(E[Y^g_{10} - Y^l_1|D_1 = 0] - E[Y^g_{00} - Y^l_0|D_1 = 0]\right) P(D_1 = 0) \\
&+ \left(E[Y^g_{11} - Y^l_1|D_1 > D_0] - E[Y^g_{00} - Y^l_0|D_1 > D_0]\right) P(D_1 > D_0),
\end{aligned}$$

where the second equality follows from independence (Assumption 1).

As a consequence of Assumption 4, we have:

$$E[Y^g_{11} - Y^l_1|D_0 = 1] = E[Y^g_{01} - Y^l_0|D_0 = 1] \qquad (11)$$

$$E[Y^g_{11} - Y^l_1|D_1 > D_0] = E[Y^g_{01} - Y^l_0|D_1 > D_0] \qquad (12)$$

and that

$$E[Y^g_{00} - Y^l_0|D_1 = 0] = E[Y^g_{10} - Y^l_1|D_1 = 0] \qquad (13)$$

$$E[Y^g_{00} - Y^l_0|D_1 > D_0] = E[Y^g_{10} - Y^l_1|D_1 > D_0] \qquad (14)$$

Equations (11) and (13) imply that the first two terms in equation (10) are zero. Equation (14) and the fact that, by independence and monotonicity, $E[D|Z = 1] - E[D|Z = 0] = P(D_1 > D_0)$ imply further that

$$\begin{aligned}
\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} &= E[Y^g_{11} - Y^l_1|D_1 > D_0] - E[Y^g_{10} - Y^l_1|D_1 > D_0] \\
&= E[Y^g_{11} - Y^g_{10}|D_1 > D_0].
\end{aligned}$$

From (12), it follows similarly that

$$\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} = E[Y^g_{01} - Y^l_0|D_1 > D_0] - E[Y^g_{00} - Y^l_0|D_1 > D_0]$$
$$= E[Y^g_{01} - Y^g_{00}|D_1 > D_0]$$

□

We assume throughout that Assumptions 1-4 hold conditional on a set of mutually-exclusive and exhaustive matching cell dummies, d_j. These covariates add a layer of cross-cell averaging to the within-cell average-causal-effects interpretation of the 2SLS estimand. With matching-cell fixed effects as the only controls, the covariate parameterization is saturated. Therefore, as shown by Abadie (2003), a 2SLS regression of $Y^g - Y^l$ on D and $\{d_j\}$ that instruments D by Z identifies the treatment coefficient in a regression of $Y^g - Y^l$ on $\{d_j\}$ and D for compliers (this follows from the linearity of the propensity score in a saturated model). Moreover, Angrist (1998) shows that, in general, a regression of this sort (saturated controls with a single additive treatment effect) generates a weighted-average of cell-specific treatment-control comparisons, with weights proportional to the within-cell variance of treatment. Thus, the IV estimand for this specification can be written as a variance-weighted average of cell-specific LATEs.

In practice, the grandfathering estimates reported here come from models that include additive controls for baseline covariates and year-of-test controls, as well as a full set of matching-cell fixed effects. When the additional controls are independent of Z within cells, the weighted average interpretation of a IV estimand with fully interacted controls is unchanged, while we can expect estimates of models that include additional controls to be more precise.

Extension of Theorem 1 to an Ordered Treatment

Suppose treatment, D, takes on values in the set $\{0, 1, ..., \bar{d}\}$. Assumption 1 is modified to accommodate this ordered treatment, below:

Assumption 1′ *(Independence)* $\{Y^l_0, Y^l_1, Y^g_{00}, ..., Y^g_{0\bar{d}}, Y^g_{10}, ..., Y^g_{1\bar{d}}, D_0, D_1\} \perp\!\!\!\perp Z$.

We similarly modify Assumption 4:

Assumption 4′ *(Strong Gains Exclusion)* $P(Y^g_{1d} - Y^l_1 = Y^g_{0d} - Y^l_0) = 1$ for $d \in \{0, 1, ..., \bar{d}\}$.

Under Assumptions 1′, 2, 3, and 4′ we can use Theorem 1 in Angrist and Imbens (1995) to show

$$\begin{aligned}\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} &= \sum_{d=1}^{\bar{d}} \frac{E[(Y^g_{1d} - Y^l_1) - (Y^g_{1d-1} - Y^l_1)|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)} \\ &= \sum_{d=1}^{\bar{d}} \frac{E[Y^g_{1d} - Y^g_{1d-1}|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}.\end{aligned}$$

Likewise,

$$\begin{aligned}\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} &= \sum_{d=1}^{\bar{d}} \frac{E[(Y^g_{0d} - Y^l_0) - (Y^g_{0d-1} - Y^l_0)|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)} \\ &= \sum_{d=1}^{\bar{d}} \frac{E[Y^g_{0d} - Y^g_{0d-1}|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}.\end{aligned}$$

As for our interpretation of the Bernoulli treatment estimand, the assumptions behind this interpretation of the ordered estimand are assumed to hold within matching cells, and our IV estimates of ordered treatment effects come from models that include a full set of matching-cell fixed effects. These models also include a set of additive controls that should be unrelated to the instruments conditional on matching controls. Angrist and Imbens (1995) show that the IV estimand in models with an ordered treatment, saturated covariate controls, and a saturated first stage (that is the first stage interacts Z with $\{d_j\}$) can be written as an average causal effect of a one-unit increase in treatment intensity for ordered-treatment compliers. In practice, we omit interactions Z with $\{d_j\}$ from the first stage, except where these are required to identify models with multiple endogenous regressors. As can be seen by comparing the estimates in columns 1 and 2 of Table 5, the omission of higher-order terms from the instrument list is of little empirical consequence for models with a single treatment effect.

Weakening Assumption 4

We modify the potential outcomes model described by equations (1) and (2) to allow legacy enrollment to change legacy-year and later potential outcomes to differing degrees. To keep the presentation compact, we assume away variation in conditional mean functions across compliant subpopulations. Specifically, we assume here that

$$E[Y^l_z|X,Z] = E[Y^l_z|X] = \alpha_1(X) + z\gamma(X) \tag{15}$$

$$E[Y^g_{zd}|X,Z] = E[Y^g_{zd}|X] = \alpha_2(X) + \lambda z\gamma(X) + d\beta, \tag{16}$$

where λ is a parameter assumed to lie in the unit interval. Covariates are required to identify the model when λ is unknown, so (15) and (16) are specified with additive covariate effects included (for the purposes of this discussion, matching cell dummies are subsumed in the covariate vector, X). As in Chamberlain's (1984) panel models with unobserved individual effects scaled by an unknown time-varying factor loading, models with $\lambda \neq 1$ are identified by quasi-differencing, then moving the lagged score term (legacy-grade scores, in this case) to the right hand side and treating this term as endogenous. The extra instruments needed to identify this model come from interacting Z with X. This leads us to omit covariate interactions from the specification of the treatment effect.[27]

Theorem 2. Suppose Assumptions 1-3 hold for a Bernoulli covariate, x, and that the conditional mean functions for potential outcomes satisfy (15) and (16) when X is replaced with x. Suppose also that the first stage varies with x, so that $P(D_1 > D_0|x = 0) \neq P(D_1 > D_0|x = 1)$. Then the IV estimand for a regression of Y^g on the pair (Y^l, D), treated as endogenous and instrumented with (Z, Zx), while controlling for exogenous x, identifies the parameters λ and β in equation (16).

Proof. Note that $Y^l = Y^l_Z$ and $Y^g = Y^g_{ZD}$. Equations (15) and (16) therefore imply

$$\begin{aligned} E[Y^l|Z,x] &= \alpha_1(x) + \gamma(x)Z \\ E[Y^g|Z,x] &= \alpha_2(x) + \lambda\gamma(x)Z + \beta E[D|Z,x] \end{aligned}$$

for $j = 0, 1$. From here, we get the reduced form

$$E[Y^g|Z,x] = [\alpha_2(x) - \lambda\alpha_1(x)] + \lambda E[Y^l|Z,x] + \beta E[D|Z,x]. \tag{17}$$

This completes the proof since (17) is the reduced form for the 2SLS procedure described in the theorem. □

In practice, our 2SLS models for RSD use 30 interactions of the grandfathering instrument with baseline year, SPED status, and SPS bins instead of the single interaction used to establish the theorem (for UP, we have over 60 interactions with baseline year, SPED status, and baseline school). Assuming away treatment effect variation with X, the over-identified model should produce a more efficient estimate of the parameters than the just-identified estimand.

Appendix Table A7 reports key parameter estimates from the setup described in Theorem 2. Consistent with differences-in-differences Figures 3 and 4, legacy year effects in RSD are estimated at about $-.09$ for math and $-.03$ for ELA. The contrast between 2SLS estimates for RSD allowing λ to be a free parameter identified by Theorem 2 and estimates under gains exclusion ($\lambda = 1$) appears in columns 2 and 3 of the table. Although λ is estimated to be about a half, the differences in treatment effects

[27] See Hull (2014) for a more general approach to identifying treatment effects under parameterized violations of the exclusion restriction.

here are modest and not significantly different from zero. As suggested by differences-in-differences Figures 5 and 6, the legacy year treatment effect for those grandfathered into UP is positive for math and negative for ELA. These estimates are reported in columns 5 and 6 of Table A7. In this case, λ is estimated to be about .63 for math and about .4 for ELA. Allowing λ to be free changes the math takeover effect little, but generates a marginally significant decline in the estimated ELA effect, from .4 to about .31. This decline leaves the comparison of lottery and grandfathering estimates qualitatively unchanged.

Estimation of Average Potential Outcomes

Our estimates of potential outcome means for lottery compliers come from an application of theoretical results in in Abadie (2002). To extend this to grandfathering instruments, let $W^1 \equiv D(Y^g - Y^l)$ and expand the within-stratum reduced form regression of W on Z as follows:

$$
\begin{aligned}
E[W^1|Z=1] - E[W^1|Z=0] =& E[D(Y^g - Y^l)|D_0 = 1, Z = 1]P(D_0 = 1|Z = 1) \\
&+ E[D(Y^g - Y^l)|D_1 = 0, Z = 1]P(D_1 = 0|Z = 1) \\
&+ E[D(Y^g - Y^l)|D_1 > D_0, Z = 1]P(D_1 > D_0|Z = 1) \\
&- E[D(Y^g - Y^l)|D_0 = 1, Z = 0]P(D_0 = 1|Z = 0) \\
&- E[D(Y^g - Y^l)|D_1 = 0, Z = 0]P(D_1 = 0|Z = 0) \\
&- E[D(Y^g - Y^l)|D_1 > D_0, Z = 0]P(D_1 > D_0|Z = 0), \\
=& \left(E[Y^g_{11} - Y^l_1|D_0 = 1] - E[Y^g_{01} - Y^1_0|D_0 = 1]\right) P(D_0 = 1) \\
&+ E[Y^g_{11} - Y^l_1|D_1 > D_0]P(D_1 > D_0) \\
=& E[Y^g_{11} - Y^l_1|D_1 > D_0]P(D_1 > D_0) \\
=& E[Y^g_{01} - Y^l_0|D_1 > D_0]P(D_1 > D_0).
\end{aligned}
$$

The first equality in this expansion follows by monotonicity; the second by independence; the third and fourth by gains exclusion (Assumption 4). As always, the first stage regression of D on Z identifies $P(D_1 > D_0)$. We've therefore shown that the IV estimand that puts W^1 on the left hand side and instruments D with Z estimates $E[Y^g_{01} - Y^l_0|D_1 > D_0]$. Likewise, swapping W^1 with $W^0 \equiv (1 - D)(Y^g - Y^l)$ in this 2SLS estimand recovers $E[Y^g_{00} - Y^l_0|D_1 > D_0]$ when $1 - D$ is taken as the endogenous variable. It remains, however, to construct an estimand for the complier mean of Y^g_{01} and of Y^g_{00}.

To identify $E[Y^g_{01}|D_1 > D_0]$ and $E[Y^g_{00}|D_1 > D_0]$ from $E[Y^g_{01} - Y^l_0|D_1 > D_0]$ and $E[Y^g_{00} - Y^l_0|D_1 > D_0]$, we need to know $E[Y^l_0|D_1 > D_0]$. Unfortunately, Y^l_0 is a potential outcome indexed against Z, that is, legacy enrollment, while compliance is a condition determined by takeover enrollment. This rules out identification of the average Y_0^l for compliers based on exclusion of Z. We therefore assume instead that the mean of Y_0^l is the same for grandfathering compliers and for never-takers:

Assumption 5 $E[Y_0^l|D_1 > D_0] = E[Y_0^l|D_1 = D_0 = 0]$.

Because never-takers are rare relative to compliers (compare the first stage of around .69 in Table 5 with the proportion of never takers, $1 - E[D|Z = 1] = P[D_1 = 0] = .22$]), modest violations of Assumption 5 should be of little empirical consequence.

Using Assumption 5 to construct the missing mean of Y_0^l for compliers, we have,

$$\begin{aligned} E[Y^l|D = 0, Z = 0] =& E[Y_0^l|D_0 = 0] \\ =& E[Y_0^l|D_1 > D_0], \end{aligned}$$

where the first equality follows by Independence and the second from the fact that those with $D_0 = 0$ are either never-takers or compliers and from Assumption 5. We can therefore estimate $E[Y_0^l|D_1 > D_0]$ by an estimate of $E[Y^l|D = 0, Z = 0]$. Adding this to $E[Y_{01}^g - Y_0^l|D_1 > D_0]$ and $E[Y_{00}^g - Y_0^l|D_1 > D_0]$ produces the desired mean outcomes.

B Data Appendix

B.1 New Orleans RSD

The New Orleans RSD grandfathering analysis file is constructed from student enrollment, demographic, and outcome data provided by RSD for school years 2007-2008 through 2013-2014. Enrollment and demographic data include information on all students enrolled in the New Orleans RSD. 4th and 8th grade test score outcomes are from the Louisiana Educational Assessment Program (LEAP); all other outcome grades measure achievement by the integrated LEAP (iLEAP) exam.

Student enrollment and demographics

RSD enrollment data include a June (end-of-year) file for school years 2007-2008 through 2012-2013, and an October file for school years 2011-2012 through 2013-2014. For school year 2013-2014, an additional February file is available. Each enrollment file is a snapshot of the enrollment records at each New Orleans RSD school, and contains information on the first and last dates of attendance of each student in each school and grade over the academic year up to the given month. Enrollment files also include a unique student identifier, the "student ID." We use information on student names and dates of birth to confirm consistency of student IDs across enrollment files. After resolving any inconsistencies, student IDs are used to match students to LEAP/iLEAP test score files.

Enrollment files contain information on student sex, race, special education status, limited English proficiency status, subsidized lunch eligibility, and school attended.[28] We construct a panel dataset capturing demographic and enrollment information for every student in each grade, keeping information from the first calendar year spent in each grade and recording the number of times a grade is subsequently repeated. A student is counted as attending a takeover charter if she enrolls in one for any amount of time. If a student attends multiple non-takeover charter schools within the same year and grade, she is counted as enrolled in the longest-attended school. Attendance duration ties are broken by giving preference to the most recent enrollment. All other attendance ties are broken randomly. Students classified as special education, limited English proficient, or eligible for a free- or reduced-price lunch in any record within a grade retain that designation for that grade.

[28]Race is coded as black, white, asian, hispanic, and other. In RSD these are not mutually-exclusive categories. Race and gender are grade-invariant characteristics, while SPED, LEP, and free/reduced price lunch status are grade-specific.

LEAP/iLEAP

The LEAP/iLEAP outcomes of interest are math, English Language Arts (ELA), sciences, and social sciences test scores for grades 5 through 8. Tests at the end of the legacy grade and at the end of the baseline matching grade are also used to calculate test score gains and as controls, respectively. The grade configuration of legacy and baseline tests is summarized in Table 1.

Each observation in the LEAP/iLEAP data files corresponds to a student's test results in a particular subject, grade, and year. For each grade, we use scores from the first attempt at a given subject test. The raw test score variables are standardized to have mean zero and standard deviation one within a subject-grade-year in the New Orleans RSD. The standardization excludes scores from students enrolled in alternative schools.

School Performance Scores

School performance scores ("SPS") for school years 2007-2008 through 2011-2012 are used to construct matching cells for New Orleans RSD grandfathered students. These scores are obtained from `http://www.louisianabelieves.com/resources/library/performance-scores`.

During the analysis sample years, SPS scores ranged from 0 to 200, where a score below 75 corresponded to an "F" letter grade, and a score above 120 corresponded to an "A" letter grade. Matched cells are constructed by splitting the 0 to 200 range into 5-point bins.

Grandfathering eligibility and matching

The grandfathering eligibility instrument is based on fall enrollment. For the New Orleans RSD, the instrument is coded by the enrollment designation procedure described above, using attendance data up to October 31. Students that leave a legacy school prior to October 31 are not considered grandfathering-eligible.

For each legacy school, grandfathering-eligible students are those enrolled in grades 4-7 in the fall immediately prior to takeover. These students are matched to grandfathering-ineligible students that share the sex, race, special education status, subsidized lunch eligibility, and SPS 5-point bin of the grandfathered student in their baseline grade and year. Students who are eligible for grandfathering into a takeover charter in multiple grades or who are matched to such a student are retained in the first grade by which they enter the analysis sample.

Appendix Table A3 describes the construction of the RSD grandfathering sample. We identify a total of 1,657 grandfathering-eligible students across all legacy grades of the 11 schools in our analysis. Excluding students without baseline information and those not enrolled in a direct-run school at baseline reduces this sample to 1,019. Matching these students and retaining the first grade observation produces our analysis sample of 763 grandfathering-eligible and 2,410 grandfathering-ineligible students.

B.2 Boston

Our analysis of the UP Charter School of Boston uses student enrollment, demographic, and outcome data provided by the Massachusetts Department of Elementary and Secondary Education for school years 2007-2008 through 2013-2014. We construct two analysis files, one for the UP grandfathering analysis and one for the UP lottery analysis. Boston enrollment and demographic data come from the Student Information Management System (SIMS), a centralized database that covers all public school students in Massachusetts. Test score outcomes are from the Massachusetts Comprehensive Assessment System (MCAS). For the lottery sample, lists of first-time applicants and lottery winners are provided by UP.

Student enrollment and demographics

SIMS data include an end-of-year file and an October file for each school year. As with RSD, each observation in the SIMS refers to a student in a grade of a school in a year. While length of attendance is recorded, the data do not contain exact dates of enrollment. The SIMS also includes a unique student identifier, known as the SASID, which is used to match students to MCAS test score files.

SIMS variables used in our analysis include student sex, race, special education status, limited English proficiency status, subsidized lunch eligibility, and school attended.[29] We construct a panel dataset capturing demographic and enrollment information for every Massachusetts public school student enrolled in each grade, keeping information from the first calendar year spent in each grade and recording the number of repeated attempts. A student is counted as attending UP if she enrolls for any amount of time. If a student attends multiple schools within the same year and grade, she is counted as enrolled in the longest-attended school. All other attendance ties are broken randomly. Students classified as special education, limited English proficient, or eligible for a free- or reduced-price lunch in any record within a grade retain that designation.

MCAS

The MCAS outcomes of interest are math and ELA test scores in 7th and 8th grade for the UP grandfathering analysis sample, and in grades 6-8 for the UP lottery sample. For the grandfathering

[29] Race is coded as black, white, asian, hispanic and other, and is mutually-exclusive.

analysis, tests at the end of the legacy grade and at the end of the baseline matching grade are also used to calculate test score gains and as controls, respectively. The grade configuration of legacy and baseline tests is summarized in Table 1. For the lottery analysis, baseline tests (from 5th or 6th grade, depending on the application grade) are used as controls.

Each observation in the MCAS data files corresponds to a student's test results in a particular subject, grade, and year. For each grade, we use scores from the first attempt at a given subject test. The raw test score variables are standardized to have mean zero and standard deviation one within a subject-grade-year in Boston. The standardization excludes scores from students enrolled in alternative schools.

Grandfathering eligibility and matching

The grandfathering instrument is based on fall semester enrollment. For UP the instrument is coded by the enrollment designation procedure described above using data from the October 2010 SIMS. Grandfathering-eligible students are those enrolled in 6th and 7th grade at Gavin middle school, and are matched to grandfathering-ineligible students that share the sex, race, special education status, subsidized lunch eligibility, and 5th grade school and year.

We identify a total of 334 students eligible for grandfathering into UP (Appendix Table A3). Excluding students without baseline information and those not enrolled in BPS at baseline reduces this sample to 290. These grandfathering-eligible students and their 913 ineligible matches constitutes our Boston grandfathering analysis sample.

Lottery sample

We obtained a list of students applying to UP in Spring 2011 through Spring 2013 for entry in grades 6-8 from school officials. The raw lottery records include each applicant's name, date of birth, contact information, lottery priority group, and lottery number. Three of these admission cohorts were found to have been oversubscribed by first-time applicants: 6th and 7th grade entry in Spring 2011, and 6th grade entry in Spring 2012. Appendix Table A5 describes these cohorts: from a total of 1,418 student names we find 1,015 first-time, non-sibling BPS students whose admission to UP was determined solely by lottery number.

We use student lottery numbers to construct two indicator variables for whether applicants were eligible to receive an offer to attend UP. The *immediate offer* instrument indicates admission offers made on the day of the lottery in March. The *waitlist offer* instrument indicates that a student has a lottery number better than the student with the worst lottery number who was offered admission to

UP from the waitlist by the end of September. Overall immediate and waitlist offer rates were 30 and 21 percent, respectively.

UP's lottery rosters do not include SASIDs; these records are matched manually to the SIMS by name, date of birth, application year and application grade. In some cases, this procedure did not produce a unique match and information on town of residence was used to break ties. Our matching procedure successfully located 96% (972) of UP Boston applicants in the SIMS/MCAS database. Excluding student not enrolled in BPS at baseline produces the UP lottery sample of 962 students.

Table A1: RSD full charter takeovers from 2008-09 to 2012-13

Closure year (spring)	Legacy school	Charter legacy?	Legacy grades	Takeover school	Takeover charter network	"No Excuses" network?	Takeover grades	Study takeover?
2010	A.D. Crossman: Esperanza Charter	Yes	K–8	Esperanza Charter School	Choice		K–8	
	John Dibert Elementary		PK–8	John Dibert Community School	FirstLine	Yes	PK–8	Yes
	Laurel Elementary		PK–8	SciTech Academy at Laurel Elementary	ReNEW	Yes	PK–8	Yes
	Live Oak Elementary		PK–8	Batiste Cultural Arts Academy at Live Oak Elementary	ReNEW	Yes	PK–8	Yes
	Harney Elementary		PK–8	Edgar P. Harney Spirit of Excellence Academy	Spirit of Excellence		K–8	Yes
	Gentilly Terrace Elementary		PK–8	Gentilly Terrace School	New Beginnings		PK–8	Yes
2011	Harriet Tubman Elementary	Yes	PK–8	Harriet Tubman Charter School	Crescent City	Yes	K–8	
	Joseph S. Clark Senior High		9–12	Joseph S. Clark Preparatory High School	FirstLine	Yes	9–12	
	Sarah Towles Reed Elementary		PK–8	Dolores T. Aaron Elementary	ReNEW	Yes	PK–8	Yes
2012	McDonogh #42 Charter	Yes	PK–8	McDonogh 42 Elementary Charter School	Choice		PK–8	
	Joseph A. Craig School		PK–8	Joseph A. Craig Charter School	Friends of King	Yes	PK–8	Yes
	John McDonogh Senior High		9–12	John McDonogh High School	Future is Now		9–12	
2013	Pride College Preparatory Academy	Yes	K-5	Mildred Osborne Charter School	Arise Academy		PK–6	
	Crocker Arts and Technology School	Yes	PK-5	Lawrence D. Crocker College Prep	New Orleans College Prep		PK–5	
	Paul B. Habans Elementary School		PK-6	Paul Habans Charter School	Crescent City	Yes	PK–6	Yes
	Murray Henderson Elementary School		1–5	Paul Habans Charter School	Crescent City	Yes	PK–6	Yes
	H.C. Schaumburg Elementary School		PK-8	Schaumburg Elementary	ReNEW	Yes	PK–8	Yes
	Abramson Science and Technology School		K-8	Schaumburg Elementary	ReNEW	Yes	PK–8	Yes

Notes: This table lists RSD's full charter takeovers from the 2008-09 to the 2012-13 academic years. Study takeovers are those involving a public-to-charter middle school takeover. "No Excuses" networks are identified using charter applications and school or network web sites. There were no full charter takeovers in the 2008-09 academic year.

Table A2: RSD and Boston school and teacher characteristics

	RSD				Boston			
	RSD direct-run (1)	RSD charter (2)	Legacy (3)	Takeover (4)	BPS (5)	Boston charter (6)	Gavin (7)	UP (8)
A. School characteristics								
Student-teacher ratio	--	--	--	--	12.3	11.8	13.5	11.7
Average class size	20.3	19.4	19.9	19.7	--	--	--	--
Per-pupil expenditures								
Reported	$13,104	$11,056	$11,682	$10,934	$17,948	$14,938	$15,054	$14,586
Adjusted	$11,104	--	--	--	$15,419	$14,000	$12,119	$13,441
B. Teacher characteristics								
Average age	--	--	--	--	42	32	41	28
Proportion inexperienced (age ≤ 28)	--	--	--	--	0.10	0.40	0.03	0.60
Average years of experience	12.4	7.0	--	--	--	--	--	--
Proportion inexperienced (exp ≤ 2)	0.22	0.41	--	--	--	--	--	--
Average salary	$48,080	$46,416	--	--	$81,963	$66,696	$77,251	$60,459
No experience	$40,400	$42,047	--	--	--	--	--	--
One year	$43,503	$43,803	--	--	--	--	--	--
Two years	$44,002	$44,695	--	--	--	--	--	--
10 years or more	$52,613	$50,530	--	--	--	--	--	--

Notes: Figures in columns 1-4 are calculated using data from http://www.louisianabelieves.com. Figures in columns 3-4 include any full takeover or legacy school for which data were available, including charter-to-charter and high school full takeovers. Average class size is based on the 2010-2011 and 2011-2012 academic years, and is calculated using the midpoint of reported class size ranges, with the exception of the class category size of "34 +", which is coded as an average class size of 34 students. Schools included in column 3's class size figure are Sarah Towles Reed Elementary, Fannie C. Williams Elementary, Harriet Tubman Elementary School, Joseph S. Clark Senior High, McDonogh 42 Elementary Charter, Joseph A. Craig, Crocker Arts and Technology, H.C. Schaumburg Elementary, Abramson Science and Technology, Pride College Preparatory Academy, Paul B. Habans Elementary, and Murray Henderson Elementary. Schools included in column 4's class size figure are SciTech Academy at Laurel Elementary, Esperanza Charter, Edgar P. Harney Spirit of Excellence Academy, Gentilly Terrace Elementary, Batiste Cultural Arts Academy at Live Oak Elementary, and Reed Elementary. New Orleans teacher experience and salary are based on operator-level data for the 2010-2011 academic year. Per-pupil expenditures (PPE) are averages over academic years 2008-2009 through 2010-2011. Column 1's PPE is based on aggregate figures reported for all RSD direct-run schools, and exclude one-time expenditures related to Hurricane Katrina. Schools included in column 3's PPE are Gentilly Terrace, E.P. Harney Spirit of Excellence Academy, Batiste Cultural Arts Academy at Like Oak Elementary, John Dibert Community School, SciTech Academy at Laurel Elementary, and Esperanza Charter. Schools included in column 4's PPE are AD. Crossman Esperanza Charter and Harriet Tubman. Adjusted PPE figures are provided for better comparability. For New Orleans, excess spending (relative to charters) in operations and management is deducted from RSD direct-run figures, as RSD direct-run expenditures include spending on building insurance premiums for all buildings overseen by RSD, including those operated by charters. (See Cowen Institute, The State of Public Education in New Orleans: School Finances, March 2011.) For Boston, adjusted PPE figures exclude special education expenses. Columns 5-6 include Boston traditional public and charter schools serving grades 6-8. Column 6 includes Academy of the Pacific Rim, Boston Collegiate, Boston Preparatory, Brooke Charter Rosindale, Excel Academy, MATCH, Neighborhood House, Roxbury Preparatory, and Smith Leadership Academy. Boston student/teacher ratio and teacher characteristics are averages for academic years 2010-2011 and 2011-2012. Student/teacher ratio is calculated based on October SIMS (Student Information Management System) and on teacher full-time equivalents obtained from EPIMS (Educational Personnel Information Management System). PPE figures for columns 5-6 are obtained from http://www.doe.mass.edu and http://profiles.doe.mass.edu, refer to fiscal year 2011-2012 and are enrollment-weighted. Gavin's PPE is calculated based on Gavin-specific instructional spending, as reported on Schedule 3 of Boston's FY11 End of Year Financial Report, and on Boston's FY11 average spending on school administration, pupil services, operations and management, and insurance and retirement program. Average teacher salary is obtained from http://www.doe.mass.edu/finance/statistics/.

Table A3: RSD and UP grandfathering cohorts

Closure Spring Year	Legacy school	Takeover charter network	Grandfathering-eligible students				Comparison students
			All	With baseline covariates	Direct-run/BPS at baseline	Unique analysis sample	
2010	John Dibert	FirstLine	140	92	86	78	204
	Laurel	ReNEW	176	102	96	76	873
	Live Oak	ReNEW	152	76	72	65	673
	Harney	Spirit of Excellence	121	63	56	55	425
	Gentilly Terrace	New Beginnings	156	111	107	106	543
2011	Sarah Towles Reed	ReNEW	215	143	137	81	267
2012	Joseph A. Craig	Friends of King	216	154	142	86	112
2013	Paul B. Habans	Crescent City	232	180	168	87	326
	Murray Henderson	Crescent City	135	85	66	53	168
	H.C. Shaumburg	ReNEW	111	81	71	55	274
	Abramson	ReNEW	25	24	23	21	45
	Pooled RSD:		1,657	1,105	1,019	763	2,410
2011	Gavin	Unlocking Potential	334	307	290	290	913

Notes: This table describes the sample of grandfathering-eligible students and their ineligible matches for the RSD study takeovers and for UP. Legacy school students are matched to comparison students by race, sex, baseline grade and year, special education status, subsidized lunch eligibility, and baseline school SPS scores in five-point bins.

Table A4: RSD and UP grandfathering attrition

	Sample means				Balance coefficients	
	RSD/Boston		Analysis sample		Analysis sample	
	RSD/Boston students	Charter-bound students	Takeover charter students	Grandfathering-eligible students	No score controls	Baseline score controls
	(1)	(2)	(3)	(4)	(5)	(6)
			A. RSD grandfathering			
Has legacy grade outcomes	0.750	0.855	0.958	0.936	0.003	
					(0.011)	
N	14,554	11,358	1,040	763	3,503	
Has first exposure year outcomes	0.672	0.776	0.828	0.710	0.008	
					(0.021)	
N	14,554	11,358	1,040	763	3,503	
Has second exposure year outcomes	0.629	0.728	0.814	0.738	0.031	
					(0.027)	
N	9,368	7,510	775	443	2,353	
Has third exposure year outcomes	0.603	0.708	0.755	0.669	0.027	
					(0.034)	
N	5,856	4,685	534	290	1,505	
Has fourth exposure year outcomes	0.481	0.573	0.599	0.488	0.000	
					(0.050)	
N	2,545	2,091	252	162	788	
			B. UP grandfathering			
Has legacy grade outcomes	0.909	0.973	0.938	0.893	-0.012	0.008
					(0.014)	(0.011)
N	8,506	1,552	225	290	1,203	1,100
Has first exposure year outcomes	0.884	0.938	0.933	0.855	-0.025	-0.013
					(0.018)	(0.017)
N	8,506	1,552	225	290	1,203	1,100
Has second exposure year outcomes	0.867	0.915	0.877	0.817	-0.065*	-0.054*
					(0.033)	(0.033)
N	7,996	1,444	130	164	626	565

Notes: This table reports sample means and coefficients from regressions of the variable in each row on a grandfathering eligibility dummy indicating enrollment in an takeover legacy school in the fall of the academic year prior to takeover, controlling for matching strata. Regressions in column 6 also control for 4th grade MCAS scores. The sample in columns 3-6 is restricted to students enrolled in an RSD direct-run school (panel A) or BPS school (panel B) at baseline. Column 1 reports means for a sample of RSD/Boston students in the same baseline years as the analysis sample, while column 2 is restricted to those students that enroll in an RSD/Boston charter school in grades following the baseline. Column 3 reports means for students that enroll in a takeover charter in potential takeover grades, while column 4 describes students enrolled in a legacy school. Robust standard errors are reported in parentheses.

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

Table A5: UP lottery records

	2011		2012	
	6th grade	7th grade	6th grade	Total
	(1)	(2)	(3)	(4)
Total number of records	791	170	457	1,418
Excluding late applicants	698	81	361	1,140
Excluding applicants from outside of BPS	666	79	323	1,068
Excluding siblings of UP students	652	61	302	1,015
Excluding records not matched to SIMS	621	53	298	972
In a BPS school at baseline	619	51	292	962

Notes: This table summarizes the sample restrictions imposed for the analysis of UP's lottery applicants.

Table A6: UP lottery descriptive statistics, balance, and attrition

	Sample means				Balance coefficients (Lottery applicants)	
	Boston students	Grandfathering-eligible students	Lottery applicants			
			6th grade	7th grade	Immediate offer	Waitlist offer
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Balance			
Hispanic	0.374	0.241	0.327	0.294	-0.023	0.035
					(0.033)	(0.037)
Black	0.375	0.469	0.483	0.373	0.004	-0.026
					(0.036)	(0.040)
White	0.124	0.152	0.090	0.196	0.014	-0.005
					(0.022)	(0.023)
Asian	0.073	0.100	0.055	0.098	0.025	-0.012
					(0.019)	(0.019)
Female	0.487	0.483	0.504	0.471	0.003	-0.017
					(0.036)	(0.040)
Special education	0.222	0.317	0.231	0.275	0.032	-0.001
					(0.031)	(0.034)
Free/reduced price lunch	0.794	0.928	0.802	0.843	0.019	-0.024
					(0.029)	(0.032)
Limited English proficient	0.280	0.307	0.248	0.275	0.023	-0.035
					(0.032)	(0.034)
N	6,744	290	911	51	962	962
Baseline math test score	0.003	-0.253	-0.054	-0.081	0.003	-0.064
					(0.066)	(0.073)
N	6,501	258	897	48	945	945
Baseline ELA test score	0.006	-0.235	-0.030	-0.169	-0.060	0.018
					(0.066)	(0.074)
N	6,387	254	890	47	937	937
			B. Attrition			
Has first exposure year outcomes	0.917	0.855	0.924	0.843	-0.016	0.033
					(0.020)	(0.021)
N	6,744	290	911	51	962	962
Has second exposure year outcomes	0.878	0.817	0.872	0.784	-0.034	0.014
					(0.025)	(0.028)
N	6,744	164	911	51	962	962
Has third exposure year outcomes	0.826		0.814		-0.047	-0.038
					(0.039)	(0.044)
N	4,294		617		617	617

Notes: This table reports sample means and coefficients from regressions of the variable in each row on either an immediate or waitlist offer dummy. The immediate offer dummy indicates that a lottery applicant was offered a seat in the March lottery, while the waitlist offer dummy indicates that an applicant was eligible for the offer of a seat off the waitlist from March to the end of September. All regressions include lottery risk set dummies. The sample in columns 2-6 is restricted to students enrolled in a BPS school at baseline. Column 1 reports means for a sample of Boston students in the same baseline grades and years as the analysis sample. Robust standard errors are reported in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table A7: Relaxing gains exclusion

	RSD			UP		
	Legacy score (1)	Outcome score (2)	Outcome gain (3)	Legacy score (4)	Outcome score (5)	Outcome gain (6)
			A. Math			
Legacy enrollment	-0.088**			0.106**		
	(0.040)			(0.046)		
Takeover enrollment		0.198***	0.217***		0.370***	0.337***
		(0.032)	(0.037)		(0.035)	(0.036)
Legacy score (λ)		0.527***			0.629***	
		(0.132)			(0.104)	
Gap to gain estimate		-0.020			0.034	
		(0.039)			(0.043)	
Instruments		31	31		61	61
N	2,553	5,625	5,625	1,028	1,543	1,543
			B. ELA			
Legacy enrollment	-0.030			-0.149***		
	(0.042)			(0.049)		
Takeover enrollment		0.139***	0.154***		0.309***	0.403***
		(0.030)	(0.037)		(0.037)	(0.041)
Legacy score (λ)		0.515***			0.399***	
		(0.129)			(0.090)	
Gap to gain estimate		-0.015			-0.093*	
		(0.037)			(0.051)	
Instruments		31	31		64	64
N	2,553	5,621	5,621	1,025	1,539	1,539

Notes: This table compares 2SLS estimates of takeover enrollment effects on math and ELA test scores under differing assumptions about the persistence of legacy score effects. The outcomes, sample and enrollment endogenous variables are as in Table 4 (RSD) and Table 8 (UP). The instruments in columns 2-3 and 5-6 are grandfathering eligibility interacted with baseline year, special education status, and baseline school SPS bin (RSD) or school (UP). The estimates in columns 2 and 5 treat legacy scores and takeover enrollment as endogenous. Column 1 reports the average effect of grandfathering eligibility on legacy scores, estimated by OLS. All models control for matching strata, limited English proficiency, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses.

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