Charters Without Lotteries: Testing Takeovers in New Orleans and Boston*

Atila Abdulkadiroğlu

Joshua D. Angrist

Peter D. Hull

Parag A. Pathak†

November 2014

Abstract

Lottery estimates suggest oversubscribed charter schools boost student achievement in urban districts. But these estimates needn't capture treatment effects for students who haven't applied to charter schools or for students attending charters where 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 impressive gains. In Boston, where we can compare takeover and lottery estimates, takeover charters generate achievement gains as large or larger than the gains for students assigned seats in lotteries.

*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, Olivia Kim, and Mayara Silva for exceptional research assistance and to 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 Arnold Foundation. Thanks also go to seminar participants at the Federal Reserve Bank of New York for helpful comments. The views expressed here are those of the authors alone.

†Abdulkadiroğlu: Duke University, e-mail: atila.abdulkadiroglu@duke.edu. Angrist: MIT and NBER, e-mail: angrist@mit.edu. Hull: MIT, e-mail: hull@mit.edu. Pathak: MIT and NBER, e-mail: ppathak@mit.edu.

No child's chances 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 decimated New Orleans' public schools in August 2005, along with the rest of the city's infrastructure. The ensuing scramble to reopen New Orleans schools prompted further legislative action in November of that year. 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. In the following years, as enrollment increased rapidly 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

[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.

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 direct-run 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 plots post-2008 achievement trends in RSD, OPSB, and the rest of Louisiana.

An important and distinctive feature of New Orleans charter expansion is 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 if they so desire. By contrast, most charter schools in other districts open as *startups*, that is, as new schools, with no seats guaranteed by virtue of previous enrollment.[3]

A similar though smaller-scale takeover intervention is also under way in the Boston Public School (BPS) district. Boston's experiment with charter takeovers has unfolded with less urgency than in New Orleans, but the forces behind it are similar. At the end of the 2010-2011 school year, 9 BPS schools 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]

In this paper we evaluate the causal effect of RSD and Boston takeover schools on student achieve-

[3]Most New Orleans startups either shared facilities with existing schools, without special treatment of students at the "roommate school," or opened in school buildings that had been closed since Katrina (Simon (2008) describes some of the startups).

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

ment 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 pool of estimates exploiting charter school admissions lotteries, while consistently showing large gains for urban charters, necessarily capture causal effects only for charter applicants, a self-selected population that may be especially likely to see gains from the charter treatment.[5] Lottery estimates also comes from schools that are in-demand: they have more applicants than seats. By contrast, grandfathered enrollment in charter takeovers is passive, with an existing population 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 charters also identifies effects for schools that are not otherwise over-subscribed.

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 technical 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 estimator 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 a rapidly diminishing share of direct-run school enrollment over the study period. A simple 2SLS modification addresses the problem of a possibly mixed 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."[6] The FY2011 federal budget operationalized this goal by adding three billion dollars to around $500 million previously appropriated for School Improvement 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):

[5]Lottery estimates are reported in, e.g., Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak (2011), Angrist, Dynarski, Kane, Pathak, and Walters (2012), Angrist, Pathak, and Walters (2013), Dobbie and Fryer (2011), Dobbie and Fryer (2013), Hoxby, Murarka, and Kang (2009), and Tuttle, Gill, Gleason, Knechtel, Nichols-Barrer, and Resch (2013). Rothstein (2004) challenges the external validity of charter treatment effects, writing (on page 82): "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)

[6]This is from a message from Education Secretary Duncan, in a message to school boards, uploaded May 27, 2010 (available through ed.gov).

> A restart model is one in which an 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 schools qualify for federal support under this heading.[7]

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 the grandfathering 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, and compares grandfathering and lottery estimates of achievement effects at UP Academy. The last section summarizes the 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 only 13 are startups. Not surprisingly, charter takeovers are especially controversial. In a recent critical account of school reform, for example, Ravitch (2010) discounts the premise behind takeovers, writing (on page 137):

> Success, whether defined as high test scores or graduation rates or student satisfaction, cannot be bottled and dispensed at will. This may explain why there are so few examples of low-performing schools that have been 'turned around' into high-performing schools ... Certainly schools can improve and learn from one another, but school improvements – if they are real – occur incrementally, as a result of sustained effort over years.

[7]SIGs were originally authorized under the 1965 Elementary and Secondary Education Act, but remained unfunded until 2007. The 2009 American Recovery and Reinvestment Act generated a substantial jump in federal SIG appropriations. Districts applied for this new money in the first half of 2010. In addition to Restart, other SIG-qualified interventions include Transformation, which replaces the principal and introduces multiple instructional and personnel reforms, and Turnaround, which adds the replacement of at least 50 percent of school staff to the Transformation model. SIGs can also be used to support outright closure. In a study discussed further below, Dee (2012) evaluates California's SIG experience. USDOE (2009) details SIGs further.

Others have associated charter takeovers with racial imbalances in New Orleans enrollment, a concern that has prompted legal action.[8] Such critical concerns notwithstanding, the perception that RSD's takeover policy has been fruitful has prompted ongoing explorations of similar approaches in Michigan and Tennessee.

Table 1 lists the 18 New Orleans RSD schools that experienced 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.[9] Alternatives to the full takeover model include a rolling grade-by-grade charter conversion, a principal-led conversion, and school mergers. We focus on full takeovers because this intervention is broad, well-defined, and easy to date, 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, since charter-to-charter conversion doesn't contribute to a charter school first stage. 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 to adjust for violations of the exclusion restriction are readily available for middle schoolers.[10]

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. Table 1 shows that the 11 takeover schools in our study are run by six charter management organizations (CMOs), with the Crescent City and ReNEW CMOs each operating multiple schools. Note also that in two cases in 2013, two legacy schools were merged into a single takeover school. Table 1 also shows that 9 out of 11 study schools describe themselves as embracing No Excuses pedagogy. The No Excuses model for urban education is characterized by tutoring and targeted remedial support, extensive reliance on data and feedback, college-preparation, a curriculum focused on basic skills, high expectations from students as well as staff, and an emphasis on discipline and comportment.[11]

[8] 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).

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

[10] 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 a 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.

[11] An online appendix table documents our sources for this classification. Two schools have proven difficult to classify.

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 A1 compares teacher characteristics at RSD direct-run and charter schools. Teachers at RSD charters tend to be younger, have less experience, and earn lower 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.[12]

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 middle schools opened in Lawrence, Massachusetts.[13] 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.[14]

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. 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 the only CMO 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).[15]

[12]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).

[13]Unlocking Potential schools in Lawrence enroll students in the relevant Lawrence Public Schools catchment area, and are not considered charters, though they operate with some of the same autonomy as Boston in-districts.

[14]BGA's founding headmaster and chairman of the board came from Boston Fenway High School, a Pilot School (see Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak (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).

[15]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 attended private day schools at district expense (as do many high-needs special education students in the district). Theses cases notwithstanding, the overall UP enrollment take-up rate for grandfathered special education students is close to that for other grandfathered

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 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 A1. 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 identifies with the No Excuses model.[16] 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 design that reveals the causal effects of SIG awards. Dee's estimates suggest that SIG-funded interventions improve performance for students at treated schools. A 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. Its 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.

students. Our IV strategy treats all grandfathered students similarly.

[16] 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, 2010).

Fryer (2014)'s quasi-experimental analysis uses baseline enrollment zones to construct instruments for exposure to treatment. Our grandfathering strategy is similarly founded, but uses matching to adjust for baseline differences associated with the grandfathering instrument. We'll soon document the value of matching in our context, while our 2SLS identification strategy 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: one high school, one middle school, and one K-8. 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. But these sorts of comparisons do not appear to fit into a causal framework except under stronger conditional independence assumptions than those invoked here.

Somewhat farther afield, though still in the takeover ballpark, Epple, Jha, and Sieg (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 enrolled in the fall 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 have school performance scores (SPS) in the same five-point bin as the legacy school.[17] In addition to baseline

[17]SPS scores range from 0-200 and are used for accountability purposes within RSD. 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.

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

In practice, the RSD grandfathering experiment involves multiple grades, schools, and years. The relationship between legacy, baseline, and takeover grades in each RSD grandfathering scenario is described in Table 2. Because the earliest baseline information available is from 3rd grade, the 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 or stratum may contain students that are eligible for grandfathering into multiple takeover charters in the same year. The grandfathering instrument indicates eligibility at any of the 11 takeovers 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 30% the size of the matched comparison group (Table A2 reports sample sizes by legacy school).

Math and English Language Arts (ELA) achievement data for RSD students come from 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 2008 through 2014.[18] 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 3 reports descriptive statistics for the RSD analysis sample and for a broader sample 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 falls more than 0.27σ below the corresponding 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

[18]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.

positively selected on baseline characteristics.[19]

The RSD comparison group appears to be well-matched to the RSD grandfathering cohort. This is documented in column 5 of Table 3, 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 in column 5 of Table 3 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 likelihood of qualifying for a subsidized lunch, in limited English proficiency rates, or in baseline scores.

Appendix Table A3 reports follow-up rates and gauges differential attrition from 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 3, baseline covariates remain balanced in the analysis subsample for which we can measure outcomes.[20]

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 grandfathered cohort and the comparison group move in parallel in the pre-takeover period, differences in score growth between the grandfathered group and the comparison group in the post-takeover period offer compelling evidence of a takeover treatment effect.

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 relative to this group. Consistent with RSD's focus on low-performing schools, 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-treatment 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 here effectively eliminates baseline differences, so that simple post-treatment comparisons

[19] In the middle school sample analyzed in Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak (2011), for example, the baseline math gap between charter applicants and Boston students is around 0.36σ.

[20] The availability of legacy year scores and grade repetition are similarly unrelated to grandfathering eligibility.

by grandfathering status 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 almost perfectly centered on zero, while achievement contrasts after the legacy year strongly favor the grandfathered cohort. Since over 65% of grandfathering-eligible students matriculated at a takeover charter, 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 design therefore allows for direct effects of legacy-school attendance when using legacy enrollment to instrument takeover attendance.[21]

3.3 Econometric Framework

We use an indicator for enrollment at a legacy school in the fall of the year preceding takeover as an instrument for takeover charter attendance, while removing direct effects of pre-takeover legacy enrollment with an additive structure for potential outcomes. 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: immediately prior to the takeover (grade l) and after the takeover (grade g), with the grandfathering-eligible group mostly enrolled in the takeover school in grade g. 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.

[21]As far as we know, other studies using persistent enrollment at legacy schools to construct instruments have not addressed this problem.

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 changes takeover exposure from D_0 to D_1. In this setup, observed outcomes are determined by potential outcomes and by the instrument according to

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

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 3 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, Imbens, and Rubin (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 $d = 0, 1$. 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 which 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 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.[22]

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

$$E[Y^l_z|T] = \alpha_1(T) + z\gamma(T) \tag{1}$$

$$E[Y^g_{zd}|T] = \alpha_2(T) + z\gamma(T) + d\beta(T). \tag{2}$$

The parameters $\alpha_t(T)$ in these expressions are subgroup-specific potential outcome means with both legacy and takeover enrollment indicators switched off; $\gamma(T)$ is an additive legacy school enrollment effect, common to grades l and g; and $\beta(T)$ is the causal effect of takeover attendance for subgroup T (this is identified only for $T = c$ in our setup). This model rules out interactions between legacy and takeover attendance, while also allowing legacy effects to be persistent.

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

Theorem 1 Under Assumptions 1-4,

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

This theorem allows us to recover average causal effects of takeover exposure on test score *levels*, while allowing for violations of the exclusion restriction due to legacy school enrollment (in the additive

[22] Assumption 4 provides an appealing alternative to regression control for Y^l: assuming $Y^l_1 \neq Y^l_0$, achievement in grade l is an outcome variable in our causal framework, so its inclusion in regression models for grade g outcomes likely induces selection bias (Angrist and Pischke, 2008).

model of equations (1) and (2), the theorem identifies $\beta(c)$).

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 variable measuring years of takeover exposure. The latter use is supported by a theoretical extension 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.

Motivated by these theoretical results, we estimate the causal effect of takeover attendance on achievement in an IV setup that links post-legacy score gains with treatment. The second-stage estimating equation is

$$Y_{it}^{g} - Y_{i}^{l} = \alpha_t + \sum_{j} \delta_j d_{ij} + \gamma' X_i + \beta D_{it} + \eta_{it}, \tag{3}$$

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.[23] 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). Finally, X_i is a vector of controls, described further below, and the d_{ij} are indicators for matching strata, indexed by j.

The corresponding first stage equation can be written

$$D_{it} = \lambda_t + \sum_{k} \mu_j d_{ij} + \iota' X_i + \pi Z_i + \nu_{it}, \tag{4}$$

where Z_i is the excluded instrument, indicating legacy school enrollment in the fall of the legacy school's final year in operation. 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, special education status, baseline school, and the legacy grade, equations (3) and (4) include a dummy for each of these cells. The empirical first- and second-stage models also include dummies for subsidized lunch status, English proficiency, and year-of-test (with coefficients α_t and λ_t). Finally, although baseline score controls appear to be uncorrelated with grandfathering exposure in RSD, we include them to boost precision.

[23]Test scores are from the first recorded attempt in the relevant grade. See the Data Appendix for details

4 Charters without Lotteries in New Orleans RSD

4.1 Grandfathering Results

Attendance at an RSD takeover charter increases math and ELA scores by an average of 0.2σ and 0.14σ, respectively, per year enrolled. These precisely estimated IV estimates, reported in the last column of Table 4, are generated by a first stage of about 1.1 years of takeover exposure (first stage estimates are reported in column 3).[24] Analyses that disaggregate by outcome grade and by potential takeover exposure time 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 two-thirds of legacy school students offered a takeover seat took it, at least initially.

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 4. This suggests that uninstrumented comparisons by enrollment status, such as those reported in CREDO (2013a), suffer from 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 0.15σ and 0.11σ, respectively. The downward bias of these estimates is consistent with the negative legacy-year treatment effect suggested by Figures 3 and 4.

The takeover-induced shift in the distribution of LEAP and iLEAP math and ELA scores is captured by Figure 5, which plots potential outcome distributions for lottery compliers, computed as in Walters (2014). Specifically, for a grid of values v in the support of an outcome variable y_i we estimated equations of the form:

$$K_h(v - y_i)(1 - D_i) = \sum_j \kappa_{0j}(v)d_{ij} + \gamma_0(v)(1 - D_i) + u_{0i}^v \tag{5}$$

$$K_h(v - y_i)D_i = \sum_j \kappa_{1j}(v)d_{ij} + \gamma_1(v)(1 - D_i) + u_{1i}^v \tag{6}$$

where the binary variable D_i, indicating takeover attendance up to and including the outcome grade, is instrumented by grandfathering eligibility. Here, $K_h(v) = \frac{1}{h}K(v/h)$ is a kernel function with bandwidth h. The resulting estimates of $\gamma_1(v)$ and $\gamma_0(v)$ describe estimated densities of outcomes for treated and untreated compliers. Difference in densities for compliers are causal effects. Note that the list of controls here is limited to matching cell fixed effects (the d_{ij}), so that differences in baseline score

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

densities are informative.[25]

Consistent with the balance in average baseline scores seen in Table 3, the distributions of Math and ELA scores for treated and control compliers in the baseline grade are virtually identical. This can can be seen in the upper panel of Figure 5. In later years, takeovers appear to shift the distribution of scores steadily rightward. With the exception of the third year of follow-up for ELA, Kolmogorov-Smirnov tests of distributional equality show that these post-treatment shifts are unlikely to be a chance finding (note that sample size declines farther out). As with the 2SLS estimates in Table 4, the distributional shifts plotted in Figure 5 are more pronounced for math than for ELA. The distributional analysis also reveals the broad-based nature of takeover effects on math: takeovers induce a marked shift throughout the distribution of scores.

4.2 Interpreting RSD Takeover Effects

The RSD grandfathering identification strategy compares students that mostly attends takeover charters with a grandfathering-ineligible comparison group that attended various sorts of schools. Many in the comparison group first attended one of RSD's direct-run public schools, but this counterfactual evolved as RSD closed its direct-run schools and as students moved across schools on their own, entering charters through lotteries instead of grandfathering. Estimates of RSD takeover effects therefore reflect a growing share of charter-to-charter comparisons. If non-study takeover charters also boost student achievement, such comparisons mask a higher overall charter treatment effect.

Table 5 describes the grandfathering attendance counterfactual in detail, focusing on the distinction between takeover charters, other RSD charters, and direct-run RSD schools. Specifically, the first two columns show the distribution of school types by grandfathering status, while column 3 describes the types of schools attended by untreated compliers. These complier attendance counterfactuals are determined by 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, described choices in non-treated and treated states. Column 3 of Table 5 reports estimates of $E[W_0|D_1 > D_0]$.[26] By definition, treated compliers enroll in a takeover school when grandfathered; column 4 in the table is included as a reminder of this fact.

The takeover first stage for enrollments contrasts a 79 percent first-year takeover enrollment rate

[25] As with other 2SLS estimates that include saturated controls for covariates, the parameters $\gamma_1(v)$ and $\gamma_0(v)$ are average densities, weighted across matching strata.

[26] These were constructed using an IV procedure that puts $(1 - D)W$ on the left hand side of an estimating equation that instruments $1 - D$ by grandfathering eligibility, controlling for matching strata and other baseline controls. The theory behind this procedure is detailed in Abadie (2002, 2003). See Table 1 in Abdulkadiroglu, Angrist, and Pathak (2014) for a similar analysis.

for those grandfathered (reported in column 2) with the 8.7 percent comparison group enrollment rate (reported in column 1).[27] The first-year increase in takeover enrollment reflects a substantial reduction in attendance at other charters (compare 40 with 18.2) and, especially, a sharp reduction in attendance at direct-run schools (compare 51 with 3.1). The counterfactual attendance distribution in column 3 shows that 36 percent of untreated compliers enrolled initially in a non-takeover charter school, while nearly two thirds attended a direct-run school.

Both the takeover first stage and the proportion of the comparison group in direct-run schools falls over time. The takeover first stage in the third year of exposure is around 0.49 (0.757 − 0.268), while the counterfactual direct-run enrollment share falls to 0.29. The balance of third-year non-treated complier enrollment was in other, non-takeover, charter schools. Reflecting RSD's complete charter transformation by Fall 2014, the counterfactual other-charter enrollment rate in our sample exceeded 98 percent after four years of exposure.[28]

The growing share of our sample enrolled in non-takeover charter schools dilutes estimated takeover effects if other charter schools generate similar gains. We therefore estimated a 2SLS model with two endogenous variables, one tracking takeover attendance and one tracking attendance at other charters. This model is

$$y_{it}^{g} - y_{i}^{l} = \alpha_t + \sum_j \delta_j d_{ij} + \gamma' X_i + \beta_D D_{it} + \beta_C C_{it} + \eta_{it}, \tag{7}$$

where C_{it} counts the number of years a student enrolls in a non-takeover RSD charter prior to testing. Equation (7) is identified by adding interactions between the grandfathering instrument 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 are not grandfathered.

Removing other charters from the counterfactual outcome distribution with the aid of equation (7) nearly doubles the estimated takeover effect on math scores. This can be seen in the contrast between the estimates in columns 1 and 2 of Table 6. Column 1 repeats the takeover effect for the all-grades sample shown in Table 4, while column 2 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 0.34σ. These results are remarkably similar to earlier estimates of charter middle school effects on

[27] The discrepancy between these estimates and the first stage in Table 4 is due to the lack of controls for matching strata and other covariates in Table 5.

[28] A few students in our comparison group attended a single direct-run school operating in Spring 2014; others repeated grades, adding a fourth year of direct-run exposure in 2012 and 2013.

Boston lottery applicants.[29] On the other hand, the other-charter ELA effect in column 2 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 2 of Table 6 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 (7) 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 the number of years enrolled in any RSD charter. The resulting estimates of β_A, reported in columns 3 and 4 of Table 6 for just-identified and over-identified specifications (that is, without and with covariate interactions in the instrument list), indeed shown a precision gain, with standard errors falling from .071 and 0.147 in column 2 to .057 in column 4. 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 2 make the pooled ELA results harder to interpret.[30]

5 Measuring UP in Boston

Our RSD analysis suggests charter takeover attendance can boost 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 backdrop. The availability of oversubscribed admission lotteries at UP also facilitates a 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, and 5th grade special education status (Table 2 describes the timing of the grandfathering research design for UP). As for RSD, each grandfathered student is matched to multiple comparison students.

[29] See, e.g., Abdulkadiroğlu, Angrist, Dynarski, Kane, and Pathak (2011).

[30] 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 3 and 4 exceed those in columns 1 and 2 reflect the negative weight this scheme assigns to the other-charter effect.

The resulting analysis sample contains 290 grandfathering-eligible Gavin students, with about 1,000 students in the comparison group.[31]

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 and 2013 (data from 2014 are used in the lottery design). 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 7, which describes the population of Boston 5th graders in years covered by the UP analysis sample and 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 in seen in the wider sample of charter-bound students in Table 7, 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 7. The estimates in column 5 are from models that control only for matching cells; these show a small but marginally significant difference in subsidized lunch rates, along with a larger marginally significant grandfathering gap in baseline ELA scores. The differences in column 5 suggest the comparison group here is not as well-matched as for RSD. Importantly, 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, though a small difference in subsidized lunch status persists.

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

[31] As can be seen in Table 2, baseline information for the Boston analysis sample comes from 5th grade for students whose last legacy grade was 7th as well as for those whose last legacy grade was 6th.

Appendix Table A3. Follow-up differences are more pronounced for the cohort seen two years out, though these differences are still only marginally significant. As with the subsidized lunch differential, this modest difference in follow-up rates seems unlikely to account for the strong 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 6 and 7, which plot achievement paths in the same format used for RSD in Figures 3 and 4. Confidence bands for difference-in-differences comparisons are plotted with dotted lines in panel b 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. The DD estimates show marked and statistically significant differences in score growth in in post-treatment years, with no significant differences earlier.

Interestingly, and in contrast with RSD, the legacy-year treatment effects generated by the Gavin experiment are not significantly different from zero. This may reflect the fact that (conditional on covariates) the Gavin experience differed little from that at other BPS schools, while the pre-closure education provided by RSD's legacy schools was indeed substantially worse than that available elsewhere in RSD.

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 8. The first stage that generates these results is around 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 reveal 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 the table, is close to 80 percent. Math results 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.

As noted in the outline of our econometric framework in Section 3, IV estimates of models with variable treatment intensity generate (or approximate) a weighted average of per-unit average causal effects. Panel B of Table 8 reports results disaggregated by potential years of exposure, contrasting estimates for the cohort exposed for at most one year, a Bernoulli treatment, with estimates for those exposed for as much as two years. The average causal effect of a year at UP on students' ELA scores falls after the second year of exposure, from 0.49σ to 0.26σ. Given the exceptionally large first-year ELA impact this seems unsurprising, and is consistent with Figure 7's difference-in-differences evidence for ELA, which shows a post-takeover achievement jump, followed by a plateau.[32]

5.4 Lottery Estimates

Since the fall of 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 (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 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 for the grandfathering estimates, the lottery analysis is limited to lottery applicants who attended a BPS elementary school in the baseline grade. Appendix Table A4 gives an account of UP lottery applicant record processing.

Table 9 describes the UP lottery sample and documents baseline covariate balance by win/loss status. 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.

In contrast with other group of Boston lottery applicants, UP lottery applicants are not positively

[32]UP results without differencing post-takeover and legacy-grade scores are similar to those reported in Table 8, at 0.43σ for math and 0.24σ for ELA.

selected: average baseline scores for lottery applicants are just below the Boston mean. Importantly, however, this differs from the Gavin cohort, which has baseline scores roughly a quarter of standard deviation below those for Boston, as can be seen in column 2 of Table 9 (this repeats column 4 in Table 7). UP lottery applicants are also less likely than the grandfathered cohort to be poor enough to qualify for a subsidized lunch. Finally, consistent with random assignment of lottery offers, the balance coefficients in columns 5-8 of the table show UP lottery winners and losers to be similar. Likewise, we see little evidence of excess loss to follow-up in the loser group (these results appear in Appendix Table A5).

The lottery estimation framework looks much like that described by equations (3) and (4), with the modification that there's no matched comparison sample. Rather, the estimation sample consists of all applicants, as described above, while the empirical models adjust for year and grade of application instead of matching cell fixed effects. The instruments in this case are dummies indicating the offer of a seat at UP immediately following the lottery, or from the waitlist (specifically, the waitlist instrument indicates applicants with lottery numbers below the highest number offered a seat in the relevant application cohort through September). As before, the endogenous variable here counts years enrolled at UP between application and test date.

The first stage effect of an immediate offer, close to .8 for the full sample, exceeds the first stage for waitlist offers, which is just under .6. These first stage estimates–for years at UP–appear at the top of columns 2 and 3 of Table 10. 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 .52 and .4, respectively. These estimates appear in the first row of panel B of the table.

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 10, 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, rising to 0.34σ in math and 0.20σ in ELA. Likewise, disaggregation by exposure time generates larger average effects after one year than two.

The results in Tables 9 and 10 suggest the benefits of UP enrollment for those enrolled there by virtue of grandfathering are at least as large as for those who won their seats in a lottery. The contrast in gains estimates also favors grandfathering when we compare students who've had equal exposure to UP: after one year, gains for the lottery cohort are 0.365σ in math and 0.22σ in ELA, while gains after one year for those grandfathered into UP come to 0.35σ in math and 0.49σ in ELA in 7th and 8th grade. Gains for the grandfathered cohort after two years of exposure are estimated to be 0.31σ

in math and 0.26σ 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 wind up at other high-performing Boston charters, thereby diluting lottery-generated treatment effects.

Roughly 87 percent of untreated compliers in the grandfathering research design enrolled in a traditional BPS school, with about 7 percent winding up in another Boston charter. This can be seen in Table 11, which details UP first stages and counterfactual school choices in the same format as Table 5. By comparison, the lottery design leaves 95 percent of untreated compliers in a traditional BPS school, with only 3 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 in the lottery counterfactual, imply that the excess of grandfathering over lottery estimates is not explained by a diluted counterfactual in the control group.

6 Summary and Conclusions

Charter school takeovers in the New Orleans Recovery School District generate impressive achievement gains for a highly disadvantaged student population that enrolled in these schools passively. The average gains for grandfathering compliers reflect a substantial and broad shift in the distribution of achievement. We find similarly large effects for 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 (Earlier results along these lines include (Angrist, Pathak, and Walters, 2013; Angrist, Cohodes, Dynarski, and Walters, 2014; Dobbie and Fryer, 2013)).

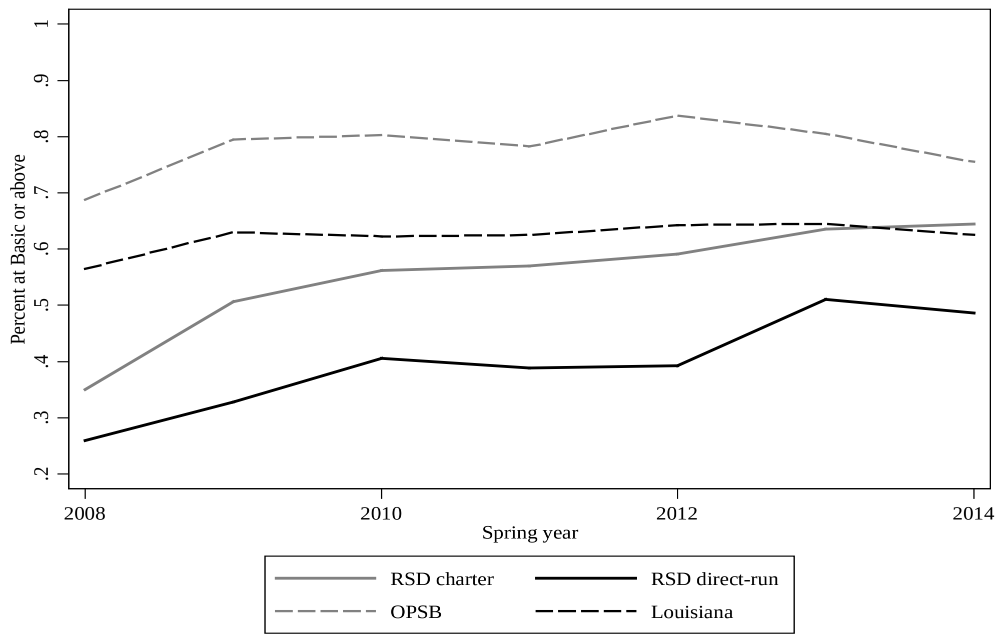
Our analysis of the takeover model highlights two important methodological issues that seem likely to crop up in other work using similar research designs. First, while legacy school enrollment provides a valuable source of exogenous variation in charter exposure, the grandfathering IV strategy should adjust for violations of the exclusion restriction due to legacy exposure *per se*. This adjustment is substantively important for RSD schools, where legacy enrollment depresses pre-takeover achievement. Second, in an environment with schools of many types, takeover and other sorts of charter treatment effects may be diluted by charter attendance in the control group. A simple 2SLS procedure allowing for multiple treatment channels generates a more homogeneous and easier-to-interpret counterfactual.

In practice, cleaning up the non-charter counterfactual substantially boosts our estimates of RSD takeover effects on math.

The strong results for RSD and the juxtaposition of estimates from grandfathering and lottery-based research designs for Boston's UP Academy weigh against the view that charter lottery applicants enjoy an unusual and unrepresentative benefit from charter attendance because they're highly motivated or uniquely primed to benefit from charter attendance. Boston and RSD takeovers generate gains for their passively enrolled students similar to the lottery estimates reported in Angrist, Pathak, and Walters (2013) for a sample of Massachusetts urban charters. The achievement gains from 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 are detailed in Fryer (2014)).

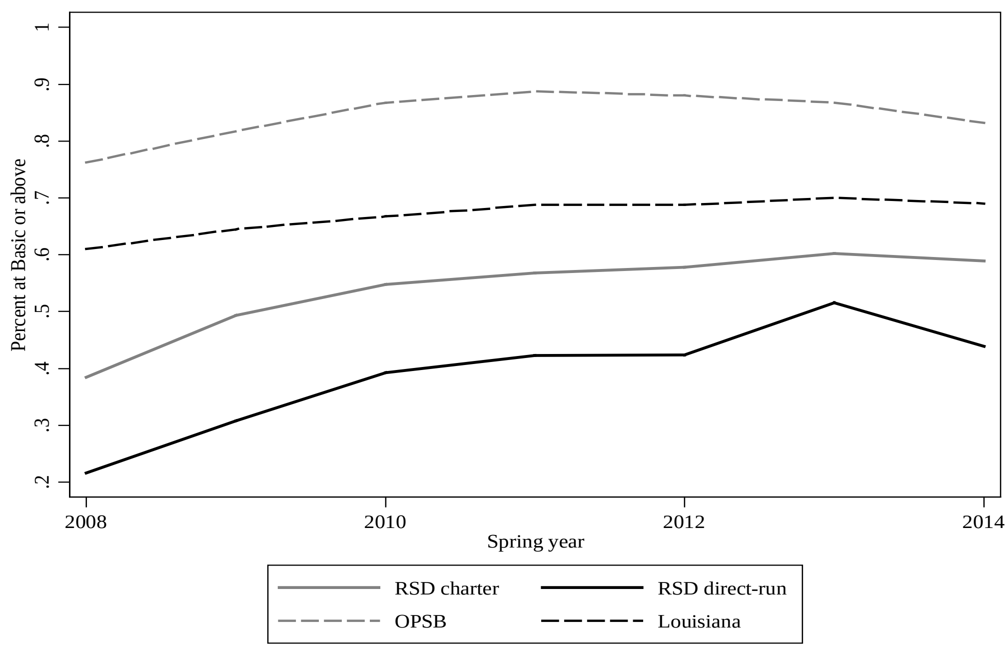
Finally, our findings highlight the subtle interplay between school choice and educational value added (Walters (2014) investigates this link in Boston). In a pioneering effort to streamline charter admissions and broaden school choice, RSD introduced a district-wide match known as OneApp in 2012. OneApp combined all RSD schools, direct-run and charter, matching students and schools using the tools of market design (Abdulkadiroğlu, Pathak, and Roth, 2014). 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. Our findings suggest the possibility of substantial gains from centralized school assignment mechanisms such as OneApp, that promote charter attendance among students who might not otherwise choose to apply.

Figure 1a: Math scores in RSD and elsewhere



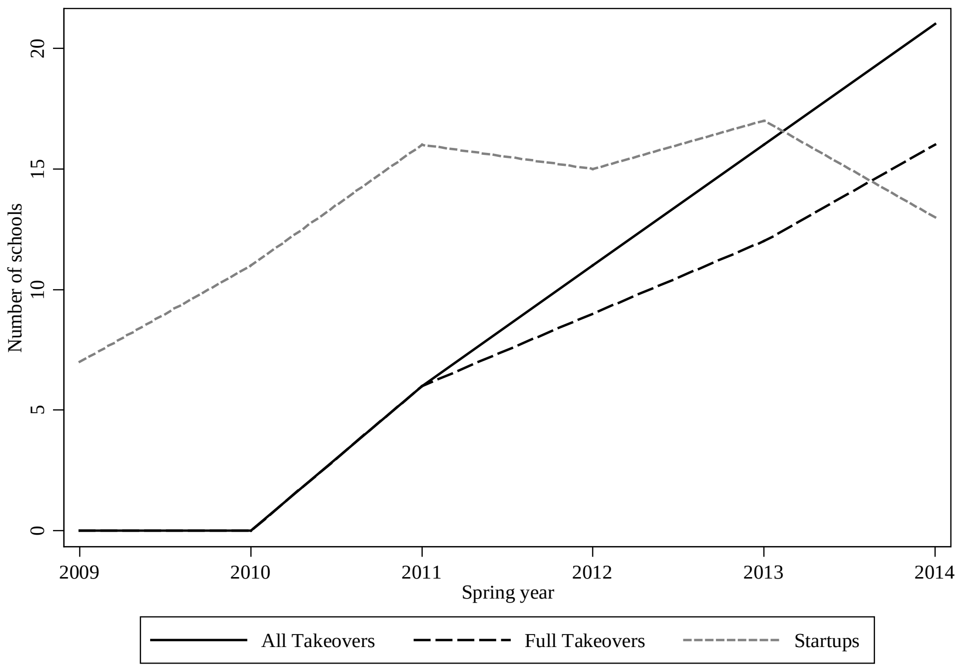
Notes: This figure plots the average percentage of RSD, OPSB, and Louisiana students that achieve Basic or above status on LEAP/iLEAP math exams in 5th-8th grades. Scores for OPSB and Louisiana are from https://www.louisianabelieves.com/resources/library/test-results. Statistics plotted are unweighted averages across grades for each year, and are computed separately for students enrolled in RSD charter and RSD direct-run schools.

Figure 1b: ELA scores in RSD and elsewhere



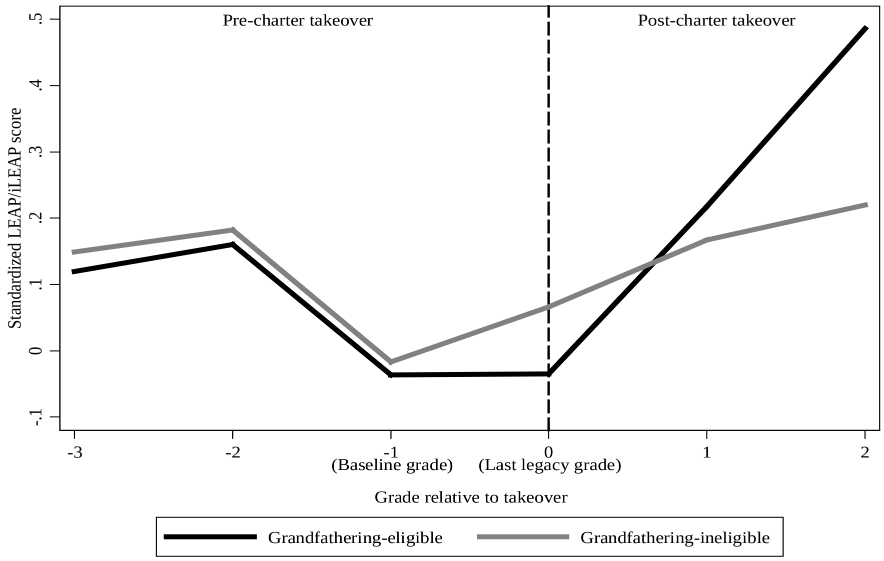
Notes: This figure plots the average percentage of RSD, OPSB, and Louisiana students that achieve Basic or above status on LEAP/iLEAP ELA exams in 5th-8th grades. Scores for OPSB and Louisiana are from https://www.louisianabelieves.com/resources/library/test-results. Statistics plotted are unweighted 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 the New Orleans Recovery School District



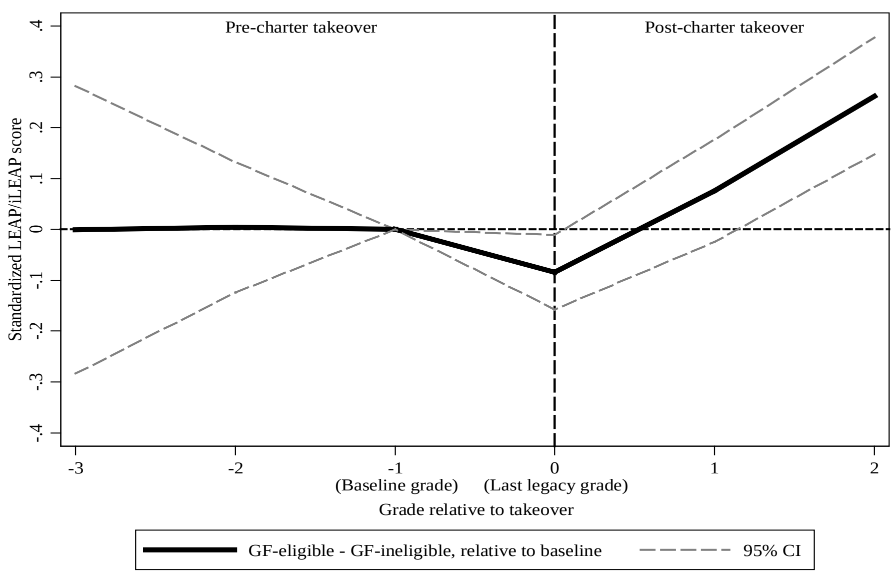
Notes: This figure plots the number of New Orleans RSD charter schools (serving any grades) created between academic years 2008-09 and 2013-14, excluding alternative schools. Takeovers are schools whose creation was directly tied to the closure of a particular school, with seats reserved in the first year of operation for students from that legacy school. Full takeovers are those with grandfathering provisions for all grades, excluding charter mergers and principal-led conversions. 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



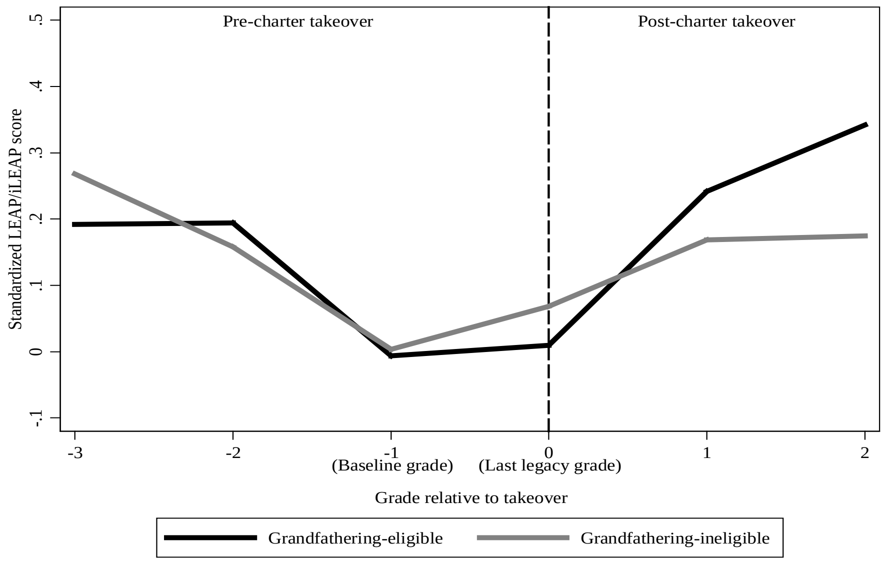
Notes: This figure plots average LEAP/iLEAP math scores of students in the RSD legacy middle school matched sample, controlling for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in the set of direct-run schools in New Orleans RSD.

Figure 3b: Grandfathering DD for math



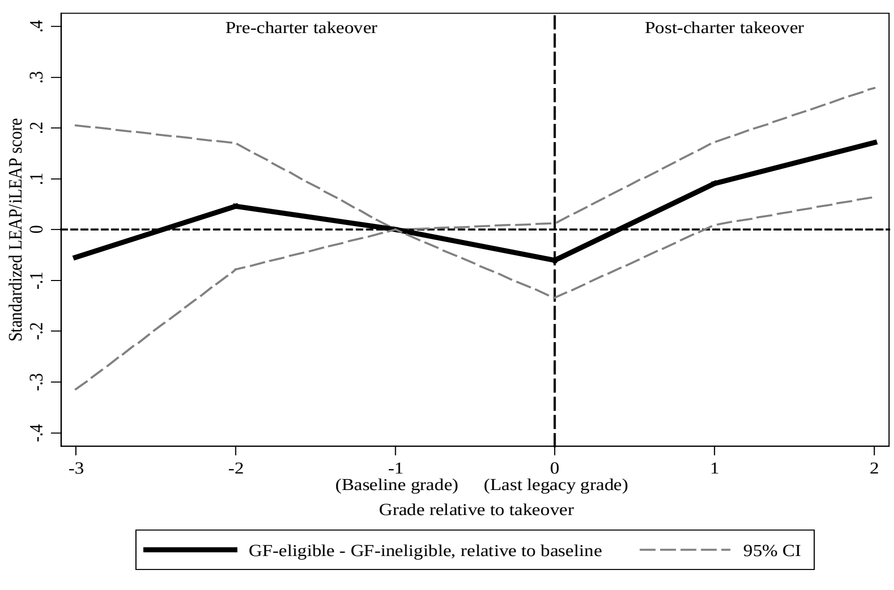
Notes: This figure plots estimated difference-in-differences effects of RSD takeover charter eligibility on LEAP/iLEAP math scores for students in the RSD legacy middle school matched sample. All estimates control for matching cell fixed effects. Scores are standardized to have mean zero and standard 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



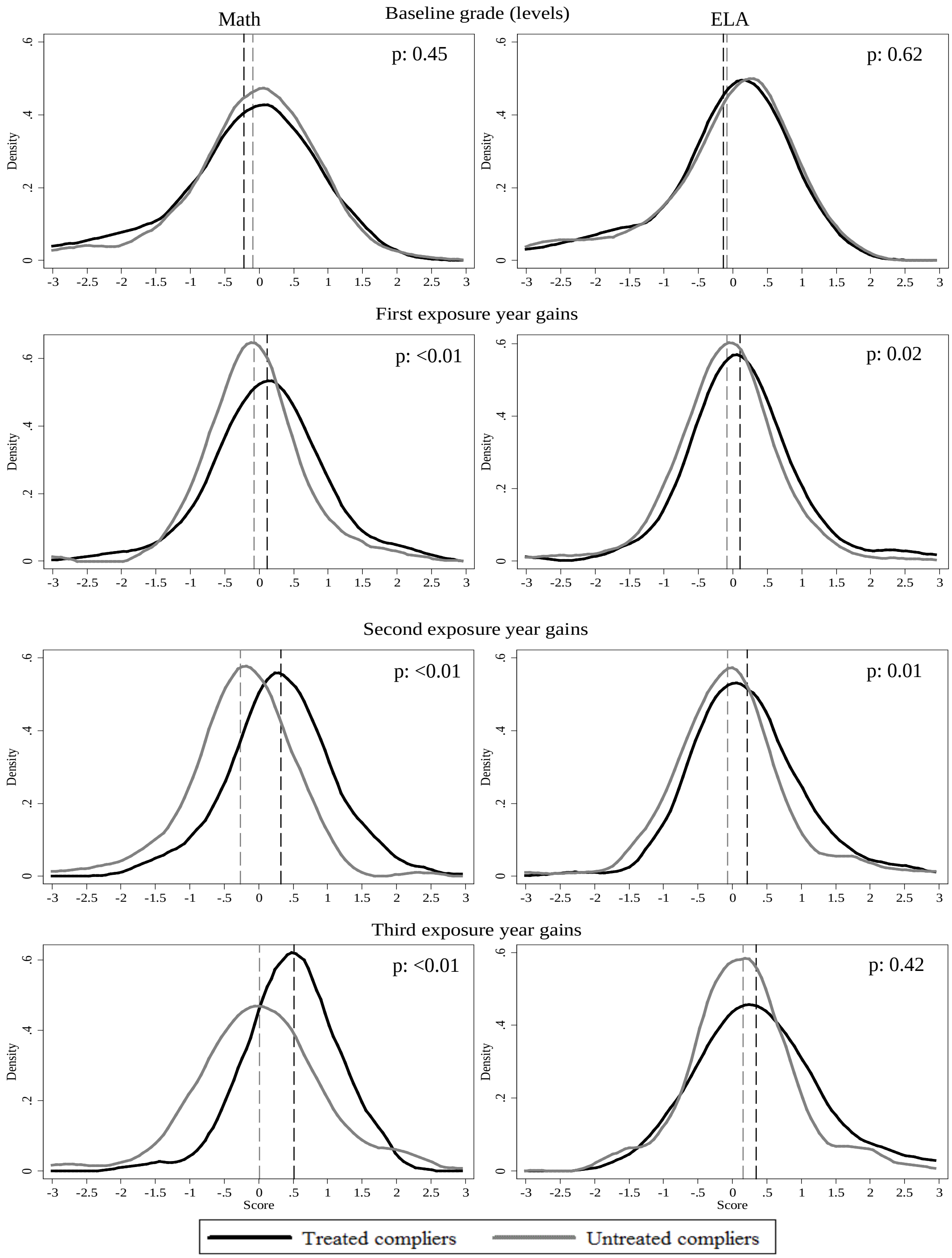
Notes: This figure plots average LEAP/iLEAP ELA scores of students in the RSD legacy middle school matched sample, controlling for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in the set of direct-run schools in New Orleans RSD.

Figure 4b: Grandfathering DD for ELA



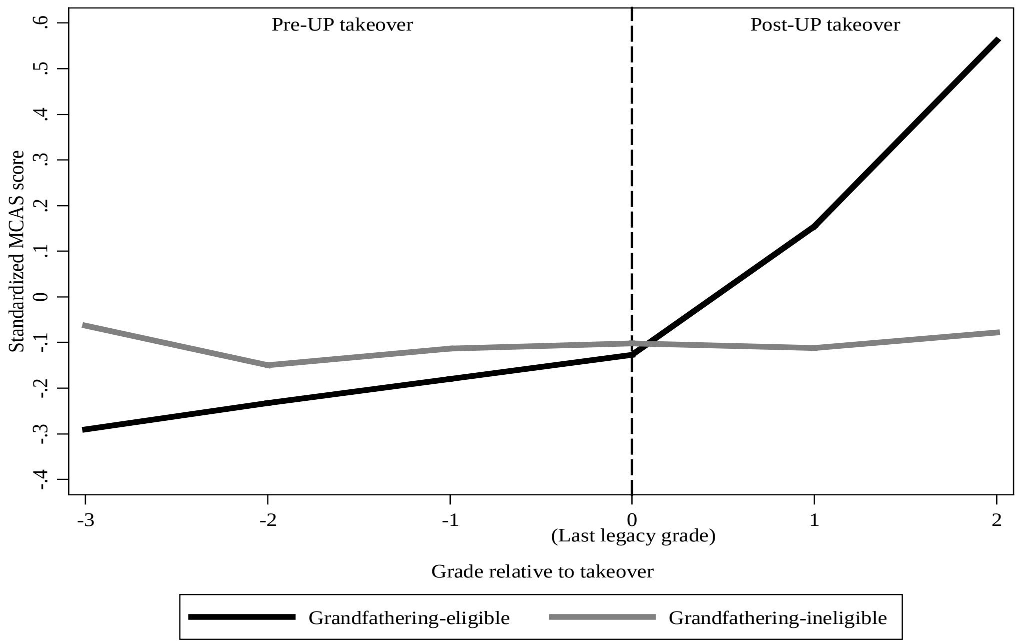
Notes: This figure plots estimated difference-in-differences effects of RSD takeover charter eligibility on LEAP/iLEAP ELA scores for students in the RSD legacy middle school matched sample. All estimates control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in the set of direct-run schools in New Orleans RSD.

Figure 5: RSD grandfathering complier densities



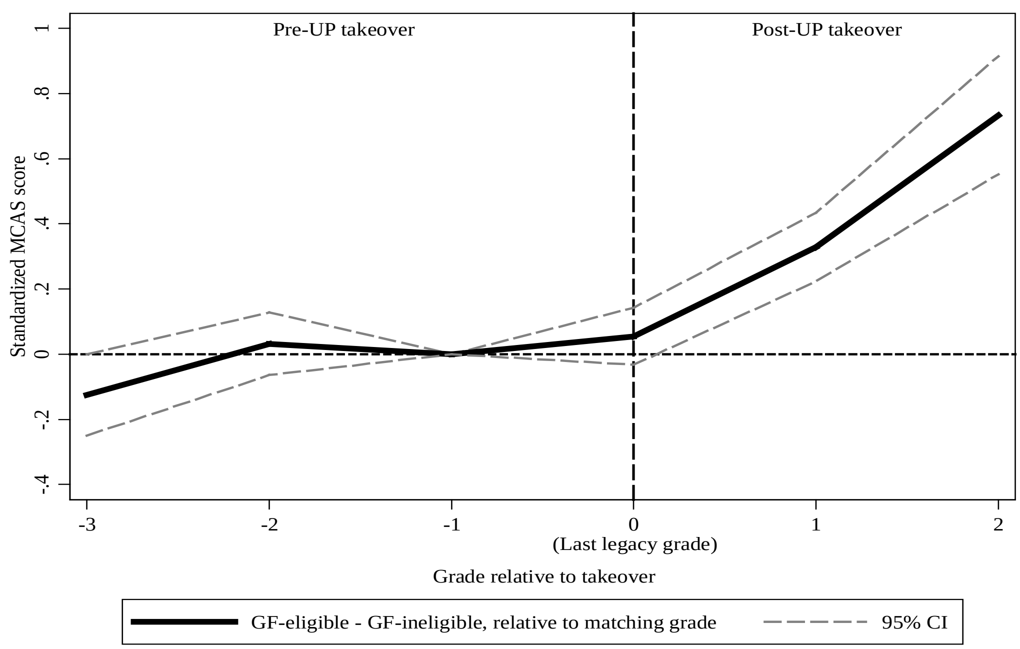
Notes: This figure plots smoothed LEAP/iLEAP math and ELA score levels and gains distributions for treated and untreated grandfathering compliers, estimated by instrumental variables. The first panel shows baseline score levels; other panels show gains from legacy to outcome grades. Densities are estimated using an Epanechnikov kernel function with a bandwidth equal to 1.25 times the Silverman (1986) rule-of-thumb. Means of each distribution are represented by dashed lines. Bootstrapped Kolmogorov-Smirnov test statistic p-values for distributional equality are reported in each panel.

Figure 6a: Math Scores for the UP grandfathering sample



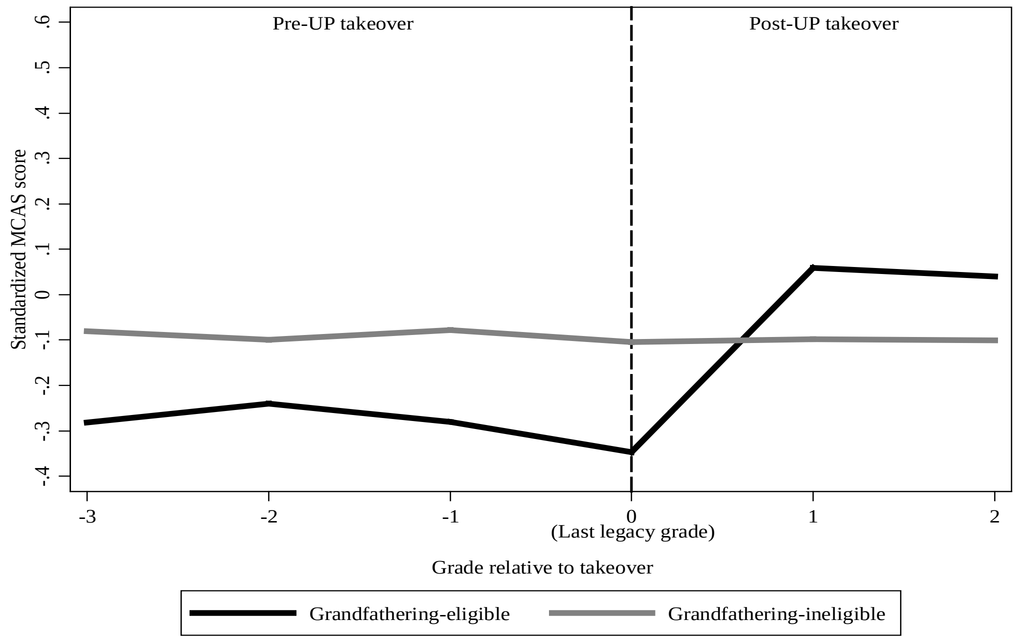
Notes: This figure plots average MCAS math scores of students in the Gavin Middle School matched sample, controlling for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in Boston.

Figure 6b: Grandfathering DD for math



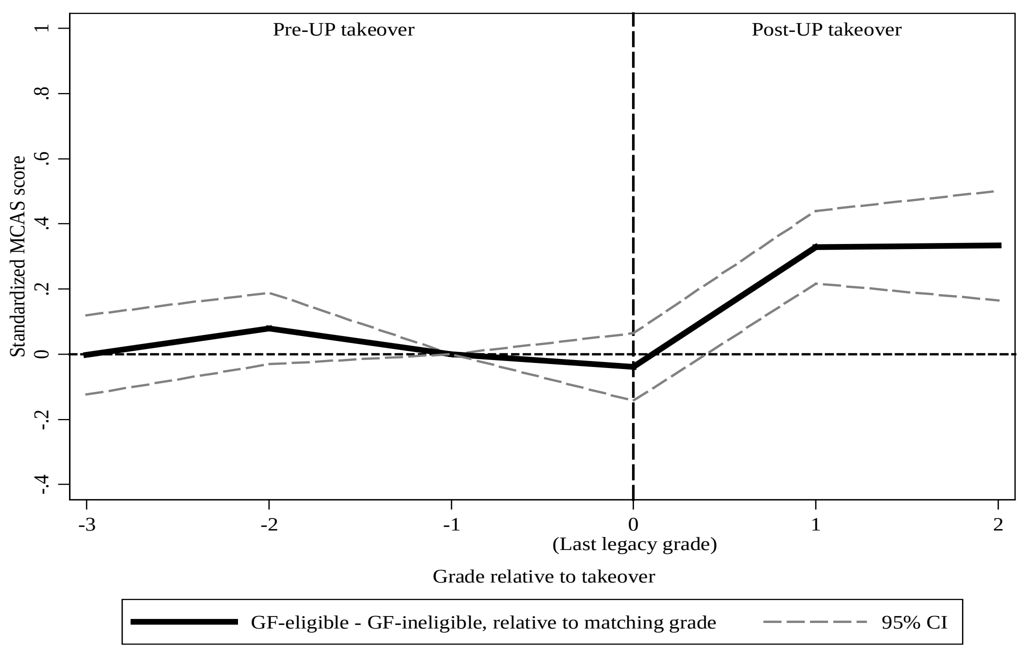
Notes: This figure plots estimated difference-in-differences effects of UP grandfathering eligibility on MCAS math scores for students in the Gavin Middle School matched sample. All estimates control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in Boston.

Figure 7a: ELA Scores for the UP grandfathering sample



Notes: This figure plots average MCAS ELA scores of students in the Gavin Middle School matched sample, controlling for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in Boston.

Figure 7b: Grandfathering DD for ELA



Notes: This figure plots estimated difference-in-differences effects of UP grandfathering eligibility on MCAS ELA scores for students in the Gavin Middle School matched sample. All estimates control for matching cell fixed effects. Scores are standardized to have mean zero and standard deviation one within each year and grade in Boston.

Table 1: RSD full charter takeovers: 2008-09 to 2012-13

Closure spring year	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 describes 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 those self-identified with the term as part of their educational philosophy, as described in each network's RSD charter applications or website, and documented in the online appendix. There were no full charter takeovers in the 2008-09 academic year.

Table 2: Timing in the grandfathering research design

	RSD				UP	
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 3: 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	0.023	0.018	0.029	--	--
Black	0.963	0.971	0.994	0.981	--	--
White	0.019	0.016	0.008	0.016	--	--
Asian	0.008	0.008	0.001	0.009	--	--
Female	0.476	0.474	0.486	0.503	--	--
Special education	0.069	0.066	0.070	0.095	--	--
Free/reduced price lunch	0.913	0.927	0.938	0.920	-0.002	-0.004
					(0.013)	(0.015)
Limited English proficient	0.017	0.016	0.012	0.020	0.000	-0.001
					(0.001)	(0.001)
N	14,478	11,303	1,052	746	3,650	2,670
Baseline math score	0.006	0.022	-0.315	-0.274	-0.018	-0.050
					(0.048)	(0.052)
N	12,882	10,488	1,050	743	3,647	2,668
Baseline ELA score	0.006	0.028	-0.304	-0.272	-0.009	-0.032
					(0.048)	(0.055)
N	12,885	10,491	1,052	745	3,649	2,670
Baseline science score	0.008	0.020	-0.219	-0.221	-0.045	-0.083
					(0.049)	(0.058)
N	12,853	10,471	1,045	741	3,632	2,658
Baseline social science score	0.006	0.025	-0.229	-0.227	-0.041	-0.034
					(0.047)	(0.055)
N	12,847	10,466	1,044	741	3,624	2,657

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, baseline grade and year, and baseline school SPS scores in five-point bins). The sample for 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 4: Grandfathering IV estimates of RSD takeover attendance effects

		Comparison group mean	OLS	2SLS estimates	
				First stage	Attendance effect
		(1)	(2)	(3)	(4)
			A. By grade		
All grades	Math	0.009	0.119***	1.088***	0.202***
	(N: 5,853)		(0.019)	(0.052)	(0.038)
	ELA	-0.039	0.077***	1.089***	0.139***
	(N: 5,851)		(0.018)	(0.052)	(0.039)
5th & 6th grades	Math	-0.008	0.085**	0.740***	0.137**
	(N: 2,675)		(0.035)	(0.042)	(0.068)
	ELA	-0.057	0.011	0.747***	0.099
	(N: 2,675)		(0.033)	(0.042)	(0.071)
7th & 8th grades	Math	0.023	0.131***	1.374***	0.233***
	(N: 3,178)		(0.020)	(0.069)	(0.036)
	ELA	-0.023	0.100***	1.372***	0.167***
	(N: 3,176)		(0.018)	(0.070)	(0.036)
			B. By potential exposure		
First exposure year	Math	-0.039	0.178***	0.657***	0.194***
(5th-8th grades)	(N: 2,647)		(0.044)	(0.023)	(0.069)
	ELA	-0.063	0.100**	0.657***	0.191***
	(N: 2,648)		(0.042)	(0.023)	(0.068)
Second exposure year	Math	-0.003	0.163***	1.164***	0.323***
(6th-8th grades)	(N: 1,746)		(0.030)	(0.061)	(0.058)
	ELA	-0.043	0.090***	1.174***	0.163***
	(N: 1,747)		(0.028)	(0.061)	(0.052)
Third & fourth exposure year	Math	0.108	0.091***	1.718***	0.116***
(7th & 8th grades)	(N: 1,460)		(0.022)	(0.130)	(0.041)
	ELA	0.009	0.073***	1.718***	0.087**
	(N: 1,456)		(0.020)	(0.131)	(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 for columns (2)-(4) includes RSD direct-run school students matched to a pre-takeover year legacy school student as described in Table 3. The endogenous regressor counts the number of years enrolled in a takeover charter prior to testing. All models control for matching strata, limited English proficiency, subsidized price lunch eligibility, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Means in column (1) are for grandfathering-ineligible matched students.

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

Table 5: 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 full takeover charter	0.087	0.787	--	1.000
… in other RSD charter	0.401	0.182	0.362	--
… in RSD direct-run	0.511	0.031	0.638	--
N	2,137	516		
	B. Second exposure year			
Enrolled in full takeover charter	0.201	0.717	--	1.000
… in other RSD charter	0.503	0.236	0.475	--
… in RSD direct-run	0.296	0.048	0.525	--
N	1,437	314		
	C. Third exposure year			
Enrolled in full takeover charter	0.268	0.757	--	1.000
… in other RSD charter	0.572	0.217	0.715	--
… in RSD direct-run	0.160	0.026	0.285	--
N	831	189		
	D. Fourth exposure year			
Enrolled in full takeover charter	0.306	0.646	--	1.000
… in other RSD charter	0.660	0.342	0.982	--
… in RSD direct-run	0.033	0.013	0.018	--
N	359	79		

Notes: This table describes school enrollment in the RSD analysis sample by grandfathering eligibility status. 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. Complier means are estimated by 2SLS, controlling for matching strata, limited English proficiency, subsidized price lunch eligibility, baseline test scores, and year/grade effects.

Table 6: Grandfathering IV estimates of RSD charter attendance effects

	(1)	(2)	(3)	(4)
		A. Math (N:5,853)		
Takeover charter	0.202***	0.358***		
	(0.038)	(0.071)		
	[438.0]	[424.0]		
Other RSD charter		0.336**		
		(0.147)		
		[49.2]		
Any RSD charter			0.367***	0.369***
			(0.069)	(0.057)
			[231.6]	[485.0]
Instruments	1	31	1	31
		B. ELA (N:5,851)		
Takeover charter	0.139***	0.136*		
	(0.039)	(0.080)		
	[437.0]	[422.5]		
Other RSD charter		-0.023		
		(0.171)		
		[48.1]		
Any RSD charter			0.251***	0.208***
			(0.072)	(0.059)
			[236.6]	[494.5]
Instruments	1	31	1	31

Notes: This table reports 2SLS estimates of the effects of 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 3. The endogenous regressors count the number of years enrolled in a takeover, other, or any RSD charter prior to testing. The instrument for columns (1) and (3) is grandfathering eligibility. For columns (2) 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, subsidized lunch eligibility, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Angrist-Pischke first-stage F statistics are reported in brackets.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 7: 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	0.345	0.272	0.238	0.241	--	--	--
Black	0.407	0.519	0.511	0.469	--	--	--
White	0.136	0.153	0.128	0.152	--	--	--
Asian	0.072	0.024	0.088	0.100	--	--	--
Female	0.484	0.505	0.489	0.483	--	--	--
Special education	0.227	0.185	0.269	0.317	--	--	--
Free/reduced price lunch	0.804	0.749	0.943	0.928	0.052**	0.051**	0.057**
					(0.021)	(0.022)	(0.022)
Limited English proficient	0.230	0.131	0.339	0.307	0.045	0.036	0.029
					(0.031)	(0.032)	(0.033)
N	8,353	1,535	227	290	1,341	1,172	1,107
Baseline math score	0.008	0.177	-0.292	-0.253	-0.102	-0.019	-0.013
					(0.067)	(0.050)	(0.052)
N	7,912	1,504	212	258	1,270	1,147	1,090
Baseline ELA score	0.012	0.184	-0.279	-0.235	-0.160**	0.011	0.000
					(0.063)	(0.051)	(0.052)
N	7,795	1,500	210	254	1,226	1,146	1,089

Notes: This table reports sample means and coefficients from regressions of the variable on each row on a grandfathering eligibility dummy indicating enrollment in Gavin Middle School in 6th or 7th grade in fall 2010. All regressions include cell fixed effects (cells are defined by race, sex, special education status, and 5th grade school and year). Regressions in column (6) and (7) also control for 4th grade MCAS scores. The sample in columns (3)-(7) 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 years as the analysis sample, while column (2) is restricted to those students that enroll in a Boston charter school in grades 6-8. Column (3) reports means for students that enroll in UP in grades 7 and 8 in the analysis sample, while column (4) describes students enrolled in Gavin Middle School in fall 2010. Robust standard errors are reported in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 8: Grandfathering IV estimates of UP attendance effects

		Comparison group mean	OLS	2SLS estimates: First stage	2SLS estimates: Attendance effect
		(1)	(2)	(3)	(4)
			A. By grade		
All grades	Math	-0.007	0.400***	1.032***	0.332***
	(N: 1,727)		(0.030)	(0.041)	(0.038)
	ELA	0.028	0.285***	1.026***	0.388***
	(N: 1,723)		(0.034)	(0.041)	(0.044)
7th grade	Math	-0.009	0.405***	0.791***	0.336***
	(N: 601)		(0.060)	(0.037)	(0.066)
	ELA	0.016	0.568***	0.793***	0.635***
	(N: 597)		(0.070)	(0.037)	(0.080)
8th grade	Math	-0.006	0.393***	1.177***	0.334***
	(N: 1,126)		(0.032)	(0.052)	(0.038)
	ELA	0.034	0.230***	1.164***	0.283***
	(N: 1,126)		(0.036)	(0.052)	(0.044)
			B. By potential exposure		
First exposure year (7th & 8th grades)	Math	0.001	0.374***	0.809***	0.350***
	(N: 1,143)		(0.045)	(0.026)	(0.049)
	ELA	0.027	0.455***	0.800***	0.490***
	(N: 1,140)		(0.054)	(0.027)	(0.060)
Second exposure year (8th grade)	Math	-0.023	0.401***	1.507***	0.313***
	(N: 584)		(0.037)	(0.087)	(0.043)
	ELA	0.029	0.211***	1.509***	0.263***
	(N: 583)		(0.041)	(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 includes BPS students matched to a 2010-11 6th or 7th grade Gavin Middle School student as described in Table 7. The endogenous regressor counts the number of years enrolled in UP prior to testing. The instrument is a grandfathering eligibility dummy. All models control for matching strata, limited English proficiency, subsidized price lunch eligibility, baseline test scores, and year/grade effects. Robust standard errors, clustered by student, are reported in parentheses. Means in column (1) are for grandfathering-ineligible matched students.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 9: UP lottery descriptive statistics, balance and attrition

	Sample means				Balance coefficients			
	Boston	Grandfathering-	Lottery applicants		Lottery applicants		First exposure year sample	
	students	eligible students	6th grade	7th grade	Immediate offer	Waitlist offer	Immediate offer	Waitlist offer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hispanic	0.374	0.241	0.327	0.294	-0.023	0.035	-0.033	0.039
					(0.033)	(0.037)	(0.035)	(0.039)
Black	0.375	0.469	0.483	0.373	0.004	-0.026	0.012	-0.037
					(0.036)	(0.040)	(0.038)	(0.041)
White	0.124	0.152	0.090	0.196	0.014	-0.005	0.007	0.001
					(0.022)	(0.023)	(0.023)	(0.024)
Asian	0.073	0.100	0.055	0.098	0.025	-0.012	0.030	-0.013
					(0.019)	(0.019)	(0.021)	(0.020)
Female	0.487	0.483	0.504	0.471	0.003	-0.017	0.003	-0.027
					(0.036)	(0.040)	(0.038)	(0.041)
Special education	0.222	0.317	0.231	0.275	0.032	-0.001	0.028	0.019
					(0.031)	(0.034)	(0.032)	(0.035)
Free/reduced price lunch	0.794	0.928	0.802	0.843	0.019	-0.024	0.008	-0.019
					(0.029)	(0.032)	(0.031)	(0.034)
Limited English proficient	0.280	0.307	0.248	0.275	0.023	-0.035	0.025	-0.033
					(0.032)	(0.034)	(0.034)	(0.036)
N	6,744	290	911	51	962	962	885	885
Baseline math test score	0.003	-0.253	-0.054	-0.081	0.003	-0.064	-0.023	-0.053
					(0.066)	(0.073)	(0.069)	(0.074)
N	6,501	258	897	48	945	945	884	884
Baseline ELA test score	0.006	-0.235	-0.030	-0.169	-0.060	0.018	-0.076	0.002
					(0.066)	(0.074)	(0.067)	(0.076)
N	6,387	254	890	47	937	937	878	878

Notes: This table reports sample means and coefficients from regressions of the variable on 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 being offered a seat off the waitlist from March to the end of September. All regressions include lottery cohort dummies. The sample in columns (2)-(8) 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 10: Lottery IV estimates of UP attendance effects

		Comparison group mean	OLS	2SLS estimates: First stage: Immediate offer	Waitlist offer	Attendance effect
		(1)	(2)	(2)	(3)	(4)
A. By grade						
All grades	Math	0.050	0.301***	0.760***	0.562***	0.270***
	(N: 2,202)		(0.022)	(0.063)	(0.067)	(0.056)
	ELA	0.075	0.148***	0.759***	0.562***	0.118**
	(N: 2,205)		(0.020)	(0.063)	(0.067)	(0.051)
6th grade	Math	0.065	0.334***	0.505***	0.407***	0.350***
	(N: 838)		(0.045)	(0.035)	(0.040)	(0.090)
	ELA	0.043	0.234***	0.505***	0.407***	0.214**
	(N: 839)		(0.045)	(0.035)	(0.040)	(0.091)
7th & 8th grades	Math	0.041	0.296***	0.919***	0.653***	0.248***
	(N: 1,364)		(0.021)	(0.085)	(0.088)	(0.054)
	ELA	0.093	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 (6th & 7th grades)	Math	0.058	0.347***	0.519***	0.397***	0.365***
	(N: 881)		(0.044)	(0.034)	(0.038)	(0.086)
	ELA	0.046	0.239***	0.521***	0.394***	0.220**
	(N: 882)		(0.044)	(0.034)	(0.038)	(0.088)
Second & third exposure year (7th & 8th grades)	Math	0.045	0.294***	0.921***	0.665***	0.242***
	(N: 1,321)		(0.021)	(0.088)	(0.091)	(0.054)
	ELA	0.092	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 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 in 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 for applicants not given an immediate or waitlist offer.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 11: School choice in the UP analysis samples

	All students		Compliers	
	Z=0	Z=1	Z=0	Z=1
	(1)	(2)	(3)	(4)
	A. UP grandfathering sample			
Enrolled in UP	0.013	0.794	--	1.000
...other Boston charter	0.075	0.009	0.072	--
…BPS	0.827	0.137	0.872	--
…other Massachusetts	0.085	0.060	0.056	--
N	920	233		
	B. UP lottery sample			
Enrolled in UP	0.040	0.513	--	1.000
...other Boston charter	0.177	0.156	0.047	--
…BPS	0.751	0.309	0.947	--
…other Massachusetts	0.033	0.022	0.005	--
N	429	456		

Notes: This table describes school enrollment of students in the UP grandfathering and lottery samples, for the first exposure year, by instrument status. 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 offered (Z=1) and not offered (Z=0) students, while columns (3)-(4) show the same for lottery offer compliers. Complier means in panel A are estimated by 2SLS controlling for matching strata, limited English proficiency, subsidized price lunch eligibility, baseline test scores, and year/grade effects. Complier means in panel B are estimated by 2SLS controlling for lottery cohort dummies and student race, sex, special education status, limited English proficiency, subsidized lunch status, baseline test scores, baseline school, and year/grade effects.

References

ABADIE, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variables Models," *Journal of the American Statistical Association*, 97(457).

——— (2003): "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, 113(2).

ABDULKADIROGLU, A., J. ANGRIST, AND P. PATHAK (2014): "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools," *Econometrica*, 82(1), 137–196.

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(2), 699–748.

ABDULKADIROĞLU, A., P. A. PATHAK, AND A. E. ROTH (2014): "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. COHODES, S. M. DYNARSKI, AND C. WALTERS (2014): "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice," *Journal of Labor Economics*, forthcoming.

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(4), 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(430), 431–442.

ANGRIST, J. D., P. A. PATHAK, AND C. R. WALTERS (2013): "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, 5(4), 1–27.

ANGRIST, J. D., AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Ash, K. (2013): "Charters Adopt Common Application Systems," Education Week, September 25.

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

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, https://credo.stanford.edu/documents/NCSSAccessed: 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.

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

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(3), 158–187.

Dobbie, W., and R. G. Fryer (2013): "Getting Beneath the Veil of Effective Schools: Evidence from New York City," *American Economic Journal: Applied Economics*, 5(4), 58–75.

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

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

Fryer, R. (2014): "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments," Available at: http://scholar.harvard.edu/files/fryer/files/2014_injecting_charter_school_best_practices_into_traditional_public_schools.pdf, Last Accessed: October 10, 2014.

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.

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

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

NACPS (2013a): "Back to School Tallies: Estimated Number of Public Charter Schools and Students, 2012-2013," .

——— (2013b): "A Growing Movement: America's Largest Charter School Communities," .

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," .

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

Simon, D. (2008): "Five new N.O. charter schools now have homes," Available at: http://www.nola.com/news/index.ssf/2008/03/charter_school_sites_chosen.html, Last accessed: November 7, 2014.

Toness, B. V. (2010): "Charter School Will Do Things the Hard Way," 90.9WBUR, December 9, 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," Mathematica Policy Research, Princeton, NJ.

UP (2010): "Application for a Massachusetts Horace Mann III Public Charter School: UP Academy Charter School of Boston," Available at: http://www.doe.mass.edu/charter/finalists/10/UPApp.doc'.

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.

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 that

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

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

Under Assumption 4, the first two terms of this expression are equal to zero and

$$E[Y^g_{00} - Y^l_0|D_1 > D_0] = E[Y^g_{10} - Y^l_1|D_1 > D_0].$$

Since $E[D|Z = 1] - E[D|Z = 0] = P(D_1 > D_0)$, we have

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

Since Assumption 4 also implies $E[Y^g_{11} - Y^l_1|D_1 > D_0] = E[Y^g_{01} - Y^l_0|D_1 > D_0]$, we have

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

We assume throughout that Assumptions 1-4 hold within a set of mutually-exclusive and exhaustive matching cells $\{d_j\}$. As shown by Abadie (2003), an IV regression of $Y^g - Y^l$ on D and $\{d_j\}$ that instruments D by Z would then identify a fixed-effects regression of $Y^g - Y^l$ on D for compliers. By the derivation in Angrist (1998) this becomes a weighted average of cell-specific LATEs, with weights proportional to the within-cell variance of the instrument and the share of compliers in each cell. In practice we estimate this regression with additional baseline and year-of-test controls. When such controls are independent of Z within cells and are fully interacted with $\{d_j\}$ the weighted average LATE estimand is unchanged, though it may be more precisely estimated. We approximate this model with an additive specification for controls, nothing that our estimates are very similar to those obtained with fully-interacted covariates.

Extension of Theorem 1 to an Ordered Treatment

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

Assumption 1′ *(Independence)* $Y_0^l, Y_1^l, Y_{00}^g, ..., Y_{0\bar{d}}^g, Y_{10}^g, ..., Y_{1\bar{d}}^g, D_0, D_1 \perp\!\!\!\perp Z$.

We also modify Assumption 4, assuming that the instrument Z is excludable from the treatment effect on gains for all individuals:

Assumption 4′ *(Strong Gains Exclusion)* $P(Y_{1d}^g - Y_1^l = Y_{0d}^g - Y_1^l) = 1$ for $d \in \{0, 1, ..., \bar{d}\}$.

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

$$\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} = \sum_{d=1}^{\bar{d}} \frac{E[(Y_{1d}^g - Y_1^l) - (Y_{1d-1}^g - Y_1^l)|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}$$
$$= \sum_{d=1}^{\bar{d}} \frac{E[Y_{1d}^g - Y_{1d-1}^g|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}.$$

Likewise,

$$\frac{E[Y^g - Y^l|Z=1] - E[Y^g - Y^l|Z=0]}{E[D|Z=1] - E[D|Z=0]} = \sum_{d=1}^{\bar{d}} \frac{E[(Y_{0d}^g - Y_0^l) - (Y_{0d-1}^g - Y_0^l)|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}$$
$$= \sum_{d=1}^{\bar{d}} \frac{E[Y_{0d}^g - Y_{0d-1}^g|D_1 \geq d > D_0]P(D_1 \geq d > D_0)}{\sum_{d=1}^{\bar{d}} P(D_1 \geq d > D_0)}.$$

As in the binary treatment case, we let these assumptions hold within matching cells, and control linearly for cell indicators and other baseline controls. As shown by Angrist and Imbens (1995) a 2SLS specification where Z is fully interacted with $\{d_j\}$ in the first stage produces an interpretable weighted average of cell-specific Average Causal Responses. Again we approximate this model with an additive specification for covariates, verifying that these estimates are quite close.

B Data Appendix

The data used in this study come from a variety of sources.

New Orleans enrollment, demographics, and outcomes data are provided by the New Orleans Recovery School District (RSD). New Orleans student enrollment and demographics data ("NOLA enrollment") include information on all students enrolled in the New Orleans RSD. Test scores used in the analysis of New Orleans RSD students are from the Louisiana Educational Assessment Program (LEAP) and integrated LEAP (iLEAP), also provided by the New Orleans RSD.

Boston student enrollment, demographics, and outcomes are provided by the Massachusetts Department of Elementary and Secondary Education. Boston enrollment and demographics data come from the Student Information Management System (SIMS), a centralized database that covers all public school students in Massachusetts. Test scores used in analysis of Boston students are from the Massachusetts Comprehensive Assessment System (MCAS). Finally, lists of UP Boston charter applicants and lottery winners are constructed from records provided by UP Academy of Boston. This Appendix describes each data source and details the procedures used to clean and match them.

Student enrollment and demographics data

Data description and sample restrictions

Our analysis of the New Orleans RSD and UP Boston uses student enrollment and demographics data from school year 2007-2008 through 2013-2014.

The NOLA 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, a February file is available. Each record in the NOLA enrollment files refers to a student in a school in a year, though there are student-school-year duplicates for students that switch grades within a school and year. In addition, the NOLA enrollment files include the first and last dates of enrollment for each student in each school-year-grade program. This information is combined with other variables to determine school attendance, as described in detail below. NOLA enrollment files include a unique student identifier, the "student ID." We use information on student names and dates of birth to confirm whether student IDs are consistent within and across enrollment files. After resolving any inconsistencies, student IDs are used to match students to the LEAP/iLEAP files, as described below.

The Massachusetts SIMS data include an October file and an end-of-year file for every school year. Similar to the NOLA enrollment files, each observation in the SIMS refers to a student in a school in a year, though there are some student-school-year duplicates for students that switch grades within a school and year. The SIMS includes a unique student identifier known as the SASID, which is used to match students to the MCAS data files, as described below.

Coding of demographics and attendance

The NOLA enrollment and SIMS variables used in our analysis include grade, year, name, date of birth, gender, race, special education and limited English proficiency status, free or reduced price lunch and school attended. We construct two wide datasets, one for the New Orleans RSD and one for Massachusetts, that capture demographic and attendance information for every student. These wide datasets store information from the first calendar year spent in each grade. If a student attended multiple schools within the same year and grade program, preference was given first to attendance at a takeover charter (NOLA enrollment) or at UP Boston (SIMS), and second to the longest-attended school. For the NOLA enrollment wide dataset, information on enrollment dates were further used to give preference to the school of latest enrollment. Any other attendance ties were broken at random.

In the grandfathering analysis, the grandfathering instrument is based on Fall semester attendance. For the New Orleans RSD, the instrument is coded following the procedure described above as applied to all enrollment records, but restricting the sample to schools enrolled by October 31, provided the student did not leave the school by that same date. For UP Boston, the instrument is coded following the procedure described above applied to enrollment data from the October enrollment files.

For both NOLA enrollment and SIMS, students classified as special education, limited English proficiency, or eligible for a free- or reduced-price lunch in any record within a school-year-grade retain that designation for the entire school-year-grade.

LEAP/iLEAP and MCAS Data

We use LEAP/iLEAP and MCAS data from the 2007-2008 school year through the 2013-2014 school year. Each observation in these outcome databases corresponds to a student's test results in a particular grade and year.

The LEAP/iLEAP outcomes of interest are math, English Language Arts (ELA), sciences, and social sciences tests for grades 5 through 8. We also use baseline tests taken prior to charter takeover, which are from grades 3 through 6 depending on a student's grade at the time of takeover. 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. We use the first test taken in a particular subject and grade.

The MCAS outcomes of interest are math and English Language Arts (ELA) tests in grade 10. We also use baseline tests taken prior to charter application, which are from 5th grade or 6th grade depending on a student's application grade. 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. We use the first test taken in a particular subject and grade.

UP Boston Lottery Data

Data description and sample restrictions

Our sample of applicants is obtained from records of lotteries held at UP Academy of Boston for enrollment in the fall semester of academic years 2011-2012 and 2012-2013. The raw lottery records include applicant names, dates of birth, contact information and other information used to define lottery groups, such as sibling status. Appendix Table A4 shows the sample restrictions we impose on the raw lottery records. We exclude late applicants, applicants from outside Boston Public Schools (BPS), and sibling applicants, as these groups are typically not included in the standard lottery process. We further exclude applicants that cannot be matched to SIMS, and applicants with multiple applications. Imposing these restrictions reduces the number of lottery records from 1,406 to 968.

Lottery offers

In addition to the data described above, the lottery records also include information regarding offered seats. We used this information to reconstruct indicator variables for whether lottery participants received randomized offers. We make use of two sources of variation in UP Boston's offers, which differ in timing. The *immediate offer* instrument captures offers made on the day of the lottery. The *waitlist offer* instrument captures offers made later, as a consequence of movement down a randomly sequenced waiting list. Immediate and waitlist offer rates were 30 an 21 percent, respectively, in our UP Boston lottery analysis sample.

Matching Data Sets

Grandfathering samples

Two grandfathering master datasets are constructed based on the files mentioned above: one for New Orleans, and one for Boston. Outcome data files in each site (MCAS or LEAP/iLEAP) are merged to the respective master enrollment and demographic files using the unique SASID/student ID. We used information on student names, dates of birth, and town of residence, whenever available, to confirm that matches were successful.

UP Boston lottery sample

UP Boston's lottery records do not include SASIDs; these records are matched manually to the SIMS by name, application year and application grade. In some cases, this procedure did not produce a unique match. We accepted some matches based on fewer criteria where the information on grade, year and town of residence seemed to make sense.

Our matching procedure successfully located most UP Boston applicants in the SIMS database. The fifth row of Appendix Table A4 reports the number of applicant records matched to the SIMS in each applicant cohort. The overall match rate across all cohorts is 97 percent (983/1,014).

Our lottery analysis also excludes students who did not attend a Boston Public Schools (BPS) school at baseline, as students applying from private schools have lower follow-up rates. This restriction eliminates 15 applicants to UP Boston. Of the remaining 967 applicants, 889 (92 percent) contribute a score to our MCAS analysis.

Table A1: 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				
Avg 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	--	--	--	--	--	--
Avg 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 based on 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-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 class size figure of column (3) 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 class size figure of column (4) 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. Per-pupil expenditures are averages over academic years 2008-2009 through 2010-2011. Per-pupil expenditures in column (1) are based on aggregate figures reported for all RSD direct-run schools, and exclude one-time expenditures related to Hurricane Katrina. Schools included in PPE figure of column (3) 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 PPE figure of column (4) are AD. Crossman Esperanza Charter and Harriet Tubman. Adjusted PPE figures are provided for better comparability. For Boston, adjusted PPE figures exclude special education expenses. 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. (For details, see Cowen Institute, The State of Public Education in New Orleans: School Finances, March 2011.) New Orleans teacher experience and salary are based on operator-level data for the 2010-2011 academic year. 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. Student/teacher ratio and Boston 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). Average per-pupil expenditures (PPE) 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 A2: RSD and UP grandfathering cohorts

Closure Spring Year	Legacy school	Takeover charter network	Grandfathering-eligible students				Comparison students	Enrollment retention
			All	With baseline covariates	Direct-run/BPS at baseline	Unique analysis sample		
2010	John Dibert	FirstLine	140	92	86	78	218	65.58%
	Laurel	ReNEW	176	102	96	74	938	51.16%
	Live Oak	ReNEW	152	76	72	64	670	52.90%
	Harney	Spirit of Excellence	121	63	56	54	402	59.80%
	Gentilly Terrace	New Beginnings	156	111	107	106	571	63.64%
2011	Sarah Towles Reed	ReNEW	215	143	137	79	277	55.50%
2012	Joseph A. Craig	Friends of King	216	154	142	85	119	52.48%
2013	Paul B. Habans	Crescent City	232	180	168	82	337	59.21%
	Murray Henderson	Crescent City	135	85	66	53	173	33.77%
	H.C. Shaumburg	ReNEW	111	81	71	50	278	61.27%
	Abramson	ReNEW	25	24	23	21	47	29.60%
	Pooled RSD:		1,657	1,105	1,019	746	2,534	
2011	Gavin	Unlocking Potential	334	307	290	290	1,051	69.06%

Notes: This table summarizes 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, baseline special education status, and baseline school SPS scores in five point bins. Enrollment retention is computed as the proportion of students enrolled in the legacy school in the academic year prior to takeover that are enrolled in the takeover charter the following fall.

Table A3: 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	0.956	0.934	0.004	
					(0.012)	
N	14,478	11,303	1,052	746	3,650	
Has first exposure year outcomes	0.671	0.776	0.827	0.706	0.014	
					(0.021)	
N	14,478	11,303	1,052	746	3,650	
Has second exposure year outcomes	0.635	0.735	0.809	0.736	0.037	
					(0.027)	
N	9,509	7,617	795	439	2,499	
Has third exposure year outcomes	0.607	0.711	0.750	0.669	0.034	
					(0.034)	
N	5,817	4,670	540	287	1,572	
Has fourth exposure year outcomes	0.484	0.580	0.598	0.488	0.008	
					(0.049)	
N	2,545	2,088	254	162	828	
			B. UP grandfathering			
Has legacy grade outcomes	0.911	0.973	0.938	0.893	-0.018	0.003
					(0.014)	(0.011)
N	8,353	1,535	227	290	1,341	1,221
Has first exposure year outcomes	0.884	0.936	0.934	0.855	-0.034*	-0.022
					(0.018)	(0.017)
N	8,353	1,535	227	290	1,341	1,221
Has second exposure year outcomes	0.867	0.913	0.878	0.817	-0.074**	-0.063**
					(0.032)	(0.032)
N	7,813	1,416	131	164	708	638

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. Columns (5)-(6) include cell fixed effects as described in Tables 3 and 7. 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 A4: UP lottery records

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

Notes: This table summarizes the sample restrictions imposed for the UP lottery analysis.

Table A5: UP lottery attrition

	Sample means			Balance coefficients	
	Boston	Lottery applicants		Lottery applicants	
	students	6th grade	7th grade	Immediate offer	Waitlist offer
	(1)	(2)	(3)	(4)	(5)
Has first exposure year outcomes	0.914	0.924	0.843	-0.016	0.033
				(0.020)	(0.021)
N	5,922	911	51	962	962
Has second exposure year outcomes	0.880	0.872	0.784	-0.034	0.014
				(0.025)	(0.028)
N	5,922	911	51	962	962
Has third exposure year outcomes	0.831	0.814		-0.047	-0.038
				(0.039)	(0.044)
N	3,810	617		617	617

Notes: This table reports sample means and coefficients from regressions of the variable on 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 being offered a seat off the waitlist from March to the end of September. All regressions include lottery cohort dummies. The sample in columns (2)-(7) 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%