Charter School Closure and Student Achievement: Evidence from Ohio*

Deven Carlson
University of Oklahoma
Department of Political Science
455 W. Lindsey St., 205 Dale Hall Tower
Norman, OK 53706
decarlson@ou.edu

Stéphane Lavertu
The Ohio State University
John Glenn School of Public Affairs
1810 College Road, 110 Page Hall
Columbus, OH 43210
lavertu.1@osu.edu

June 2016

Abstract

The closure of low-performing schools is an essential feature of the charter school model. Our regression discontinuity analysis uses an exogenous source of variation in school closure—an Ohio law that requires charter schools to close if they fail to meet a specific performance standard—to estimate the causal effect of closure on student achievement. The results indicate that closing low-performing charter schools eventually yields achievement gains of around 0.2-0.3 standard deviations in reading and math for students attending these schools at the time they were identified for closure. The study also employs mandatory closure as an instrument for estimating the impact of exiting low-quality charter schools, thus providing plausible lower-bound estimates of charter school effectiveness. These results complement the more common lottery-based estimates of charter school effects, which likely serve as upper-bound estimates due to their focus on oversubscribed schools often located in cities with high-performing charter sectors. We discuss the implications for research and policy.

Keywords: charter schools, school closure, student achievement, regression discontinuity

JEL codes: H75; I21; I28

* The authors contributed equally and are listed in alphabetical order. We thank the Ohio Department of Education and the Ohio Education Research Center—particularly Jenny Gnagey, Josh Hawley, Lisa Neilson, Randy Olsen, and Jaron Shook—for providing data for this study. We thank Stephanie Moulton, Sally Sadoff, and attendees at the 2016 Annual Meeting of the Association for Education Finance and Policy for helpful comments on earlier drafts of this paper.

1. Introduction

Charter schools continue to proliferate. During the 2013-14 school year, there were over 6,400 charter schools serving over 2.5 million students nationwide—more than double the 3,000 charter schools in operation just a decade earlier (NAPCS, 2015). These publicly-funded schools, which operate under a contract (or "charter") that they develop in collaboration with a state-approved authorizing organization, enjoy greater operational independence from state and local regulation than traditional public schools.[1] However, the charter school model is designed to couple such autonomy with greater accountability for service provision. Charter schools must compete for students—and the public funding that accompanies them—in order to stay open and, thus, must meet the quality demands of parents (Hanushek et al., 2007). Additionally, authorizers or state regulators may hold charter schools formally accountable for educational outcomes. For example, 15 states now have laws requiring the automatic closure of charter schools that fail to meet minimum performance requirements (Ziebarth, 2015).

We exploit exogenous variation generated by Ohio's automatic closure law to identify the effect of charter school closure on the achievement of students attending these schools at the time they were identified for required closure. The specific metric providing the identifying variation is a school's score on Ohio's value-added "gain index." Conditional on failing to meet requirements on other performance metrics, a charter school is required to close if its gain index score in math or reading results in it being classified as having "below expected gains." We

[1] Charter schools were initially conceived in the early 1990s as a set of public schools that would be free from many of the rules and regulations governing traditional public schools. This autonomy was intended to promote innovations that could then be imported back to traditional public schools. In this conception, charter schools were intended as a complement, rather than an alternative, to traditional public schools. In practice, however, charter schools are often seen as competition for traditional public schools, and there is often considerable tension between the two sectors. Likely due in part to the increased autonomy of charter schools, the quality of these schools is highly heterogeneous. There are a number of very high-performing charter schools for which there is substantial excess demand. However, there are also some very low-quality charter schools. Recent work by CREDO (2013, 2015) reveals the significant variation in charter school quality across the country.

employ regression discontinuity (RD) techniques and individual-level data from over 6,000 students attending 36 charter schools at risk of closure on the basis of their gain index scores to estimate the achievement effects of closure. We estimate these effects up to three years after a school was informed that it would be required to shut down within one year.

The results indicate that requiring poor-performing schools to close has a positive effect on the achievement of their students. Three years after schools are identified for closure—and two years after schools are required to shut down—students from closing schools post reading and math scores that are typically between 0.2 and 0.3 standard deviations higher than those of students whose schools just avoid mandatory closure. The analysis also indicates that these gains are associated with displaced students ending up in higher-quality schools as measured by school value-added in math and reading. Finally, using mandatory closure as an instrument for student exit from our sample of charter schools, we show that exiting these low-performing schools leads to substantial achievement gains—often estimated to be in the range of 0.5 standard deviations, although these estimates are quite imprecise.

The analysis is relevant to several policy issues and related scholarly literatures. First and foremost, the analysis contributes to debates surrounding the optimal approach to charter school accountability. The initial charter school model assumed that oversight from authorizing organizations, together with the choice and competition inherent in the model, would ensure a charter sector with schools of consistently high quality. Although recent evidence indicates that this approach to accountability may generate quality improvements over the long term (Baude et al. 2014), charter schools in many states have been characterized by highly variable quality. Importantly, the state on which we focus, Ohio, has been singled out for its lack of charter school oversight and the poor performance of its charter sector (O'Donnell, 2015). Charter sectors such as Ohio's have contributed to a view that markets themselves may provide an inadequate mechanism for ensuring charter school quality and that state regulatory interventions may be necessary. Consistent with some other recent studies focused on private school markets (e.g. Witte et al. 2014; Carlson, Cowen, and Fleming 2014) our results indicate that state interventions in charter school markets can generate substantial achievement benefits.

Second, the analysis contributes to a body of research estimating the impact of charter school attendance on student achievement. The most convincing studies exploit admissions lotteries at oversubscribed schools to identify the effect of charter school attendance on student achievement (see Abdulkadiroglu et al., 2009, 2011; Angrist et al. 2010, 2012; Angrist, Pathak, and Walters 2013; Curto and Fryer 2014; Dobbie and Fryer 2011, 2013; Gleason et al., 2010; Hoxby, Muraka, and Kang, 2009). These studies often focus on high-performing charter sectors such as those in New York City or Boston (see Angrist, Pathak, and Walters 2013; CREDO, 2015) and find positive effects that are sometimes substantial in magnitude. For example, Abdulkadiroglu et al. (2011) found that attending Boston charter schools, as opposed to traditional public schools, resulted in annual achievement gains of up to 0.4 standard deviations.

Like admissions lotteries, the arbitrary performance requirement in Ohio's automatic closure law provides an exogenous source of variation in student attendance at a particular set of charter schools. Unlike lottery-based studies, however, the set of charter schools for which the closure law provides such variation consists of the lowest performing schools in a state with a relatively poor-performing charter sector.[2] Thus, whereas the lottery-based estimates likely represent an upper bound of charter school effectiveness, we generate among the first plausible

[2] Ohio has low-performing charter schools compared to Boston and New York. But it is worth noting that there are charter sectors that may be worse, such as those in Texas and Nevada (see CREDO, 2013; CREDO 2015). Nevertheless, the schools on which we focus—the worst performing schools in Ohio—are almost surely near the bottom of the national distribution.

lower-bound estimates of charter school effectiveness as measured by student achievement. Our results suggest that the negative impact of attending very poor-performing charter schools in Ohio is comparable in magnitude to the positive impact of attending oversubscribed charter schools in some other contexts (see Angrist, Pathak, and Walters, 2013).

Third, our analysis contributes to a small literature on school closure, which generally finds that closure ultimately has a negligible impact on the achievement of the students it displaces (e.g., see Brummet, 2014; de la Torre and Gwynne, 2009; Engberg et al., 2012; Young et al., 2009). These few studies, however, employ difference-in-differences designs that require arguably stronger assumptions to identify causal effects than our RD approach. In addition, these existing studies focus exclusively on the closure of traditional public schools. The impact of closure might differ in the charter sector, where closure is much more commonplace.[3] The impact of closure is also likely to differ in our study of Ohio's automatic closure law because closure decisions are formally tied to school "value added" quality metrics. Given that the closed schools in our sample are among the lowest performing in the state, the subsequent schooling options for students displaced by closure should be of relatively higher quality. This increases the likelihood that improved school quality can compensate for the negative effects of closure-induced student mobility (e.g., Brummet, 2014; Engberg et al., 2012). Indeed, the results of our analysis are consistent with the literature on the value-added measurement of school quality (e.g. Deming et al., 2014; Deming, 2014), as they suggest that the superior quality of students' new schools may explain the estimated educational benefits of closure.

The paper proceeds as follows. Section 2 provides background on Ohio charter schools and the state's automatic charter school closure law as it applied during the years of this study.

[3] Indeed, 200 charter schools closed in 2012-13 alone—about 3.4 percent of all charter schools in operation during that year (NAPCS, 2015).

Section 3 describes our research design. Section 4 describes our data and provides some descriptive analyses of trajectories in student achievement and school quality. Section 5 describes the RD analysis, including the running variable, the statistical models, and the results. Finally, we offer some concluding thoughts in Section 6.

2. Ohio Charter Schools

2.1 Overview

Ohio has approximately 400 charter schools that serve around 7 percent of public school students in the state. As in the rest of the country, these schools are publicly funded, non-sectarian, and enjoy more freedoms than traditional public schools when it comes to designing their curricula, managing their human resources, and developing a school environment. Compared to traditional public schools in Ohio, charter schools disproportionately serve minority, low-achieving, and impoverished students from urban communities (CREDO 2014).[4]

Historically, a defining feature of Ohio's charter sector was a lack of regulation. For many years there was little oversight of charter school authorizers, including few restrictions on the entities allowed to serve as authorizers and the number of schools a given organization could authorize.[5] During this period of lax regulation several cross-state studies revealed that Ohio's charter sector performed particularly poorly (CREDO 2009, 2013, 2014; Zimmer et al. 2009, 2012). Partially in response to this poor performance, state oversight strengthened. The state limited the number of charter schools that could be established, the Ohio Department of

[4] Ohio law allows new, "start-up" charter schools to open in any school district that the state has identified as "challenged," which includes the state's largest urban districts and those with poor performance on measures of student achievement. Additionally, any school district in the state can convert one of its traditional public schools to a charter school, but there are few such "conversion" charter schools.

[5] The entities permitted to authorize charter schools vary across states, but the two most common entities are local school boards and state departments of education (Carlson, Lavery, and Witte 2012). Ohio is one of a small number of states that also allow nonprofit organizations to authorize charter schools. Studies estimating the achievement effects of different authorizing entities in Ohio reach competing conclusions (CREDO 2014; Zimmer et al. 2014).

Education's (ODE) oversight role increased, and charter school authorizers are now held accountable for the educational outcomes of the schools they oversee. Perhaps the most controversial aspect of this increased oversight, however, was legislation requiring the automatic closure of charter schools that fail to meet minimum performance requirements.

2.2. Ohio's Automatic Charter School Closure Law

Ohio initially relied primarily on authorizer oversight and market forces—choice and competition—to regulate the charter sector. During this initial period, the state experienced a rapid proliferation of charter schools that were highly variable in quality, which sparked debate over whether the state's initial, hands-off approach to charter school accountability was sufficient to ensure a high-quality charter sector. One result of this debate was a 2005 Ohio law that required charter schools to close if their students failed to meet minimum academic performance standards. This automatic closure law applied to schools that were beyond their first two years of operation and that served the "general" student population.[6] Although the law required only a small number of schools to close while in effect from 2008 through 2014, it targeted the worst charter schools and was designed in a manner that provides leverage for estimating the impact of closure on student achievement.

The law's specific performance requirements have evolved over time and differ depending on whether a school serves grades K-3 (but no higher grade level), grades 4-8 (but not grades 10-12), or grades 10-12. Between the 2007-08 school year (SY2008) and SY2012—the period during which we observe automatic closures in our data—the performance requirements for schools serving grades K-3 and those serving grades 10-12 were based solely on the designations generated by Ohio's school accountability system, which annually assigned schools

---

[6] New schools and those serving primarily at-risk or disabled student populations were exempted from the provisions of the law.

one of the following six designations: "excellent with distinction," "excellent," "effective," "continuous improvement," "academic watch," and "academic emergency."[7] The provisions of the automatic closure law in place for SY2008 required charter schools serving grades K-3 or 10-12 to shut down if they had received the "academic emergency" designation in four consecutive school years. Modifications to the law that took effect in SY2009—and were in place through SY2012—required charter schools serving grades K-3 or 10-12 to shut down if they had received the "academic emergency" designation in three of the four most recent school years.

The automatic closure criteria for charter schools serving grades 4-8 incorporated both a school's accountability designation as well as its performance on Ohio's value-added measure of school quality. Specifically, at the end of SY2008, a charter school serving grades 4-8 was required to close within one year if it had received the "academic emergency" designation in three consecutive school years and also achieved "below expected gains" on Ohio's value-added "gain index" in two of the three most recent school years. From SY2009 through SY2012 the automatic closure requirement was triggered if a school received the "academic emergency" designation in two of the three most recent school years and also achieved "below expected gains" in two of the three most recent school years. Table 1 summarizes the closure criteria in place during the time period we study.

[Insert Table 1 about here]

[7] During the period of our study, the performance designations generated by Ohio's school accountability system were based on four primary metrics. The first was the federal "Adequate Yearly Progress" (AYP) metric required by No Child Left Behind. The second was its score on Ohio's "performance index," a continuous 0-120 scale measuring the school's overall student proficiency levels in math, reading, writing, science, and citizenship. The third input to the accountability system was a measure of the school's percent of quality indicators met, which measured student proficiency, graduation, and attendance rates. Based on their combined performance on these three metrics, schools were assigned a preliminary designation. This preliminary designation could then be adjusted by one unit based on the school's performance on Ohio's measure of school value-added, which was the fourth input into the accountability system. Specifically, a school's preliminary designation was increased by one level if the school achieved "above expected gains" and decreased by one unit if the school achieved "below expected gains" on the value-added gain index. This adjusted designation was the school's final accountability designation.

The value-added "gain index" is the effective source of identifying variation in our RD analysis and, thus, merits further description. Beginning in SY2007, ODE contracted with SAS to estimate Ohio public schools' contributions to student achievement gains in math and reading. SAS standardized all test scores by grade, year, and subject and then calculated annual student-level changes in these scores, so that estimates greater (less) than zero indicate that a school's students had greater (lower) achievement gains than the typical Ohio student.[8] The "gain index" for each school was calculated by dividing the value-added estimate by the standard error of that estimate. These gain index scores were then classified as "below expected gains," "met expected gains," or "above expected gains." Between SY2008 and SY2010, a gain index score below -1 was classified as "below expected gains" and a gain index score above 1 was classified as "above expected gains." For SY2011 and SY2012, the respective thresholds were set to -2 and 2.

Using the above criteria, the state began to identify schools for closure in the summer of 2008 based on performance through SY2008. The law permitted schools to remain in operation for one year after receiving notification of their required closure. For example, schools identified for closure in the summer of 2008 were required to close prior to SY2010, although they could close immediately if they chose. Between the summers of 2008 through 2012, the state notified 23 schools that they must close for failing to meet the law's minimum performance requirements. As described in further detail below, we focus on the impact of requiring the closure of 18 elementary and middle schools serving grades 4-8, as student achievement data are available for multiple consecutive years before and after closure of these schools.

[8] SAS accounted for multiple prior years of student test scores whenever possible. See SAS Institute Inc. (2012) for technical documentation. See Deming (2014) regarding the validity of such estimates of school quality.

3. Research Design

We use an RD design that exploits variation in charter school closure generated by the performance cutoff in Ohio's automatic closure law to perform three analyses. First, we estimate the impact of requiring a charter school to close on the academic achievement of students attending schools at the time they were identified for closure. This estimate can be interpreted as the impact of the automatic closure law. However, because of noncompliance stemming from the closure of charter schools that avoided mandatory closure, this estimate cannot be interpreted as the effect of closure *per se*. Thus, in our second analysis we estimate is the effect of closure itself. We estimate this parameter using an instrumental variable (IV) technique commonly used in RD designs with imperfect compliance. Finally, using a similar IV technique, we take advantage of the fact that the automatic closure law forces students to leave low-performing charter schools to estimate the effect of exiting low-quality charter schools.

3.1. Conceptual Model

The logic of the mandatory charter school closure law is that closing low-performing schools will lead to student achievement gains by increasing the quality of schools that students attend. In our context, school quality is measured in terms of student achievement growth in math and reading. Specifically, the analysis examines the impact of requiring charter schools to close if the year-to-year achievement gains of their students are significantly below the state average for multiple years. Research indicates that student achievement and attainment are increased by attending schools of higher quality, as measured by value-added performance (Deming, 2014; Deming et al., 2014). Thus, in our context, the closure of low-quality charter schools should lead displaced students to attend schools of higher quality as measured by school value-added, which, in turn, should lead to improvements in these students' math and reading achievement.[9]

[9] Although increased school quality is the mechanism through which closure is most commonly theorized to affect student outcomes (e.g. Brummet 2014), there are other potential mechanisms. In particular, closure could affect

Research also indicates, however, that student mobility can have a significant negative impact on student achievement (e.g., see Hanushek, Kain, and Rivkin, 2004; Booker et al., 2007; Xu, Hannaway, and D'Souza, 2009). Indeed, research that examines the achievement of students displaced by closure indicates that achievement declines temporarily in the years immediately following closure (Brummet 2014; de la Torre and Gwynne, 2009; Engberg et al. 2012; Young et al., 2009). The negative impact of mobility therefore might counteract the achievement benefits of attending a higher quality school. Our empirical analysis below estimates the net effect of closure-related influences on achievement.

3.2. Analytical Strategy

We employ an RD design to estimate the impact of closing low-quality schools, as measured by school value-added in math and reading. The validity of the design rests on the assumption that schools with value-added gain index scores that place them near but on either side of the value-added closure threshold should, on average, be similar in every way except whether or not they were required to close. To implement this design, we limit the analysis to a sample of schools that were at risk of closure based solely on the value-added metric. Specifically, we limit the analysis to charter schools serving grades 4-8 that met the "academic emergency" closure criterion, so that their value-added gain scores are the sole remaining determinant of closure.

Limiting the sample to schools that met the "academic emergency" closure criterion is necessary because the metrics underlying the state accountability designations exhibit signs of manipulation. In particular, schools near the "academic emergency" threshold were far more likely to receive the superior "academic watch" designation, which violates an assumption of the

student attitudes or behaviors in ways that affect their educational outcomes. We are unaware of studies that examine such responses to closure, and our data do not allow us to examine this possibility.

RD design.[10] The value-added gain index, on the other hand, is a metric that Ohio schools are likely unable to manipulate precisely (Kogan, Lavertu, and Peskowitz, 2016). Indeed, below we find little evidence contradicting the RD design's assumption that potential confounders are smooth across the closure threshold when the running variable is based on the school value-added metric.

The strength of our design is its internal validity. A potential weakness is that we focus exclusively on schools serving grades 4-8 that met the "academic emergency" closure criterion. Focusing on schools serving students in these grades and that have low proficiency rates obviously limits the generalizability of our results. But, in this case, the implications may be more significant. The closure law in Ohio was designed such that schools were aware if they were at risk of closure. Indeed, schools were notified if receiving an "academic emergency" or "below expected gains" designation during the upcoming school year would result in closure. If such awareness affected their performance trajectories—e.g., if staff and students in schools receiving the "academic emergency" designation for multiple years became demoralized and their performance declined, or if they were incentivized to improve their achievement—then the law's effect went beyond the impact of required closure. Our analysis does not capture such potential effects.

4. Data and Descriptive Statistics

4.1. Data

The analysis draws on both school-level and student-level data from ODE. The school-level data are publicly available on the ODE website. They include charter school characteristics, the list of

[10] For example, the first panel of Figure A1 in the appendix demonstrates a discontinuity in the density of charter schools on the state performance index near the cutoff determining an "Academic Watch" versus "Academic Emergency" designation. There are significantly more schools that receive the superior designation. Kogan, Lavertu, and Peskowitz (2016) found evidence of similar manipulation among traditional public schools across the state.

required closures for each of the five years we analyze, the school-level performance ratings that enable us to create yearly samples of "at risk" schools that met the closure criterion related to the "academic emergency" designation, and schools' value-added gain index scores, which we use to create the running variable that captures schools' proximity to the closure threshold.

We obtained annual, individual-level records for all students attending public schools in the state of Ohio between SY2006 and SY2013 from ODE via the Ohio Education Research Center. These data files include information on student demographic characteristics (such as student sex, race/ethnicity, economic disadvantage, and disability status), academic achievement, and identifiers for the schools and districts students attended. The achievement data consist of student scale scores on the Ohio Achievement Assessment (OAA), which is administered in grades 3-8. The analysis focuses on achievement in mathematics and reading because student test scores in these subjects are observed in consecutive grades across the entire panel.[11] To facilitate valid comparisons of student performance across grades and years, we standardized students' scale scores using the statewide mean and standard deviation by subject, grade, and year.

4.2. Analytic Sample

For each of five years from SY2008-SY2012—which we refer to as "focal" years (f) following Cellini, Ferreira, and Rothstein (2010)—we identified a set of "at risk schools" for which scores on the value-added gain index were the sole remaining determinant of mandatory closure.[12] For SY2008 this included all charter schools that received the "academic emergency" designation in the three previous years. For SY2009 through SY2012 this included schools that received the "academic emergency" designation in two of the three most recent years. This procedure resulted

[11] We limit the analysis to students who spent the majority of the school year in a particular school.
[12] We restricted the set of "at-risk schools" to those that failed the "academic emergency" closure criterion and had value-added estimates for at least two of the three prior years, which ensures that the school could fail on the value-added criterion.

in the identification of 55 “at risk” school-focal year combinations and 36 unique schools, as some schools met our definition of “at risk” in multiple years. Of these “at-risk” schools, 18 were required to close on the basis of their scores on the value-added gain index.

Based on this set of 55 “at risk” school-focal year combinations, we identified all students who attended an “at risk” school during the corresponding focal year and retrieved all of their available records between SY2006 and SY2013. For example, for schools at risk of required closure based on SY2009 focal year data, we have student data histories up to three years prior and that extend four years after SY2009. After creating SY2006-SY2013 student-level files for each of the five focal years (SY2008-SY2012), we stacked the data into a single SY2006-SY2013 file of student-level observations. We then limited the data file to student observations between two years prior to the relevant school-focal year combination and three years after that combination ($f-2$ through $f+3$). The final sample, in which we observe students over time within school-focal year combinations, consists of 21,416 observations from 6,027 unique students.

4.3. Descriptive Statistics

Table 2 presents descriptive statistics for students in our sample using focal-year observations of the covariates we employ in the analysis. The table illustrates that the students in our sample are low-achieving—their test scores place them approximately one standard deviation below the Ohio average—disproportionately African American, and overwhelmingly economically disadvantaged. It also reveals that the characteristics of students in schools required to close were similar to those of students whose schools were not required to close.

[Insert Table 2 about here.]

The RD analysis below examines whether student characteristics and pre-treatment achievement trends are comparable for students who, in the focal year, attended “at risk” schools on either side of the performance cutoff determining mandatory closure. However, because a school's being "at risk" could have affected the performance of its students, as well as the composition of the school itself, it is important to consider how the pre-treatment achievement trends of students in our sample compare to those of students that attended an "at risk" school but exited prior to the focal year. Table 3 presents the standardized achievement scores in math and reading of students in our sample who had attended their "at risk" school for two years prior to the focal year (stayers), students in our sample who entered an "at risk" school sometime over the previous two years and remained enrolled in the focal year (enterers), and students not in our sample who attended an "at risk" school in year f - 1 or f - 2 but left before the focal year (leavers). It also presents these scores separately for students attending schools that were and were not required to close.

[Insert Table 3 about here.]

First, it is notable that all of these groups have comparable average achievement levels two years prior to the focal year—around one standard deviation below the statewide mean on math and reading achievement tests. Second, it is notable that students who exit "at risk" schools exhibit larger gains than students who remain in, or enter, those schools. This suggests that significant achievement benefits may accrue from leaving the schools we identified as at risk of closure based on the value-added criterion. Third, we note that students who remain enrolled in schools required to close exhibit achievement losses while students who stay enrolled in schools that avoid required closure exhibit achievement gains. This pattern is consistent with a scenario where schools far from the closure cutoff—either above or below—may have a sense of their fate under the automatic closure law and respond accordingly. As such, it suggests the importance of the RD design we employ below.

To examine what happens in the period following the year a school was at risk of required closure, Table 4 presents—separately by required closure status—the percent of "at risk" schools that ended up closing, the percent of students attending "at risk" schools who later changed schools, and the percent of students in "at risk" schools who remained in charter schools. The table reveals that three of the 18 schools identified for closure shut down prior to the beginning of the following school year, whereas the other 15 schools identified for closure remained open for the one additional year the law allows. Among "at risk" schools not required to shut down, two closed prior to the beginning of the next school year and an additional two ceased operations in each of the next two years.

[Insert Table 4 about here.]

Table 4 also reveals that the student mobility rate was 36 percent among at-risk schools in the first post-focal year, whether or not schools were required to close. However, by the beginning of the second and third years after a school was at risk of closure identification, every student in a school identified for closure had switched schools while 62 and 75 percent, respectively, of students in schools allowed to remain open had done so.[13] Finally, Table 4 reveals that students in each class of "at risk" schools had a comparable likelihood of attending charter schools in subsequent years.[14]

Finally, Table 5 compares changes in both average student achievement and a value-added measure of the quality of schools students attended. It presents these changes separately for students attending schools that were and were not required to close, breaking down the sample of students attending schools not required to close into two sub-samples: 1) students who

[13] We calculated these mobility rates using students who have observations in both years necessary for the calculation. Students who age out of, or otherwise exit, the sample do not inform the calculations.
[14] These numbers are consistent with the post-closure schooling choices observed in a study of all charter school closures in Ohio, not just those at-risk under the automatic closure law (Carlson and Lavertu 2015).

nonetheless left by year $f + 2$ (the year by which all schools required to close would have shut down) and 2) students who remained in their “at risk” school in year $f + 2$.[15] The results demonstrate a large average increase in school quality—nearly one standard deviation—for students who attended a school required to close. Students who exited a school that avoided required closure also exhibit a meaningful increase in school quality, albeit one smaller in magnitude than the quality increase for students whose school was required to shut down. In contrast, students who remained enrolled in their focal-year school experienced little change in school quality by year $f + 2$.[16] These changes in school quality are consistent with changes in average student achievement over the same three-year period (reported in the middle and bottom panels of Table 5). In particular, the results demonstrate that students who exited their “at risk” school—either by choice or because it shut down—exhibited larger gains than students who remained enrolled in their “at risk” school through year $f + 2$.

[Insert Table 5 about here.]

This descriptive analysis suggests that Ohio’s mandatory charter school closure law had some achievement benefits for the students it displaced. Indeed, Tables 3 and 5 illustrate that students who exited their “at-risk” school exhibited larger achievement gains than students who

[15] To estimate school value-added we use annual records from all students in grades 3-8 (i.e. tested grades) attending a public school in Ohio from 2005-06 to 2012-13 and estimate the following model separately for each school year from 2006-07 to 2012-13 to obtain annual measures of school quality:

$$A_{ist} = R_{ist-1}\beta_1 + M_{ist-1}\beta_2 + X_{iskt}\tau + V_s\gamma + \varepsilon_{ist}$$

where A is the average standardized achievement in mathematics and reading for student i in school s in school year t. In this model, R_{ist-1} represents first- and second-order terms measuring a student’s prior reading achievement, M_{ist-1} represents an identical set of prior math achievement measures, X is a vector of observable student characteristics, V is a vector of school fixed effects, and ε is the error term. Covariates contained in the vector of observable background characteristics include indicators of whether or not a student is female, white, black, Asian, Hispanic, disabled, or economically disadvantaged, and whether or not a student changed schools while not in a terminal grade. After estimating this model we recover the coefficient estimates for the school fixed effects, which are parameterized using sum-to-zero constraints and thus estimated relative to the average school (see Mihaly et al. 2010), and then standardize the estimates such that the distribution of school value-added for each year is $N(0,1)$.

[16] Table A1 in the appendix illustrates how a within-student analysis of changes in school quality yields similar results. Those who stay in at-risk schools experience no significant post-focal year improvements in school quality, whereas those who exit experience substantial gains in school quality.

remained enrolled in those schools. Additionally, Table 5 shows that students who left their "at risk" schools on average ended up in schools of significantly higher quality as measured by student achievement growth, whereas those who stayed did not experience changes in school quality. Finally, Table 4 points to the importance of accounting for non-compliance—both in terms of schools that shut down without the state requiring it, as well as students who exit poor-performing "at risk" schools even though their schools had not closed.

5. Regression Discontinuity Analysis

We employ an RD design to estimate the impact of closing low-quality schools as measured by school value-added in math and reading. In this section, we first describe how we constructed a running variable that captures the "at risk" schools' distances from the closure requirement based on the value-added gain index score (section 5.1). We then test the RD assumptions that the density of this running variable is smooth across the closure cutoff and that our covariates are balanced at the closure threshold (sections 5.1 - 5.2). Finally, we describe the statistical models and review the results of analyses estimating the impact of identifying schools for mandatory closure (section 5.3), the impact of closure *per se* (section 5.4), and the impact of student exit from our sample of "at risk" charter schools (section 5.5).

5.1. Running Variable

The "at risk" schools in our sample were required to close if their value-added gain index scores generated a classification of "below expected gains" in two of the three most recent school years. Constructing the running variable for the 55 "at risk" school-focal year combinations involved collecting—separately for math and reading—schools' scores on the value-added gain index from the focal year, as well as from each of the two prior years. Then, for each of the three years, we identified the lower of the schools' two gain index scores (in math or reading) and calculated the difference between that score and the relevant threshold for being classified as achieving "below expected gains."[17] We then identified the second largest value of the calculated differences, as this value is pivotal for determining closure. A negative value indicates that a school failed to score above the closure threshold in at least two years, whereas a positive value indicates that the academic gains were sufficiently high to avoid closure. Finally, we multiplied this middle value by -1 so that a positive number indicates a school was required to close due to poor performance.[18]

The validity of our RD design could be threatened if schools were able to precisely manipulate the value-added gain index score that serves as the basis of the running variable (Lee and Lemieux, 2010). Such manipulation is highly unlikely given the complexity of the value-added gain index. Schools would need to be able to manipulate precisely their estimated value-added in both reading and math, as well as the standard errors of those estimates. Additionally, as we note above, research has found no evidence of manipulation of this value-added metric for traditional public schools in Ohio (see Kogan, Lavertu, and Peskowitz, 2016). We nonetheless look for evidence of such manipulation by looking for a discontinuity in the density of the running variable at the cutoff. The histogram in Figure 1 illustrates that "at risk" schools avoided mandatory closure more often than not, but there are cases very close to the closure threshold on each side of the cutoff.

[Insert Figure 1 about here.]

[17] Schools can be classified as achieving "below expected gains" on the basis of either its math or reading gain index score. Failure in both subjects is not necessary. Between SY2008 and SY2010, a gain index score below -1 indicated "below expected gains," whereas the SY2011 and SY2012 threshold was set to -2.
[18] The running variable measures how close a school came to required closure during its at-risk period. In sensitivity analyses we assess the robustness of our results to an alternative construction of the running variable that measures a school's proximity to the automatic closure threshold in the focal year, rather than the full at-risk period.

This visual evidence is corroborated by the results of McCrary's (2008) statistical test, which fails to reject the null hypothesis of no difference in the density on either side of the cutoff (difference of -0.505 and a standard error of 0.491). Additionally, to test for manipulation in the metric underlying the running variable, we conducted a density test of the value-added gain index across all Ohio charter schools and obtained similar results (difference of -0.143 with a standard error of 0.174).[19] Although the density tests yield results that do not approach conventional levels of statistical significance, there is some suggestive visual evidence that, near the cutoff, schools are somewhat less likely to qualify for closure.

5.2. Testing for Covariate Balance

We examine this issue further by testing whether students attending schools just above the closure cutoff are systematically different from students attending schools just below the cutoff. If such differences exist, then any estimated effect of closure on student outcomes could be attributable to differential student composition across schools, rather than the policy treatment.

We checked for differences in observable student characteristics across the closure cutoff in the focal year—the year treatment status is determined. Specifically, we estimated

$$O_{ij} = f(R_j) + \tau C_j + \varepsilon_{ij} \qquad (1)$$

where *O* represents an observable student characteristic for student *i* in school-focal year combination j, $f(R_j)$ is a flexible function of a school's distance from the closure threshold (i.e. the running variable), *C* is an indicator for scoring above the threshold, and ε is the error term. We employ two specifications of $f(R_j)$. The first includes a linear term interacted with the closure indicator *C*, and the second includes linear and quadratic terms, both of which are

[19] Figure A1 in the appendix provides visual evidence of the results of these tests.

interacted with the closure indicator. In both cases we clustered standard errors by focal-year school.

[Insert Table 6 about here]

The results in Table 6 reveal no imbalances in pre-treatment student characteristics. Across the 11 characteristics we examine, there are no differences with a *p*-value less than 0.10. Perhaps most importantly, the estimated differences in the focal-year test scores of students in schools above and below the closure threshold do not approach statistical significance. Finally, for each specification we estimated a Seemingly Unrelated Regression and conducted a Chi-squared test of the hypothesis that the estimated coefficients for the indicator of scoring above the closure threshold are jointly equal to zero across the 11 regressions (see, for example, Chin, Daysal, and Imberman 2013). These tests are unable to reject the null hypothesis that the coefficients are jointly equal to zero.[20]

5.3. Estimating the Impact of Identifying Schools for Mandatory Closure

We estimated the effect of identifying charter schools for mandatory closure on achievement by estimating a series of OLS models, one for each relative focal year from two years prior to the focal year ($f-2$) to three years after the focal year ($f+3$). In other words, we estimated six separate regressions—one for each relative year $f-2$ through $f+3$. The general model specification for each relative focal year is

$$Y_{ij} = f(R_j) + \tau C_j + A_j\delta + \varepsilon_{ij} \tag{2}$$

where Y is student achievement in reading or math for student i in school-focal year combination j, $f(R_j)$ is a flexible function of a school's distance from the closure threshold (i.e. the running

[20] The *p*-value for the test that the estimated coefficients for the indicator of scoring above the closure threshold across the 11 regressions are jointly equal to zero is 0.4490 for the linear specification. The *p*-value is 0.1977 when the running variable is specified as linear and quadratic, with each term allowed to have different slopes on each side of the threshold.

variable), C is an indicator for scoring above the closure threshold and thus being required to shut down, A is a vector of school-year fixed effects, and ε is an error term.

Our parameter of interest, τ, is the effect of required closure on student achievement—the estimates of which will be unbiased as long as $f(R_j)$ is properly specified. We specify $f(R_j)$ three different ways in the results presented below: 1) a linear term interacted with the closure indicator, 2) linear and quadratic terms, and 3) linear and quadratic terms interacted with the closure indicator.[21] Standard errors are clustered by focal-year school in all models.

[Insert Table 7 about here.]

The results appear in Table 7. Consistent with the assumptions of the RD design, the table reveals that there are no statistically significant differences in pre-treatment achievement between students who attended schools that were and were not required to close. Post-treatment, the reading results indicate positive but statistically insignificant estimates of the effect of closure that are between 0.092 and 0.152 standard deviations beginning two years after the focal year—the year by which schools identified for closure must shut down. The following year reveals estimated reading achievement benefits of 0.103, 0.206, and 0.255 standard deviations across the three respective specifications, with the latter two estimates reaching conventional levels of statistical significance. Third-year effects in math range from 0.051 to 0.236 standard deviations but only one estimate reaches conventional levels of statistical significance.

We sought to increase the precision of our estimates by estimating a variant of equation (2) that includes a covariate capturing students' focal-year achievement scores in the same subject as the outcome variable. Table 8 presents the results of this analysis. The results are generally similar, although the anomalous third-year math achievement gain of 0.051 from Table

[21] Gelman and Imbens (2014) recommend against including higher-order polynomials (i.e. third or higher) in regression discontinuity analyses.

7 increases to 0.193 standard deviations and now reaches conventional levels of statistical significance. But the results in Table 8 are generally similar to those in Table 7. We provide graphical illustrations of these results in the appendix (Figures A2 and A3).

[Insert Table 8 about here.]

The results presented in Tables 7 and 8 are robust to and oftentimes stronger under a variety of alternative analytic choices, including estimating the effects across all relative focal years simultaneously using the RD panel technique Cellini, Ferreira, and Rothstein (2010) propose (see Tables A2 and A3 in the appendix); restricting our analysis to only the first time a school was "at risk" of closure, so that each school enters the analysis only once (see Table A3 in the appendix); omitting three schools that qualified for closure based on the value-added closure criterion in the year prior to the focal year (i.e. we exclude schools for which "academic emergency" was the pivotal criterion in the focal year—see Table A4 in the appendix); specifying the running variable such that it measures an "at risk" school's proximity to the automatic closure threshold in the focal year (see Table A5 in the appendix), rather than the proximity to the automatic closure cutoff over the full "at risk" period; and limiting the sample to schools that were at risk in SY2008-SY2010 (see Table A6 in the appendix), which is the period when a gain index score of -1 (instead of -2 as in SY2011-SY2012) corresponded to a rating of "below expected gains".

Finally, Table A7 in the appendix presents results from estimating a variant of equation (2) where the dependent variable is the value-added score of the school that a student attends. The results indicate that closure identification has little effect on the quality of students' schools in year $f + 1$, but there are positive effects in years $f + 2$ and $f + 3$. These results provide evidence that a closure-induced increase in school quality is a likely mechanism generating the positive achievement effects presented in Tables 7 and 8.

5.4. Estimating the Impact of Closure

The models above estimate the impact of identifying schools for closure, but they do not capture the impact of closure *per se* on students displaced by it. As Table 4 illustrates, several "at risk" schools that did not qualify for mandatory closure nonetheless shut down. To estimate the effect of closure on students displaced by it, we employ an instrumental variables (IV) approach commonly used with RD designs with imperfect compliance. Specifically, we use scoring above the threshold for mandatory closure as an instrument for closure. Scoring above the closure threshold is a valid instrument for closure if it 1) predicts closure and 2) is uncorrelated with the outcomes of interest other than through its effect on closure. Scoring above the closure threshold clearly predicts closure, satisfying the first condition. The second condition—the exclusion restriction—is not directly testable, but will be met in an RD context if the function of the running variable is properly specified and the running variable was not manipulated.

We implemented this IV approach via a series of two-stage least squares (2SLS) models in which the first stage predicts closure using the following model:

$$L_{ij} = f(R_j) + \tau C_j + A_j\delta + \omega_{ij} \quad (3)$$

where L is an indicator that student *i* attended "at risk" school *j* that closed by year $f + 2$, C is an indicator for scoring above the closure threshold, A is a vector of school-year fixed effects, and ω is the error term. The predicted values of L resulting from the estimation of equation (2)—denoted as $\hat{L}$ below—are then inserted into the second-stage equation, taking the place of the indicator for scoring above the closure threshold from the reduced-form model above. The second-stage model can be written as:

$$Y_{ij} = f(R_j) + \tau\hat{L}_j + A_j\delta + \varepsilon_{ij} \quad (4)$$

Because $\hat{L}$ contains only the variation in closure attributable to scoring above the specified cutoff, it is uncorrelated with ε and the resulting estimate of λ thus represents the local average treatment effect (LATE) of closure on student achievement. As the first-stage model predicts closure by year $f + 2$, we only estimate the 2SLS models for relative years $f + 2$ and $f + 3$.

The results appear in Table 9 and, as expected, are somewhat larger in magnitude than the estimated effects of closure identification we presented in Table 8. The estimates in Table 9 are effectively the estimates in Table 8 rescaled by the proportion of schools that "complied" with their treatment status under the mandatory closure law. Closure has positive effects on reading achievement beginning the first year by which schools were required to close (i.e. year $f + 2$). The estimates range from 0.174 to 0.244 standard deviations in magnitude and all reach conventional levels of statistical significance. Once again, significant positive effects in math emerge in the third year, with effect sizes ranging from 0.122 to 0.233 standard deviations. Our most flexible specifications—those employing a quadratic polynomial allowed to vary on either side of the cutoff (columns 4 and 8 of Table 9)—yield statistically significant effects of 0.232 standard deviations in math and 0.280 standard deviations in reading.

[Insert Table 9 about here.]

5.5. Estimating the Impact of Exiting Low-Quality Charter Schools

Finally, as Table 4 indicates, there are students who leave "at risk" schools that did not close. To estimate the impact of students exiting low-quality, "at-risk" charter schools, we re-estimated equation (3) such that the dependent variable is an indicator that a student exited an "at risk" school by year $f + 2$, and we then used that predicted value in equation (4) to estimate the impact of exiting an "at risk" school. The results, presented in Table 10, are significantly more pronounced, although the estimates are quite imprecise and should be interpreted cautiously. With these cautions in mind, the statistically significant, second-year achievement gains in reading range from 0.306 to 0.480 standard deviations—effect sizes comparable in magnitude to some lottery-based estimates of charter school effectiveness in high-performing urban charter sectors. The estimates for two years after mandatory closure vary across specifications, but all are large and positive. They range from 0.330 to 0.764 standard deviations in math and 0.331 to 0.897 in reading—with only one estimate for math (of 0.704 standard deviations) and two for reading (of 0.733 and 0.897 standard deviations) having *p*-values below 0.10.

[Insert Table 10 about here]

6. Discussion and Conclusion

We estimated the effect of Ohio's automatic charter school closure law on the achievement of students attending schools at the time they were identified for mandatory closure. The results of our RD analysis, which we restricted to low-achieving schools for which the value-added metric was pivotal for the closure decision, reveal that students whose schools were identified for mandatory closure experienced average achievement gains between 0.2 and 0.3 standard deviations by the third year after closure identification. The estimates are robust to multiple modeling approaches, multiple specifications of the running variable, and restricting the sample to observations from the first time a school was at risk of closure. These results have several implications for research and policy.

At a basic level, our results provide evidence that automatic closure laws based on value-added estimates of school quality are a viable approach for improving the achievement outcomes of students attending very low-performing charter schools. We provide evidence that these achievement increases may be attributable to students moving to higher quality schools. Such findings are consistent with recent research examining the closure of traditional public schools, which finds that students attending low-achieving schools experience sharp gains in test scores after their schools shut down (Brummet 2014). This suggests that policymakers considering automatic closure laws should ensure that they focus on the very worst schools in terms of student achievement gains. Additionally, the effectiveness of these laws will likely be enhanced if students are encouraged—either directly or via information provision—to transfer to schools of sufficiently greater quality to compensate for disturbances associated with student mobility.

This study also contributes to the literature on the achievement effects of charter school attendance. The most convincing studies exploit exogenous variation in charter school attendance generated by admissions lotteries to estimate the effect of attending oversubscribed charter schools (see Angrist, Pathak, and Walters, 2013). Many of these studies document positive effects of attending such schools, with the effects occasionally large in magnitude. The reality of the charter school sector, however, is that there is substantial heterogeneity in school quality. Because low-quality schools are seldom over-subscribed, there is little research that estimates the effect of attending such schools using a convincing identification strategy. Our study provides such evidence by taking advantage of exogenous variation in attendance at low-quality charter schools generated by Ohio's automatic closure law. Consequently, this study provides a plausible lower-bound estimate of charter school effectiveness that complements the upper-bound estimates generated by lottery-based studies. The results indicate that attending low-quality charter schools in Ohio has negative effects that are comparable in magnitude to the positive effects of attending over-subscribed schools in high-performing charter sectors.

Finally, our results also have implications for charter school accountability more generally. The original charter model held that oversight from authorizing organizations, coupled with market forces, would provide sufficient levels of accountability. Recent evidence indicates that this approach may generate quality improvements over the long term (Baude et al. 2014), but states' experiences demonstrate that some very low-quality schools will persist. Our analysis complements other studies indicating that government intervention in educational markets can improve student achievement (Witte et al. 2014; Carlson, Cowen, and Fleming 2014).

Of course, such intervention comes with potential costs. In the case of automatic closure laws, the state is restricting the educational market by removing selected schools from a family's choice set. Such removals could be problematic if a school identified for closure excels on other dimensions—such as art or music instruction, extracurricular activities, and safety—that are highly valued by families but unrelated to achievement gains in reading and math. Yet, even if value-added measures of school quality do not account for all dimensions of quality important to parents, policymakers may determine that requiring school closure on the basis of value-added is warranted, given that these metrics correlate with superior labor-market outcomes and economic growth (e.g., see Chetty, Friedman, and Rockoff 2014; Hanushek and Kimko 2000; Hanushek 2011). Put differently, they may conclude that the achievement-related benefits that accrue to students and society from closing ineffective charter schools outweigh parental preferences. This tension between parental autonomy and government regulation of the educational market has long lurked below the surface of debates over school choice policy. And although it is ultimately an issue that must be resolved through the political process, this study suggests that automatic closure laws can potentially serve as one component of a larger portfolio of policies designed to ensure charter school quality.

7. References

Abdulkadiroglu, A., Angrist, J., Cohodes, S., Dynarski, S., Fullerton, J. Kane, T., and Pathak, P. A. (2009). *Informing the Debate: Comparing Boston's Charter, Pilot, and Traditional Schools*. The Boston Foundation (TBF). Boston, January.

Abdulkadiroglu, A., Angrist, J., Dynarski, S., Fullerton, J. Kane, T., and Pathak, P. A. (2011). Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics* 126 (2): 699–748.

Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A., and Walters, C.R. (2010). Inputs and Impacts in Charter Schools: KIPP Lynn. *American Economic Review*, 100, 239–43.

Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A., and Walters, C.R. (2012). Who Benefits from KIPP? *Journal of Policy Analysis and Management*, 31, 837–60.

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

Baude, P., Casey, M., Hanushek, E.A., & Rivkin, S.G. (2014). The evolution of charter school quality. *National Bureau of Economic Research Working Paper 20645.*

Booker, K., Gilpatric, S.M., Gronberg, T., & Jansen, D. (2007). The impact of charter school attendance on student performance. *Journal of Public Economics*, 91, 849–876.

Brummet, Q. (2014). The effect of school closings on student achievement. *Journal of Public Economics*, 119, 108-124.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015). Optimal Data-Driven Regression Discontinuity Plots. *Journal of the American Statistical Association*, 110, 1753-1769.

Carlson, D.E., Cowen, J.M., & Fleming, D.J. (2014). Third-party governance and performance measurement: A case study publicly funded private school vouchers. *Journal of Public Administration Research and Theory*, 24, 897-922.

Carlson, D., & Lavertu, S. (2015). *School closures and student achievement: An analysis of Ohio's urban district and charter schools.* Washington, DC: Thomas B. Fordham Institute.

Carlson, D., Lavery, L., and Witte, J.F. (2012). Charter school authorizers and student achievement. *Economics of Education Review*, 31, 254-267.

Cellini, S.R., Ferreira, F., & Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *Quarterly Journal of Economics*, 125, 215-261.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, *104*, 2633-79.

Chin, A., Daysal, N.M., & Imberman, S.A. (2013). Impact of bilingual education programs on limited English proficient students and their peers: Regression discontinuity evidence from Texas. *Journal of Public Economics*, 107, 63-78.

Center for Research on Education Outcomes (CREDO). (2009). *Multiple choice: Charter school performance in 16 states*. Stanford, CA.

Center for Research on Education Outcomes (CREDO). (2013). *National charter school study: 2013*. Stanford, CA.

Center for Research on Education Outcomes (CREDO). (2014). *Charter school performance in Ohio*. Stanford, CA.

Center for Research on Education Outcomes (CREDO). (2015). *Urban charter school study: Report on 41 regions*. Stanford, CA.

Curto, V.E. & Fryer, Jr., R.G. (2014). The Potential of Urban Boarding Schools: Evidence from SEE. *Journal of Labor Economics*, 32, 65-93.

de la Torre, M. & Gwynne, J. (2009). *When schools close: Effects on displaced students in Chicago Public Schools*. Chicago, IL: Consortium on Chicago School Research.

Deming, D.J., Hastings, J., Kane, T., & Staiger, D. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review*, 104, 991-1013.

Deming, D.J. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review: Papers & Proceedings*, 104: 406-411.

DePaoli , J. & van Lier, P. (2013). *Avoiding accountability: How charter operators evade Ohio's automatic closure law*. Columbus, OH: Policy Matters Ohio.

Dobbie, W., & Fryer, Jr., R.G. (2011). Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 3, 158–87.

Dobbie, W., & Fryer, Jr., R.G. (2013). *The Medium-Term Impacts of High-Achieving Charter Schools on Non-Test Score Outcomes*. (No. w19581). National Bureau of Economic Research.

Engberg, J., Gill, B., Zamarro, G., & Zimmer, R. (2012). Closing schools in a shrinking district: Do student outcomes depend on which schools are closed? *Journal of Urban Economics*, 71, 189–203.

Gelman, A., & Imbens, G. (2014). *Why high-order polynomials should not be used in regression discontinuity designs* (No. w20405). National Bureau of Economic Research.

Gleason, P., Clark, M., Tuttle, C., and Dwoyer, E. (2010). *The Evaluation of Charter School Impacts: Final Report*. U.S. Department of Education National Center for Education Evaluation and Regional Assistance (NCEE) 2010-4029. Washington, DC, June.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, 88, 1721–1746.

Hanushek, E.A., Kain, J.F., Rivkin, S.G., & Branch, G.F. (2007). Charter school quality and parental decision making with school choice. *Journal of Public Economics*, 91, 823-848.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American economic review*, 90, 1184-1208.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review* 30(3): 466-479.

Hoxby, C., Murarka, S., & Kang, J. (2009). *How New York City's Charter Schools Affect Achievement*. New York City Charter Schools Evaluation Project. Cambridge, MA, February.

Kogan, V., Lavertu, S., & Peskowitz, Z. (2016). Do school report cards produce accountability through the ballot box? *Journal of Policy Analysis & Management*, 35(3): 639-661.

Lee, D.S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48, 281-355.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142, 698-714.

Mihaly, K., McCaffrey, D.F., Lockwood, J.R., & Sass, T.R. (2010). Centering and reference groups for estimates of fixed effects: Modifications to felsdvreg. *The Stata Journal*, 10, 82-103.

National Alliance for Public Charter Schools (NAPCS). (2015). The Public Charter Schools Dashboard. Accessed December 17, 2015 from http://dashboard.publiccharters.org/dashboard/home

O'Donnell, P. (2015). "Ohio's Charter Schools Ridiculed at National Conference, Even by National Supporters" *The Plain Dealer* Accessed December 16, 2015 from http://www.cleveland.com/metro/index.ssf/2015/03/ohios_charter_schools_ridicule.html

SAS Institute Inc. (2012). *Technical Description: Multi-Year and Composite Calculations*. Retrieved August 19, 2014 from https://education.ohio.gov/getattachment/Topics/School-Improvement/Diagnostic-Review/Value-Added-Tools-Resources/Multi_year_Composite_Calculations_20121001-1.pdf.aspx

Witte, J.F., Wolf, P.J., Cowen, J.M., Carlson, D.E., & Fleming, D.J. (2014). High Stakes Choice: Achievement and Accountability in the Nation's Oldest Urban Voucher Program. *Educational Evaluation and Policy Analysis*, 36, 437-456

Xu, Z., Hannaway, J., & D'Souza, S. (2009). Student transience in North Carolina: the effect of mobility on student outcomes using longitudinal data. CALDER Working Paper 37. Washington DC: Urban Institute.

Young, V.M., Humphrey, D.C., Wang, H., Bosetti, K.R., Cassidy, L., Wechsler, M.E., Rivera, M., Murray, S., & Schanzenbach, D.W. (2009). *Renaissance schools fund-supported schools: Early outcomes, challenges, and opportunities*. Menlo Park, CA: SRI International.

Ziebarth, T. (2015). Automatic Closure of Low-Performing Public Charter Schools. *National Alliance for Public Charter Schools* Accessed December 16, 2015 from "http://www.publiccharters.org/wp-content/uploads/2015/10/policy_snapshot_autoclosure_v2.pdf"

Zimmer, R, Gill, B., Booker, K., Lavertu, S., Sass, T., & Witte, J. (2009). *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition*. RAND: MG-869

Zimmer, R, Gill, B., Booker, K., Lavertu, S., & Witte, J. (2012). Examining charter school achievement in seven states. *Economics of Education Review*, 31, 213- 224.

Zimmer, R., Gill, B., Attridge, J., & Obenauf, K. (2014). Charter school authorizers and student achievement. *Education Finance and Policy*, 9, 59-85.

8. Tables and Figures

Table 1. Automatic charter school closure provisions, by school year and grade range

	Year				
Grade Range	**SY2008**	**SY2009**	**SY2010**	**SY2011**	**SY2012**
K-3	"Academic Emergency" designation for four consecutive years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years
4-8	"Academic Emergency" designation for three consecutive years AND "Below Expected Gains" in two of three most recent years (below -1 on the Gain Index)	"Academic Emergency" designation in two of three most recent years AND "Below Expected Gains" in two of three most recent years (below -1 on the Gain Index)	"Academic Emergency" designation in two of three most recent years AND "Below Expected Gains" in two of three most recent years (below -1 on the Gain Index)	"Academic Emergency" designation in two of three most recent years AND "Below Expected Gains" in two of three most recent years (below -2 on the Gain Index)	"Academic Emergency" designation in two of three most recent years AND "Below Expected Gains" in two of three most recent years (below -2 on the Gain Index)
10-12	"Academic Emergency" designation for four consecutive years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years	"Academic Emergency" designation in three of four most recent years

Table 2. Student characteristics in focal year, by closure status

Characteristic	N	Mean	Min	Max
	All Students			
Math Score (z-score)	6,016	-0.973	-6.232	3.459
Reading Score (z-score)	6,019	-0.905	-6.859	2.601
White (proportion)	6,027	0.068	0	1
Black (proportion)	6,027	0.829	0	1
Hispanic (proportion)	6,027	0.056	0	1
Female (proportion)	6,027	0.503	0	1
Disability (proportion)	6,027	0.142	0	1
Economically Disadvantaged (proportion)	6,027	0.870	0	1
English Language Learner (proportion)	6,027	0.019	0	1
	Closed			
Math Score (z-score)	2,526	-1.022	-6.011	3.459
Reading Score (z-score)	2,524	-0.909	-6.126	2.601
White (proportion)	2,526	0.052	0	1
Black (proportion)	2,526	0.891	0	1
Hispanic (proportion)	2,526	0.017	0	1
Female (proportion)	2,526	0.524	0	1
Disability (proportion)	2,526	0.157	0	1
Economically Disadvantaged (proportion)	2,526	0.876	0	1
English Language Learner (proportion)	2,526	0.000	0	1
	Not Closed			
Math Score (z-score)	3,490	-0.937	-6.232	2.011
Reading Score (z-score)	3,495	-0.901	-6.859	2.212
White (proportion)	3,501	0.080	0	1
Black (proportion)	3,501	0.783	0	1
Hispanic (proportion)	3,501	0.084	0	1
Female (proportion)	3,501	0.488	0	1
Disability (proportion)	3,501	0.132	0	1
Economically Disadvantaged (proportion)	3,501	0.866	0	1
English Language Learner (proportion)	3,501	0.032	0	1

Table 3. Pre-Treatment Achievement Among Those Who Stay, Enter, and Leave "At Risk" Charter Schools Prior to their "At Risk"/"Focal" Year (Standard Deviation Units)

	2 years prior (f - 2)	1 year prior (f - 1)	Focal year (f)	*3-year Difference*
		All Students		
Math Achievement				
Stayers	-1.001	-0.976	-0.969	*0.031*
Enterers	-0.976	-0.988	-1.003	*-0.027*
Leavers (not in sample)	-1.034	-0.971	-0.802	*0.232*
Reading Achievement				
Stayers	-0.922	-0.935	-0.898	*0.024*
Enterers	-0.948	-0.950	-1.003	*-0.055*
Leavers (not in sample)	-0.984	-0.920	-0.802	*0.183*
		Students in Schools Required to Close		
Math Achievement				
Stayers	-0.859	-0.898	-0.993	*-0.134*
Enterers	-1.013	-0.993	-1.099	*-0.086*
Leavers (not in sample)	-0.986	-0.953	-0.831	*0.155*
Reading Achievement				
Stayers	-0.758	-0.839	-0.884	*-0.126*
Enterers	-0.935	-0.926	-0.980	*-0.045*
Leavers (not in sample)	-0.862	-0.854	-0.748	*0.114*
		Students in Schools Not Required to Close		
Math Achievement				
Stayers	-1.166	-1.052	-0.950	*0.216*
Enterers	-0.953	-0.984	-0.940	*0.013*
Leavers (not in sample)	-1.063	-0.982	-0.786	*0.278*
Reading Achievement				
Stayers	-1.114	-1.029	-0.910	*0.205*
Enterers	-0.956	-0.968	-0.912	*0.044*
Leavers (not in sample)	-1.060	-0.960	-0.746	*0.314*

NOTE. Average standardized achievement levels are presented for three groups of students in the years leading up to and including the "at risk" focal year. Stayers are students who were in a school prior to the "at risk" focal year and stayed through the "at risk" focal year. Enterers are those who were observed in the "at risk" school during the focal year but who were observed in a non-at risk school in at least one of the prior two years. "Leavers" are students who attended "at risk" schools in prior years but who left before the "at risk" focal year. Stayers and enterers are in the analytic sample for the RD analysis, but leavers are not. This table omits students we do not observe in the dataset during the "at risk" focal year.

Table 4. School closure, student mobility, and charter school attendance by required closure status

	Schools Not Required to Close	Schools Required to Close
Percent of schools closed (N)		
By beginning of *f*+1 school year	5.4 (2/37)	16.7 (3/18)
By beginning of *f*+2 school year	10.8 (4/37)	100.0 (18/18)
By beginning of *f*+3 school year	16.2 (6/37)	100.0 (18/18)
Percent of students that had changed schools		
By beginning of *f*+1 school year	35.7	36.3
By beginning of *f*+2 school year	61.5	100
By beginning of *f*+3 school year	74.7	100
Percent of students attending a charter school		
f+1 school year	78.4	78.9
f+2 school year	60.7	55.2
f+3 school year	52.9	54.0

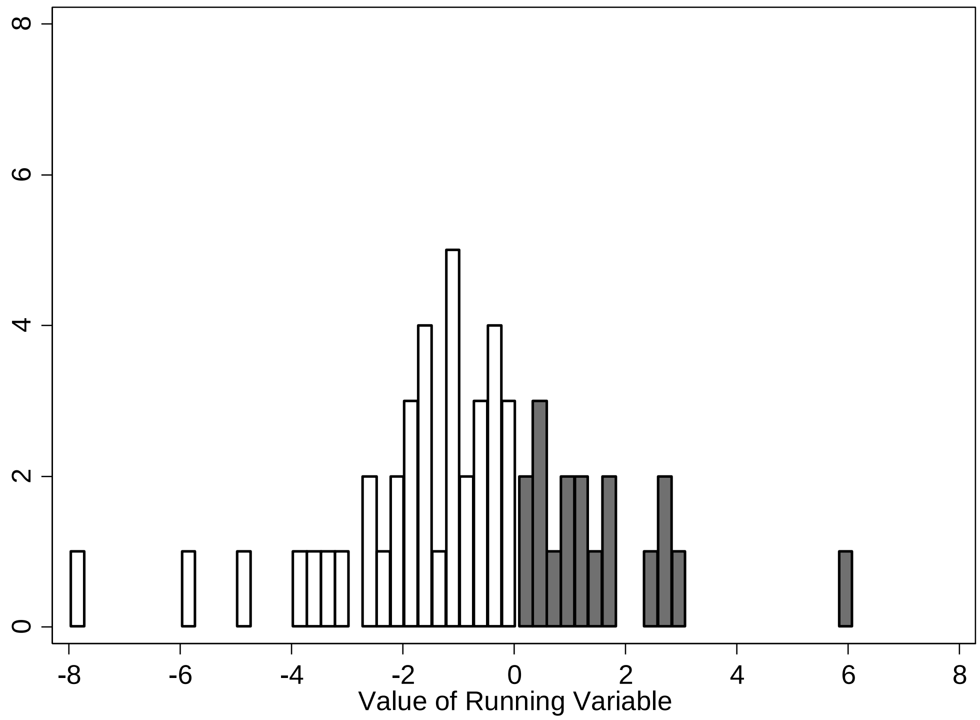
NOTE. Statistics in the middle and bottom panels of the table were calculated using students who have observations in both years necessary for the calculation. Students who age out of the sample do not inform the calculations.

Table 5. Mean School Quality for Students Whose "At Risk" Schools Were and Were not Required to Close (Standard Deviation Units).

	Focal year (f)	1 year after (f + 1)	2 years after (f + 2)	*3-year Difference*
	Average School Value-Added			
Schools Required to Close				
Students who exit by f + 2	-1.672	-1.021	-0.73	*0.942*
Schools Not Required to Close				
Students who exit by f + 2	-1.011	-0.466	-0.425	*0.585*
Students who do not exit by f + 2	-0.693	-0.713	-0.800	*-0.107*
	Average Math Achievement			
Schools Required to Close				
Students who exit by f + 2	-1.022	-0.970	-0.835	*0.187*
Schools Not Required to Close				
Students who exit by f + 2	-1.001	-0.845	-0.792	*0.209*
Students who do not exit by f + 2	-0.909	-0.850	-0.766	*0.143*
	Average Reading Achievement			
Schools Required to Close				
Students who exit by f + 2	-0.909	-0.838	-0.754	*0.155*
Schools Not Required to Close				
Students who exit by f + 2	-1.005	-0.782	-0.739	*0.266*
Students who do not exit by f + 2	-0.855	-0.791	-0.829	*0.026*

NOTE. School quality estimates in the top panel are based on estimated annual school contributions to student achievement gains in both math and reading. See footnote 14 for additional detail.

Figure 1. Distribution of the Running Variable for the 55 School-Focal Year Combinations



NOTE. The histogram presents a count of school-focal year combinations for each bin of width 0.25 on the running variable. Values of the running variable greater than 0 (gray bars) indicate that the school qualified (and was identified) for mandatory closure.

Table 6. Covariate Balance at the Threshold for Mandatory Closure using RD Design

Outcome Variable	N	[1] Linear Spec. Coefficient (S.E.)	[2] Quadratic Spec. Coefficient (S.E.)
Standardized math score	6,016	-0.024	-0.043
		(0.085)	(0.110)
Standardized reading score	6,019	-0.041	-0.100
		(0.085)	(0.120)
White	6,027	0.008	0.098
		(0.058)	(0.114)
Asian	6,027	0.001	0.002
		(0.004)	(0.004)
Black	6,027	-0.012	-0.062
		(0.109)	(0.152)
Hispanic	6,027	-0.004	-0.070
		(0.087)	(0.084)
Multiple races	6,027	0.007	0.031
		(0.019)	(0.027)
Female	6,027	0.032	0.013
		(0.022)	(0.031)
Disability	6,027	0.011	0.005
		(0.028)	(0.044)
Economic disadvantage	6,027	-0.101	-0.214
		(0.103)	(0.174)
ELL	6,027	-0.009	-0.033
		(0.032)	(0.030)

NOTE. The table presents coefficient estimates and standard errors for the indicator of scoring above closure threshold from OLS models predicting observable student background characteristics. Each coefficient is from a separate regression. In addition to the indicator for required closure, the running variable for the regressions in column [1] is specified as a linear term interacted with the closure indicator. In column [2] the running variable is specified as linear and quadratic terms, each of which is interacted with the closure indicator. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01

Table 7. Regression Discontinuity Estimates of the Impact of Mandatory Closure on Student Achievement

	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Year Relative to the Focal Year	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
Two years prior to the "at risk"/focal year	2,880	0.072	0.079	-0.016	2,872	0.034	0.048	-0.064
(*f* - 2)	(36)	(0.082)	(0.081)	(0.109)	(36)	(0.067)	(0.068)	(0.103)
One year prior to the "at risk"/focal year	3,939	-0.083	-0.087	0.083	3,936	-0.147	-0.149	-0.020
(*f* - 1)	(36)	(0.090)	(0.081)	(0.092)	(36)	(0.100)	(0.095)	(0.125)
Focal year	6,016	-0.024	-0.011	-0.043	6,019	-0.041	-0.021	-0.100
(*f*)	(36)	(0.085)	(0.091)	(0.110)	(36)	(0.085)	(0.093)	(0.120)
One year after the "at risk"/focal year	4,144	-0.002	-0.013	0.058	4,144	0.018	0.028	0.022
(*f* + 1)	(36)	(0.072)	(0.072)	(0.075)	(36)	(0.063)	(0.068)	(0.079)
Two years after the "at risk"/focal year	2,737	0.018	0.024	-0.032	2,739	0.130	0.152	0.092
(*f* + 2)	(32)	(0.084)	(0.085)	(0.106)	(32)	(0.092)	(0.092)	(0.129)
Three years after the "at risk"/focal year	1,700	0.139	0.236**	0.051	1,702	0.103	0.206**	0.255*
(*f* + 3)	(25)	(0.085)	(0.095)	(0.139)	(25)	(0.088)	(0.090)	(0.138)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

Note. The table presents coefficient estimates and standard errors for an indicator of scoring above the closure threshold from OLS models predicting standardized student achievement. Each coefficient is from a separate regression. All models control for distance from the closure threshold in the manner specified in the table and school year fixed effects. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table 8. Regression Discontinuity Estimates of the Impact of Mandatory Closure on Student Achievement (Baseline test scores included as a covariate)

	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Year Relative to the Focal Year	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
One year after the "at risk"/focal year	4,137	-0.015	-0.038	0.037	4,141	0.02	0.017	0.035
(*f* + 1)	(36)	(0.055)	(0.052)	(0.076)	(36)	(0.051)	(0.053)	(0.069)
Two years after the "at risk"/focal year	2,732	0.049	0.052	0.026	2,737	0.149*	0.169**	0.201*
(*f* + 2)	(32)	(0.074)	(0.074)	(0.105)	(32)	(0.076)	(0.074)	(0.114)
Three years after the "at risk"/focal year	1,695	0.105	0.195**	0.193*	1,701	0.106	0.203**	0.234**
(*f* + 3)	(25)	(0.091)	(0.094)	(0.111)	(25)	(0.091)	(0.084)	(0.093)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

Note. The table presents coefficient estimates and standard errors for an indicator of scoring above the closure threshold from OLS models predicting standardized student achievement. Each coefficient is from a separate regression. All models control for distance from the closure threshold in the manner specified in the table, focal-year test scores, and school year fixed effects. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table 9. Instrumental Variables (IV) Regression Estimates of the Impact of School Closure

Model	Math [1] N (N Schools)	Math [2] Coef. (SE)	Math [3] Coef. (SE)	Math [4] Coef. (SE)	Reading [5] N (N Schools)	Reading [6] Coef. (SE)	Reading [7] Coef. (SE)	Reading [8] Coef. (SE)
	Two Years After the "At Risk" Focal Year (f+2)							
2SLS--1st stage: Effect of mandatory closure identification on the probability of closure.	2,732	0.857***	0.845***	0.824***	2,737	0.858***	0.845***	0.823***
	(32)	(0.015)	(0.017)	(0.021)	(32)	(0.015)	(0.017)	(0.021)
2SLS--2nd stage: Effect of closure on standardized achievement	2,732	0.057	0.062	0.031	2,732	0.174**	0.199**	0.244*
	(32)	(0.085)	(0.086)	(0.124)	(32)	(0.083)	(0.079)	(0.128)
	Three Years After the "At Risk" Focal Year (f+3)							
2SLS--1st stage: Effect of mandatory closure identification on the probability of closure.	1,695	0.863***	0.838***	0.834***	1,701	0.863***	0.838***	0.837***
	(25)	(0.017)	(0.022)	(0.024)	(25)	(0.017)	(0.022)	(0.024)
2SLS--2nd stage: Effect of closure on standardized achievement	1,695	0.122	0.233**	0.232*	1,701	0.122	0.242**	0.280***
	(25)	(0.102)	(0.109)	(0.134)	(25)	(0.102)	(0.095)	(0.110)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

NOTE. Each coefficient is from a separate OLS regression. All models control for distance from the closure threshold in the manner specified in the table, as well as focal year test scores, and school year fixed effects. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table 10. Instrumental Variables (IV) Regression Estimates of the Impact of Student Exit from "At Risk" Charter Schools

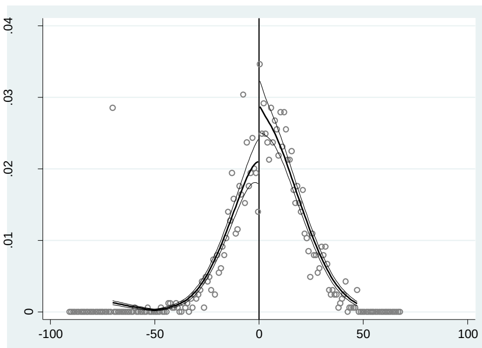
	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Model	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
	Two Years After the "At Risk" Focal Year (f+2)							
2SLS--1st stage: Effect of mandatory closure identification on the probability of closure.	2,732 (32)	0.488*** (0.021)	0.473*** (0.023)	0.420*** (0.028)	2,737 (32)	0.488*** (0.021)	0.472*** (0.023)	0.419*** (0.028)
2SLS--2nd stage: Effect of closure on standardized achievement	2,732 (32)	0.100 (0.149)	0.111 (0.153)	0.061 (0.243)	2,737 (32)	0.306** (0.150)	0.357** (0.154)	0.480* (0.288)
	Three Years After the "At Risk" Focal Year (f+3)							
2SLS--1st stage: Effect of mandatory closure identification on the probability of closure.	1,695 (25)	0.318*** (0.024)	0.278*** (0.032)	0.253*** (0.033)	1,701 (25)	0.319*** (0.024)	0.277*** (0.032)	0.261*** (0.033)
2SLS--2nd stage: Effect of closure on standardized achievement	1,695 (25)	0.330 (0.245)	0.704** (0.304)	0.764 (0.504)	1,701 (25)	0.331 (0.260)	0.733** (0.328)	0.897* (0.501)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

NOTE. Each coefficient is from a separate OLS regression. All models control for distance from the closure threshold in the manner specified in the table, as well focal year test scores, and school year fixed effects. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

9. Appendix

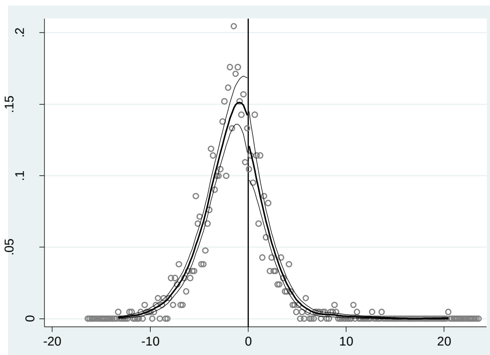
Figure A1. McCrary (2008) Density Test for the Performance Index, Gain Score, and Running Variable

Performance Index
All Charter Schools



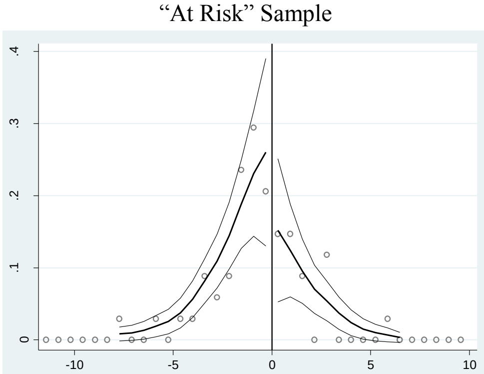
Academic Emergency (left) vs. Academic Watch (right)

Value Added Gain Index Score
All Charter Schools



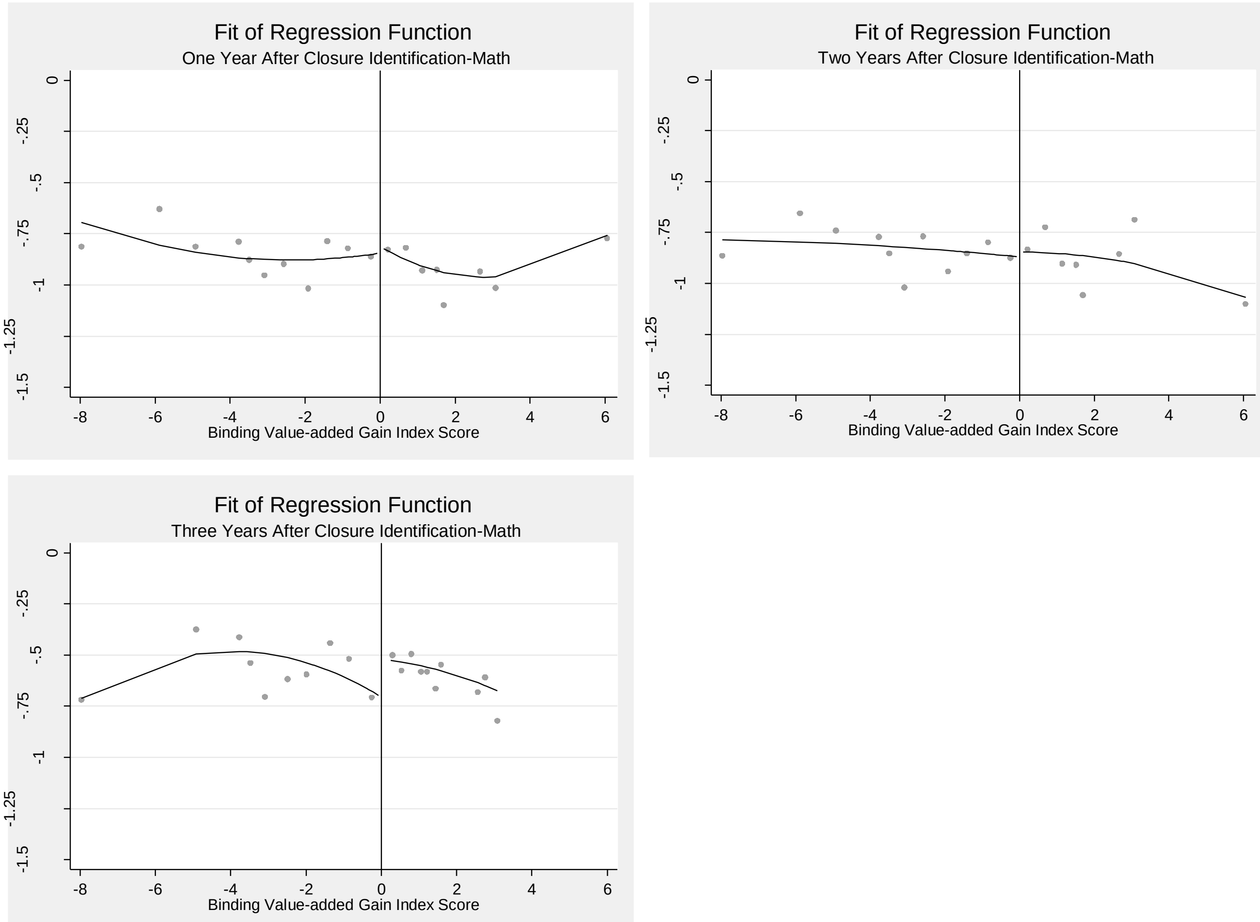
“Below Expected Gains” (left) vs. Met or Above “Expect Gains” (right)

Running Variable
“At Risk” Sample



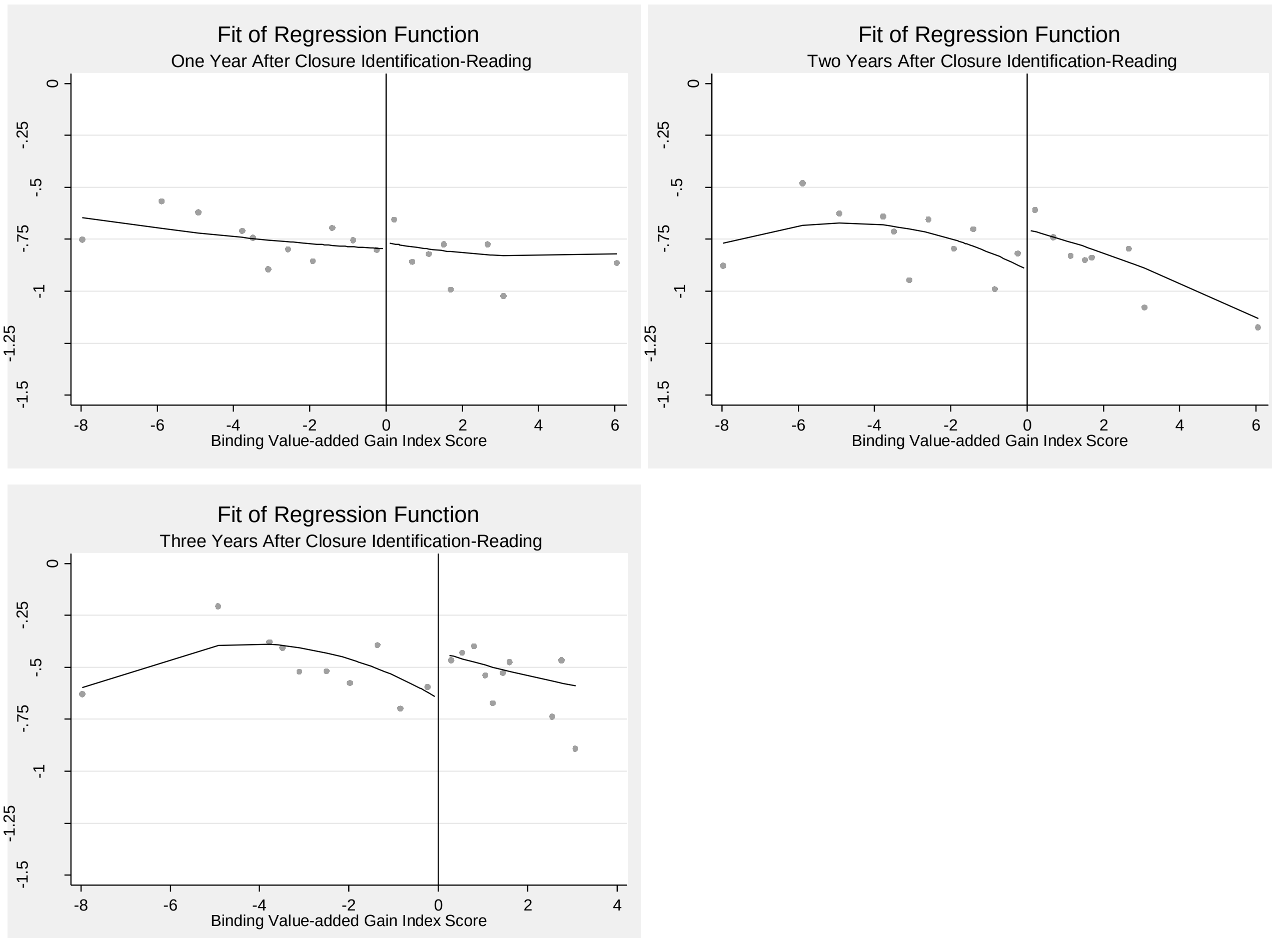
No Required Closure (left) vs. Required Closure (right)

Figure A2. Math Achievement Results (Column 4 in Table 8)



NOTE. The figure plots the fitted regression line from the math achievement model using a quadratic specification of the running variable interacted with the closure indicator (model 4 in Table 8). The markers in the figure represent local averages and are selected using the approach proposed by Calonico, Cattaneo, and Titiunik (2015) and depict the underlying variability in the data.

Figure A3. Reading Achievement Results (Column 8 in Table 8)



NOTE. The figure plots the fitted regression line from the reading achievement model using a quadratic specification of the running variable interacted with the closure threshold (model 8 in Table 8). The markers in the figure represent local averages and are selected using the approach proposed by Calonico, Cattaneo, and Titiunik (2015) and depict the underlying variability in the data.

Table A1. Within-Student Changes in School Quality Relative to the Focal Year

	At-Risk Schools Not Required to Close		At-Risk Schools Required to Close
Year	No Exit by *f*+2	Exit by *f*+2	Exit by *f*+2
One year prior to focal year	-0.441	-0.228	0.700
(*f* - 1)	(0.254)	(0.196)	(0.471)
Focal year	Omitted	Omitted	Omitted
(*f*)			
One year after focal year	0.059	0.533***	0.740**
(*f* + 1)	(0.276)	(0.170)	(0.267)
Two years after focal year	-0.094	0.563***	0.948***
(*f* + 2)	(0.425)	(0.179)	(0.258)
Three years after focal year	0.467	0.756***	1.406***
(*f* + 3)	(0.322)	(0.171)	(0.401)
N	2,422	4,005	5,114
N Students	600	1,074	1,282

NOTE. The results above are from OLS models estimating school quality (as measured by school-level value-added estimates) with student fixed effects. Robust standard errors are reported in parentheses below the coefficients. *p<0.10; **p<0.05; ***p<0.01

Table A2. Results from panel RD model

	Math		Reading	
	[1]	[2]	[3]	[4]
	Coef.	Coef.	Coef.	Coef.
Year Relative to Focal Year	(S.E.)	(S.E.)	(S.E.)	(S.E.)
One year after the focal year	-0.057	-0.078	0.004	-0.007
(*f* + 1)	(0.057)	(0.066)	(0.063)	(0.079)
Two years after the focal year	0.078	0.061	0.162**	0.191*
(*f* + 2)	(0.088)	(0.091)	(0.079)	(0.102)
Three years after the focal year	0.123	0.129	0.128	0.144
(*f* + 3)	(0.118)	(0.114)	(0.109)	(0.109)
Linear term	Yes	Yes	Yes	Yes
Quadratic term	Yes	Yes	Yes	Yes
Cubic term	No	Yes	No	Yes

NOTE. Results from panel RD method proposed by Cellini, Ferreira, and Rothstein (2010). Robust standard errors clustered by focal-year school in parentheses below coefficient estimates. *p<0.10; **p<0.05; ***p<0.01.

Table A3. Regression Discontinuity Estimates of the Impact of Mandatory Closure on Student Achievement (First time school was at risk of closure)

Year Relative to Focal Year	[1] Math Coef. (S.E.)	[2] Reading Coef. (S.E.)
	Cross-sectional Models	
One year after the focal year	0.109	0.062
(*f* + 1)	(0.087)	(0.100)
Two years after the focal year	0.041	0.195**
(*f* + 2)	(0.120)	(0.081)
Three years after the focal year	0.234*	0.269**
(*f* + 3)	(0.121)	(0.097)
N (*N* Schools)- Year 1	2,909 (34)	2,909 (34)
N (*N* Schools)- Year 2	2,153 (29)	2,156 (29)
N (*N* Schools)- Year 3	1,353 (24)	1,357 (24)
	Panel Model	
One year after the focal year	-0.007	0.008
(*f* + 1)	(0.075)	(0.092)
Two years after the focal year	0.114	0.160**
(*f* + 2)	(0.103)	(0.071)
Three years after the focal year	0.166	0.191*
(*f* + 3)	(0.130)	(0.111)
N	15,456	15,442
N Students	4,179	4,178

NOTE. Replicates models 4 and 8 from Table 8 using data only from the first time a school is identified as "at risk." The specifications all feature quadratic polynomials interacted with the closure indicator. The top panel features estimates from separate regressions (as in Table 8) whereas the bottom panel features coefficients estimated simultaneously using the panel RD method from Cellini, Ferreira, and Rothstein (2010). Robust standard errors clustered by school in parentheses below coefficient estimates. *p<0.10; **p<0.05; ***p<0.01.

Table A4. Replication of Table 8 – Omitting three schools that met the value-added gain index closure requirement prior to focal year

	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Year Relative to the Focal Year	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
Two years prior to the "at risk"/focal year	2,830	0.065	0.073	-0.029	2,822	0.027	0.043	-0.083
(*f* - 2)	(33)	(0.083)	(0.082)	(0.113)	(33)	(0.069)	(0.070)	(0.105)
One year prior to the "at risk"/focal year	3,855	-0.085	-0.084	0.081	3,853	-0.155	-0.154	-0.029
(*f* - 1)	(33)	(0.084)	(0.076)	(0.090)	(33)	(0.098)	(0.094)	(0.121)
Focal year	5,878	-0.023	-0.009	-0.041	5,881	-0.039	-0.018	-0.094
(*f*)	(33)	(0.084)	(0.091)	(0.108)	(33)	(0.085)	(0.093)	(0.120)
One year after the "at risk"/focal year	4,027	-0.027	-0.051	0.022	4,031	-0.005	-0.006	0.004
(*f* + 1)	(33)	(0.058)	(0.053)	(0.082)	(33)	(0.050)	(0.052)	(0.072)
Two years after the "at risk"/focal year	2,640	0.035	0.038	0.010	2,645	0.134	0.155*	0.171
(*f* + 2)	(29)	(0.079)	(0.078)	(0.116)	(29)	(0.079)	(0.077)	(0.121)
Three years after the "at risk"/focal year	1,615	0.079	0.183*	0.144	1,621	0.085	0.197**	0.324***
(*f* + 3)	(22)	(0.090)	(0.101)	(0.091)	(22)	(0.093)	(0.089)	(0.112)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

NOTE. Replicates Table 8 to a restricted sample that omits schools that met the value-added closure requirement prior to the focal year, such that receiving the "academic emergency" designation in the focal year led to closure. Robust standard errors clustered by school in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table A5. Regression Discontinuity Estimates of the Impact of Mandatory Closure on Student Achievement (Focal-year running variable)

	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Year Relative to the Focal Year	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
Two years prior to the "at risk"/focal year	2,058	0.052	0.044	0.130	2,053	0.094	0.058	0.085
(*f* - 2)	(25)	(0.094)	(0.100)	(0.138)	(25)	(0.074)	(0.113)	(0.123)
One year prior to the "at risk"/focal year	2,823	0.002	0.088	0.041	2,825	-0.087	-0.056	-0.151
(*f* - 1)	(25)	(0.073)	(0.089)	(0.097)	(25)	(0.107)	(0.126)	(0.141)
Focal year	4,181	0.022	0.043	-0.074	4,183	0.026	0.018	-0.230
(*f*)	(25)	(0.133)	(0.158)	(0.164)	(25)	(0.142)	(0.164)	(0.146)
One year after the "at risk"/focal year	2,923	-0.027	0.023	0.035	2,925	0.014	0.032	0.003
(*f* + 1)	(25)	(0.071)	(0.080)	(0.087)	(25)	(0.063)	(0.073)	(0.083)
Two years after the "at risk"/focal year	2,003	0.016	0.041	0.041	2,004	0.127	0.128	0.037
(*f* + 2)	(22)	(0.086)	(0.101)	(0.122)	(22)	(0.123)	(0.138)	(0.118)
Three years after the "at risk"/focal year	1,157	0.268**	0.265**	0.230**	1,156	0.197***	0.201***	0.220**
(*f* + 3)	(16)	(0.091)	(0.091)	(0.099)	(16)	(0.060)	(0.058)	(0.101)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

Note. The table presents coefficient estimates and standard errors for an indicator of scoring above the closure threshold from OLS models predicting standardized student achievement. Each coefficient is from a separate regression. All models control for distance from the closure threshold in the manner specified in the table and school year fixed effects. Models for post-focal years include baseline test score. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table A6. Regression Discontinuity Estimates of the Impact of Mandatory Closure on Student Achievement (SY2008-2010)

	Math				Reading			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	N	Coef.	Coef.	Coef.	N	Coef.	Coef.	Coef.
Year Relative to the Focal Year	(N Schools)	(SE)	(SE)	(SE)	(N Schools)	(SE)	(SE)	(SE)
One year after the “at risk”/focal year	2,828	0.051	0.043	0.070	2,829	0.060	0.072	0.149
(*f* + 1)	(26)	(0.075)	(0.075)	(0.093)	(26)	(0.079)	(0.085)	(0.123)
Two years after the “at risk”/focal year	2,192	0.028	0.034	0.115	2,197	0.061	0.107	0.262**
(*f* + 2)	(26)	(0.105)	(0.108)	(0.186)	(26)	(0.086)	(0.091)	(0.120)
Three years after the “at risk”/focal year	1,695	0.105	0.195**	0.193*	1,701	0.106	0.203**	0.234**
(*f* + 3)	(25)	(0.091)	(0.094)	(0.111)	(25)	(0.091)	(0.084)	(0.093)
Linear term	NA	Yes	Yes	Yes	NA	Yes	Yes	Yes
Linear term interacted with closure indicator	NA	Yes	No	Yes	NA	Yes	No	Yes
Quadratic term	NA	No	Yes	Yes	NA	No	Yes	Yes
Quadratic term interacted w/ closure indicator	NA	No	No	Yes	NA	No	No	Yes

Note. The table presents coefficient estimates and standard errors for an indicator of scoring above the closure threshold from OLS models predicting standardized student achievement. Each coefficient is from a separate regression. All models control for distance from the closure threshold in the manner specified in the table, student achievement in the focal year, and school year fixed effects. Robust standard errors clustered by focal-year school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.

Table A7. Regression Discontinuity Estimates of the Impact of Mandatory Closure on the Value-Added Score of the School Students Attend

Year Relative to Focal Year	[1] Coef. (S.E.)	[2] Coef. (S.E.)	[3] Coef. (S.E.)
		Cross-sectional Models	
One year after the focal year	-0.092	-0.115	0.052
(f + 1)	(0.217)	(0.213)	(0.245)
Two years after the focal year	0.269	0.317	0.236
(f + 2)	(0.275)	(0.271)	(0.388)
Three years after the focal year	0.644**	0.797***	0.078
(f + 3)	(0.242)	(0.256)	(0.315)
N (*N* Schools)- Year 1	4,093 (36)	4,093 (36)	4,093 (36)
N (*N* Schools)- Year 2	2,677 (32)	2,677 (32)	2,677 (32)
N (*N* Schools)- Year 3	1,671 (25)	1,671 (25)	1,671 (25)
Linear term	Yes	Yes	Yes
Linear term interacted with closure indicator	Yes	Yes	Yes
Quadratic term	No	Yes	Yes
Quadratic term interacted w/ closure indicator	No	No	Yes

Note. The table presents coefficient estimates and standard errors for an indicator of scoring above the closure threshold from OLS models predicting standardized value-added of the school that students attend. Each coefficient is from a separate regression. All models control for distance from the closure threshold in the manner specified in the table, student demographic characteristics, and school year fixed effects. Robust standard errors clustered by school are in parentheses below coefficient estimates: *p<0.10; **p<0.05; ***p<0.01.