



Is the Rise of High School Graduation Rates Real? High-Stakes School Accountability and Strategic Behavior

Douglas N. Harris
Tulane University

Lihan Liu
Tulane University

Nathan Barrett
Tulane University

Ruoxi Li
Tulane University

High school graduation rates have increased dramatically in the past two decades. Some skepticism has arisen, however, because of the confluence of the graduation rise and the starts of high-stakes accountability for graduation rates with No Child Left Behind (NCLB). In this study we provide some of the first evidence about the role of accountability versus strategic behavior, especially the degree to which the recent graduation rate rise represents increased human capital. First, using national DD analysis of within-state, cross-district variation in proximity to state graduation rate thresholds, we confirm that NCLB accountability increased graduation rates. However, we find limited evidence that this is due to strategic behavior. To test for lowering of graduation standards, we examined graduation rates in states with and without graduation exams and trends in GEDs; neither analysis suggests that the graduation rate rise is due to strategic behavior. We also examined the effects of “credit recovery” courses using Louisiana micro data; while our results suggest an increase in credit recovery, consistent with some lowering of standards, the size of the effect is not nearly enough to explain the rise in graduation rates. Finally, we examine other forms of strategic behavior by schools, though these can only explain inflation of school/district-level graduation rates, not national rates. Overall, the evidence suggests that the rise in the national graduation rates reflects some strategic behavior, but also a substantial increase in the nation’s stock of human capital. Graduation accountability was a key contributor.

VERSION: March 2020

Is the Rise in High School Graduation Rates Real? High-Stakes School Accountability and Strategic Behavior

Douglas N. Harris
Lihan Liu
Nathan Barrett
Ruoxi Li

February 29, 2020

Abstract: High school graduation rates have increased dramatically in the past two decades. Some skepticism has arisen, however, because of the confluence of the graduation rise and the starts of high-stakes accountability for graduation rates with *No Child Left Behind* (NCLB). In this study we provide some of the first evidence about the role of accountability versus strategic behavior, especially the degree to which the recent graduation rate rise represents increased human capital. First, using national DD analysis of within-state, cross-district variation in proximity to state graduation rate thresholds, we confirm that NCLB accountability increased graduation rates. However, we find limited evidence that this is due to strategic behavior. To test for lowering of graduation standards, we examined graduation rates in states with and without graduation exams and trends in GEDs; neither analysis suggests that the graduation rate rise is due to strategic behavior. We also examined the effects of “credit recovery” courses using Louisiana micro data; while our results suggest an increase in credit recovery, consistent with some lowering of standards, the size of the effect is not nearly enough to explain the rise in graduation rates. Finally, we examine other forms of strategic behavior by schools, though these can only explain inflation of school/district-level graduation rates, not national rates. Overall, the evidence suggests that the rise in the national graduation rates reflects some strategic behavior, but also a substantial increase in the nation’s stock of human capital. Graduation accountability was a key contributor.

Acknowledgments: This study was conducted at the Education Research Alliance for New Orleans, based in the Department of Economics at Tulane University. The authors wish to thank the organization’s funders: the John and Laura Arnold Foundation, the USDOE Institute of Education Sciences, William T. Grant Foundation, the Spencer Foundation and, at Tulane, the Department of Economics, Murphy Institute and School of Liberal Arts. For helpful comments, we thank John Papay. We also thank the staff of the Louisiana Department of Education for their helpful assistance with the data.

I. Introduction

High school graduation has long been considered a key marker of human capital development. The U.S. led the world in high school graduation through the 1960s (Murnane, 2013), which helped drive decades of economic growth throughout the first two-thirds of the twentieth century (Goldin & Katz, 2008; Gordon, 2015). However, this was followed, in the 1970s through 1980s, by slower growth (Mishel & Roy, 2006) or perhaps even declines in graduation (Heckman & Lafontaine, 2007). The rest of the world has also been catching up and the U.S. graduation ranking declined from 1st to 13th (out of 19 countries) in the OECD (Murnane, 2013).

While there remains some debate about the graduation trends prior to 2000, there is no question that U.S. graduation rates began to climb quickly afterwards. The national graduation rate, calculated from Current Population Survey (CPS) workforce surveys shown in Figure 1, increased by about 10 percentage points between 2000 and 2016 (OECD, 2012; Murnane, 2013). President Obama and Secretary of Education, Arne Duncan, touted these as important accomplishments (Duncan, 2018). However, the low-stakes CPS trend displays a somewhat different pattern than the two high-stakes trends, reported by the National Center for Education Statistics (NCES). The growth in the low-stakes graduation rate (or “status completion”) is almost linear from the 1980s onward, while the NCES graduation rate, based on school-provided data, increased sharply starting around 2000.¹ From 2000-2016, the national high-stakes graduation measure increased by 18 percentage points, eight points more than in the CPS.

¹ The two measures are not quite on the same scale, which partly explains why the NCES numbers are consistently about 10 percentage points lower than the CPS. The NCES is the average freshmen graduation rate (AFGR) and does not include GED and other forms of high school completion. See the figure notes and section II for details.

One obvious potential contributor has been a change in federal policy that coincided with the graduation rate increase. *No Child Left Behind* (NCLB) became law in 2001 and is well known for its small or null effects on test scores (Dee & Jacob, 2011). Less recognized is that the original law also required accountability for high school graduation. At first, states were only required to include graduation in some fashion; they could set any target and improvement rates they wished, but those schools that did not meet these goals were threatened with sanction. Starting in 2010, the U.S. Department of Education (USDOE) started pressuring states to set higher graduation rate goals (all states set their thresholds above 80% graduation), and requiring that states generate substantial improvement for schools and districts.² (States were also required to adopt the four-year adjusted-cohort graduation rate calculation as the basis for these calculations.) The timing of these policy changes, during 2002-2010, coincide with the high-stakes graduation rate increases in Figure 1, suggesting that accountability helped cause the change. The purpose of this study is to understand the overall effect of accountability on graduation rates, and to what degree this might reflect an increase in students' human capital versus strategic behavior of schools.

One potential effect is that the NCLB-induced pressure led schools to improve educational practice in ways that led students to stay in school longer, e.g., increasing teacher quality, doing more to engage students in school, or helping them see the value of high school for future life success (Hanushek & Rivkin, 2010; Reback, Rockoff, Schwartz, 2014; Rouse et al., 2013).³ The additional graduates, in this case, would have experienced more time on task and learned

² From 2002-2010, any improvement was a common goal, but, after 2010, typically states required a 2 percent annual increase. <https://www2.ed.gov/admins/lead/account/stateplans03/gradratechart.pdf>

³ Jackson (2018) finds evidence that teachers increase the probability of high school completion and that these effects are largely unrelated to teacher value-added to test scores.

academic, social, and other skills than they would have without the law. In other words, the law might have increased the stock of human capital.

Strategic behavior is an alternative explanation. Ample evidence suggests that test-based accountability has led schools to teach to the test (Koretz, 2009), cheat (Jacob, 2005), and even change cafeteria food (e.g., Figlio and Winicki, 2005) all in ways that increase scores without increasing human capital. Consistent with this story, low-stakes NAEP scores have been increasing much more slowly than high-stakes scores (Koretz, 2009). As Figure 1 suggests, the same could be happening with graduation rates now that they, too, have become high stakes.

Two general forms of strategic behavior are possible in this case. First, schools might make it easier for students to graduate, reducing the quality of credentials. While states set standards regarding which types of courses student take, schools have considerable autonomy for determining the content of those courses, the ease of passing them, and, more generally, what constitutes graduation. Teachers have authority to determine who passes courses and may be under pressure from school principals to inflate grades and make it easier to pass courses. Students can also choose between regular diplomas and GEDs, which are widely considered to be less rigorous than diplomas (Cameron & Heckman, 1993), even if they do seem to benefit those who obtain them relative to those who drop out with similar baseline skills (Murnane, Willett & Tyler, 2004).

Another more recent development is the availability of “credit recovery,” i.e., programs that help struggling students earn credits for courses they have failed, to stay on track to graduate. In some of these programs, students can receive course credit simply by passing a state competency-based test, without meaningful instruction or seat time, similar to the GED (Heckman & LaFontaine, 2006). Recent evidence suggests that online courses in high school

increase high school graduation, but with limited influence on academic achievement (Heinrich et al., 2019).

The above degree quality problem can be framed as a concern with the numerator of the graduation rate. The second potential distortion is with the denominator. Schools control which students count toward their graduation rate calculation through the exit codes. Since schools themselves assign these codes, they can remove students from their accountability metrics by inaccurately assigning certain codes. For example, a recent audit in Louisiana (Louisiana Department of Education, n.d.) was unable to confirm large numbers of exit codes for students who were listed as having transferred to private or out-of-state schools, or to home schooling, all of which removed them from schools' graduation rate calculations. These exit codes are more difficult for the state to verify than most others.

With both forms of strategic behavior—changing graduation requirements and changing who counts as a potential graduate—the key observation is that the data used to calculate performance are being created by the schools being held accountable for those measures, with relatively little oversight or enforcement.⁴ Anecdotal evidence on the issue does point to strategic behavior as a real problem. In Washington D.C., credit recovery was used inappropriately and grading norms began to shift so that essentially all students passed their courses (McGee, 2018; St. George, 2018).⁵ In Prince George's County, Maryland, a state investigation found improper course grade changes and graduation of students who had absences from school above the level allowed by state law (St. George and Bui, 2017). In San Diego and

⁴ Note that this form of strategic behavior is not reflected in Figure 1 because the AFGR (the only graduation measure from state-reported data that is available consistently over time) does not remove students receiving these hard-to-verify exit codes. For this reason, Figure 1 under-states potential accountability distortion. In Louisiana, for example, the state-reported graduation rate (not the AFGR) increase by 13 percentage points (Perry, Harris, Buerger & Mack, 2015).

⁵ Specifically, teachers sent students home with packets of homework and counted this as course completion; there were no rules or standards established (McGee, 2018).

other school districts in California, the definition of graduation rate did not meet state standards and inflated the outcome: students who had adult education degrees or passed high school proficiency exams were counted as graduates, and students who exited the school were sometimes incorrectly excluded from the calculation of graduation rate (Huntsberry, 2018). In New Orleans, grades were changed to make it easier for students to pass courses and graduate (Jewson, 2019).

In this paper, we test the size and nature of accountability effects using various types of difference-in-differences (DD) analysis: First, we construct a national state-level data set that goes back to several years prior to federal accountability requirements. For each state, we identify the graduation threshold and which districts were below this standard.⁶ We find that graduation rates increased more in states that had more districts threatened by sanction; in models with state fixed effects, we also find that low-graduation-rate districts responded the most to graduation accountability. These analyses provide evidence that graduation accountability increased graduation rates.

Taken together, the above analyses suggest NCLB increased high school graduation. Given prior evidence of strategic responses to test-based accountability, the divergence of the low- and high-stakes graduation metrics in Figure 1 provides *a priori* evidence that not all of the increase in graduation rates has reflected human capital improvement. Accountability tends to generate improved outcomes in part through strategic behavior. To provide more direct evidence on the source of graduation rate increases, we carried out four additional types of analysis.

⁶ School-level graduation rates were not available pre-NCLB. Also, for reasons explained later, we cannot use the same methods as Dee and Jacob (2011) because there was apparently no pre-NCLB cross-state variation in graduation accountability, as there was with test-based accountability.

First, we were able to identify which states required that students pass graduation exams in order to graduate, at three different time points (2000, 2010, and 2017). Since every student has to pass the test in order to graduate, schools in states without these exams could more easily reduce their standards. Therefore, if graduation rate increases were strictly due to strategic behavior, we would expect the growth in graduation rates to be concentrated in states that did not have graduation exams. However, we see almost the opposite pattern: graduation rates increased somewhat more in states with graduation exams. (The potential for quasi-experimental evidence on this topic appears limited for reasons explained later.) This suggests that lowering graduation standards is unlikely to be the main explanation.

Second, we identified reports showing national trends in regular high school diplomas versus GEDs. Federal graduation accountability guidelines generally exclude GEDs. Apparently as a result, the number of GEDs has been in a steady decline since NCLB (NCES, 2010; Barshay, 2018). This, again, suggests that the slower increase in CPS graduation is due to a decline in GEDs and therefore reflects improved credentials rather than lower standards.

Third, panel data from Louisiana allow us to impute credit recovery courses for individual students. Specifically, we identify students who retook a course after initially failing it, during a period when schools had the option of using credit recovery courses. We then tested whether the trends in these credit recovery courses differed before and after accountability pressures in Louisiana increased, based on whether schools were threatened with accountability sanctions. We do find some evidence of an increase in credit recovery in schools threatened by accountability, though not nearly enough to explain the overall graduation rate increases.

The above analyses focus on strategic behavior in the numerator. To understand potential distortion in the denominator, we used detailed Louisiana data to analyze the exit codes that

schools used to identify graduates and to exclude students from graduation calculations. Some of these are difficult for states to verify and are therefore likely candidates for schools seeking to increase their graduation rates artificially. We see some evidence that graduation rates increased the use of hard-to-verify exit codes, but these effects are very small relative to the size of the graduation rate increases, especially in the national numbers. School and district graduation rates are still subject to distortion through this mechanism, but our interest here is mainly in the national trends.

We continue, in Section II, by describing the national state/district-level data and Louisiana student-level data that we use in our analysis. We also start that section by discussing how different methods of calculating graduation rates are applied in our analysis, and potential distortions that arise with each measure, which is necessary later in the interpretation of results. In section III, we outline our quasi-experimental methods and present results in Section IV. Section V concludes.

II. Graduation Rate Measures and Data

II.A. Different Rates and Potential Distortion

Graduation rates can be calculated in many ways, and several of these are useful in our analysis. The three key distinguishing features are: the method of calculation, the data source, and the government/geographic unit (school, district, state, etc.).

The CPS calculates the percentage of 18- to 24-year-olds not enrolled in high school or a lower education level who hold a high school diploma or an alternative credential, such as a GED. This rate includes all civilian, noninstitutionalized individuals 18- to 24-year-olds who have completed high school, including individuals who completed their education outside of the

United States.⁷ This is what we reported in Figure 1. Because it is the only consistent measure available over many years, we also use it when we test for the role of graduation exams. While not the focus of our study, there has been debate about the degree to which the CPS status completion rate might be biased due to the exclusion of military personnel and incarcerated adults,⁸ and the degree to which the trends reflect GEDs versus regular diplomas (Heckman & Lafontaine, 2007; Mishel & Roy 2006).⁹

Next, we turn to the various graduation measures created from National Center for Education Statistics (NCES) data that are reported by schools. The *average freshmen graduation rate* (AFGR) divides the number of students receiving regular high school diplomas in a given year by the simple average of the number of 8th, 9th, and 10th graders in the same government unit four, three and two years earlier, respectively. The obvious problem with this approach is that it does not measure exits directly, or even try to distinguish between legitimate and non-legitimate ones. Instead, exits are captured, only indirectly, in the enrollment figures.¹⁰ Nevertheless, this is the only measure, other than CPS, that is available pre- and post-NCLB. Thus, it is also reported in Figure 1 and used in the national state-level and district-level DD analysis.¹¹

⁷ Data Source: U.S. Department of Commerce, Census Bureau, Current Population Survey (CPS), October. NCES Definition of status completion rate: https://nces.ed.gov/programs/dropout/ind_03.asp.

⁸ Another issue of debate has been the role of immigration in influencing the status completion rate, since maybe people in the 18-24 range had no experience in U.S. high schools (Heckman & Lafontaine, 2007; Mishel & Roy, 2006). Immigration could affect the status completion even though it does not reflect changes.

⁹ Mishel and Roy (2006) rely mainly on the National Educational Longitudinal Survey (NELS), but this data source is not useful in our analysis.

¹⁰ If a student exits for any reason, this reduces enrollment, which reduces the denominator. Because enrollments only count for 8th-10th graders, whether the exits affect the AFGR depends on which grades they occur in and the exact timing vis-à-vis the enrollment “count day.”

¹¹ While they use different definitions, the [AFGR](#) numbers closely track the [ACGR](#), at least in the aggregate. For instance, in 2012-13, there was less than one percentage point difference between the AFGR (81.9%) and the ACGR (81.4%) (Source: <https://nces.ed.gov/blogs/nces/post/what-is-the-difference-between-the-acgr-and-the-afgr>).

While the AFGR is created from school-reported data, this measure is not what most states used, nor has it ever been high-stakes. Pre-NCLB, thirty states reported the *NCES graduation rate* (NCES-GR), as recommended by the U.S. Department of Education. The NCES-GR is the number of graduates divided by the number of graduates plus the number of 12th graders who dropped out that year, plus the number of 11th graders who dropped out the prior year, and so on down to the 9th graders from four years earlier.¹² The NCES-GR reflects an assumption that all students either graduate or drop out.

The current federally mandated standard for graduation rates is the *adjusted cohort graduation rate* (ACGR). The ACGR divides the number of graduates by the number students enrolled in the same school or government unit in their freshmen year or who transferred in, while subtracting from the denominator the *legitimate exits*, i.e., those who left the government unit for legitimate reasons that completely remove the students from the graduation calculation, as if the students had never entered the exiting school (see examples below). *Non-legitimate* exits count against the government unit. Dropout is the most common non-legitimate exit. When students drop out, this reduces the graduation rate. Missing data are also counted as non-legitimate exits. The calculation of the ACGR became the federally required method of calculation in 2010, in part because of improvements in data systems that allowed schools to track individual students over time. The ACGR calculations are impossible without such detailed data and were impossible to make on a national scale before NCLB (or for many years afterwards).

¹² Pre-NCLB, another 10 states used something close to the ACGR. Finally, five states used alternative rates (see notes to Table 1).

Partly because accountability places pressure on schools to raise graduation rate, a key measurement problem is that non-legitimate exits might be miscoded as legitimate exits, e.g., a dropout is coded as a transfer. In cases where schools have incentives to inflate graduation rates, this is especially likely with legitimate exits that are difficult for states to verify. Based on our conversations with school leaders, the three most difficult-to-verify legitimate exits are: transfers to public schools outside the geographic unit, transfers to private schools, and transfers to home schooling. In all three cases, the transfers are moving the student into systems that are outside the control of state agencies (e.g., private schools), reducing verifiability. Note that these same issues apply to both the ACGR and the NCES-GR. While the latter focuses on graduates and dropouts, schools were allowed to eliminate legitimate exits. The potential for miscoding is even greater in the NCES-GR because state agencies had no way of checking any of the exit codes were indeed legitimate; with the ACGR, it is possible for state agencies to check transfers to other public schools, for example.¹³

Table 1 summarizes the calculation and potential distortions in each of the four graduation rate calculations. There is only one option for measuring graduation with the CPS data: status completion at the state/national level. The AFGR, too, is only available at the state/national level. The NCES completion rate and ACGR can be calculated at the school/district level or aggregated up to the state/national level.

As the table shows, all four measures are subject to strategic behavior in the numerator, i.e., to schools lowering their standards and making it easier for students to graduate. However, the potential strategic behavior with the denominator varies across measures. There is essentially no

¹³ This difference is less about the method of calculation than it is about the extent of data necessary to carry out the calculation. The ACGR requires more fine-grained distinctions in exit codes than the NCES-GR. Given the same data, the potential for validity checks by state agencies is the same for both measures.

way to distort the CPS denominator because it includes the entire population (or really a random sample of that population). In contrast, with the NCES-GR and the ACGR, there is an incentive to code non-legitimate exits (especially dropouts) as legitimate exits. This incentive for schools also affects the state/national calculations; dropouts who are miscoded by schools as transfers to out-of-state public, private, and homeschooling are also miscoded as such when the data are rolled up to the state level. With the AFGR, the only way to increase the rate through the denominator is to reduce 8th-10th grade enrollments, but this would be counter-productive since school funding is partially based on enrollments. So, while there may be errors in the AFGR, they should not be induced, directly or indirectly, by accountability.¹⁴

In light of the above analysis, it is worth taking a second look at the trends in graduation rates in Figure 1. The above analysis leads to the important conclusion that the denominators of the CPS and AFGR are unlikely to be distorted, so that the differences in the trends almost has to be due to distorted numerators. In addition to relaxed graduation standards for any given credential, the fact that the AFGR excludes GEDs, and the CPS includes them, could mean that the larger increase in the AFGR in Figure 1 reflects schools shifting students away from GEDs toward regular diplomas. Such a shift would not show up in the CPS figures, which counts them equally, but it would show up in the AFGR.

This discussion of different graduation calculations argues for the use of multiple graduation calculations. By analyzing multiple measures, each with different distortions and available over

¹⁴ It is conceivable, not especially likely in our view, that distortions in the NCES-GR and ACGR could show up indirectly in the AFGR. If the incentives created by the NCES-GR and ACGR led schools to push students out (e.g., transfer to private schools) in 8th, 9th, or 10th grade, then this would reduce enrollment levels and therefore inflate the AFGR. However, the incentive with the NCES-GR and ACGR is to miscode non-legitimate exits, not to increase exits overall.

different time frames, we can better understand the nature of strategic behavior and therefore the degree to which the rise in the graduation rate reflects strategic behavior versus human capital.

II.B. National District-Level and State-Level Data

Using data from the Common Core of Data (CCD), we calculated the AFGR for all districts in the United States. We restrict the sample to those districts that have at least two years of graduation rates available prior to graduation accountability (1999-2002) and afterwards (2006-2010).¹⁵¹⁶ This yields 4,557 districts in 34 states included in the district-level analysis, or about 40 percent of all districts in the country.¹⁷ The post-treatment years stop in 2010 because, at that point, the law required a switch from the AFGR to cohort graduation rates (CGR), which creates issues of comparability.¹⁸ Also, it does not appear that the cohort graduation rates are available post-2010 at the district level on a national scale.

The state-level data set analysis is constructed by taking a weighted average of the district-level graduation up to the state level. Our analysis also sometimes includes additional control variables. The district-level expenditure variable is from Common Core of Data (CCD) finance data files¹⁹ and the district-level poverty variable comes from Census Bureau.²⁰

¹⁵ This is, specifically, the “Local Education Agency (School District) Universe Survey Dropout and Completion Data.” See: <https://nces.ed.gov/ccd/drpagency.asp>.

¹⁶ The years 2003-2005 are excluded from the analysis mainly because: (a) the data available from 2003-2005 are incomplete and would thus require either an unbalanced panel with a shifting sample or a balanced, but small sample; and (b) 2003-2005 is a period of transition for many states that saw large number of revisions of state accountability policies to meet the federal standard. As a result, these years are not very informative about the policy effects. We also drop the year 1998 from all parts of the analysis because these data display unusual patterns that call into question their validity. For each district, we imputed missing graduation rates using linear interpolation (separately for pre and post treatment years). The same imputation method was also applied to poverty rate and per pupil expenditure.

¹⁷ The included states are AK, AL, AZ, AR, CT, DE, GA, ID, IL, IA, LA, MA, MD, ME, MN, MO, MS, MT, ND, NE, NJ, NM, OK, OR, PA, RI, SD, TN, UT, VT, WI, WV, WY. Within these states, roughly 5 percent of districts having missing data in some years, which we imputed (linear interpolation).

¹⁸ We considered constructing AFGR graduation results from underlying data but the USDOE stopped reporting the number of graduates by district after 2010.

¹⁹ <https://nces.ed.gov/ccd/f33agency.asp>

²⁰ <https://www.census.gov/did/www/schooldistricts/data/poverty.html>

Table 2 summarizes the years of the pre- and post-treatment data we use in the analysis, in relation to the changes in policy. As the middle column shows, we use 1998-99 and 1999-00 (spring years) as the pre-treatment years. As a robustness check, we also follow Dee and Jacob (2011) and use 2000-01 and 2001-02 academic years as the pre-treatment years. While NCLB passed in 2001, the law was not implemented until after 2003 (Swanson 2003). The results are very similar when we use 1998-99 and 1999-00.

II.C. Louisiana Credit Recovery and Exit Code Data

Louisiana is a useful state to study because the Louisiana Department of Education (LDOE) has one of the older and more extensive student-level data systems in the United States, which allows for more precise understanding of graduation rates. Below, we discuss both the credit recovery data and the exit codes.

We cannot, unfortunately, identify credit recovery courses directly because the Louisiana transcript data do not distinguish these from other in-person courses. What we can observe is whether students take more than one math class in a year, during a time when we know online credit recovery courses were available. This measure might over-count credit recovery because students might have retaken courses in traditional classroom settings and formats. However, our conversations with school principals, and other evidence below, suggest our approach is a reasonable approximation. As described later, we are primarily interested in the changes in these courses after accountability ratcheted up, so our assumption is that these marginal courses are credit recovery.

Next, we discuss the exit code data. State-defined student exit codes include dozens of ways students can leave school, including various forms of credential completion and non-completion, for the 2001-2015 spring years. As shown in Table 3, in Louisiana, the most common type of

completion by far is graduation with a diploma (49.9 percent of all exits).²¹ Other forms of completion include GED (2.7 percent), certificate of achievement (0.9 percent), and other forms of completion (sum to 1.5 percent). Some of the exits without completion are considered by the state to be “legitimate” meaning the student is removed from the denominator for purposes of the cohort graduation rate and therefore removed from the school’s official responsibility. The most common legitimate exit codes are: transfer to other public schools within state (16.0 percent of all exits), transfer out of state (7.4 percent), transfer to home study (2.0 percent), and transfer to private school (1.9 percent).²² All other exits are considered non-legitimate and are counted against the school. The most common are missing (5.8 percent), transfer to adult ed (4.7 percent of all exits), dropout (3.4 percent) and no show (1.4 percent), and, expelled (<1 percent), and others.²³

Among the legitimate exit codes, and given potential concerns about strategic behavior with the increasingly high-stakes nature of high school graduation, we also have to consider how easily the state can verify exit codes. Federal guidelines require schools to provide “confirmation in writing” of exits from the public school system, but the meaning of this was unclear and enforcement was weak (USDOE, 2008; Murnane, 2013). The simplest of these, given the state’s

²¹ Note that “percent of exits” is quite different from the graduation rate. Essentially, percent of exits that are graduates is the number of graduates divided by the number of dropouts plus the number of transfers (plus the number of graduates). We report percent of exits to allow for potentially sharp changes in each exit code vis-à-vis changes in policy in the DD analyses.

²² We are unable to determine the school that a student enters if they transfer within a year, but we can see the student’s school in the data the following year. This allows us to follow students across time and across schools and districts within the Louisiana public school system.

²³ All of these calculations include students who enter high school after they would normally be counted in high school graduation rates (i.e., “late entrants”). Also, note that some additional codes are not listed because they comprise less than .01 percent of exits. These include flags for transfer to early college, students effected by Hurricane Gustav and Ike, completed state skill, scholar – non-public, and others. However, we do include these codes in the calculations so that the estimates appropriately sum to one.

student-level longitudinal data system, is to verify that students have transferred between public schools within the state, which can be verified directly in the state data.

Three legitimate codes are particularly difficult for the state to verify: transfers out of state, transfers to private schools, and transfers to home study, which collectively comprise 11.3 percent of the all exits. The state requires schools to keep written justification for these codes²⁴ and carries out small-sample audits to check their validity. One recent audit was unable to verify large numbers of these exits, suggesting they are being used inappropriately (Louisiana Department of Education, n.d.).²⁵ When families seek to home school their children, they are required to fill out a form, but this form is so simple that some education leaders have expressed concern to us that schools are having families fill these out even when students are really dropping out. Schools seeking to reduce their denominator for graduation index and cohort graduation rate calculations could overstate the number of legitimate-unverifiable exits codes to increase their measured cohort graduation rates.

Because there are several important broad categories of exit, we organize the variables into six groups. We define three different forms of completion: on-time with regular diploma, delayed graduation with a regular diploma, and alternative completion such as GED (regardless of whether completion is late). Among the non-completion exits are legitimate exits and these are further separated into the verifiable and non-verifiable groups. All remaining exit codes are

²⁴ The federal government is unclear on what constitutes written justification. In private communication, LDOE indicates that, for transfers to home study (home schooling), schools are supposed to provide a copy of parents' applications for home. For transfers to private schools or other states, they are supposed to provide a written request for the students' records from the receiving school.

²⁵ The audit involved asking schools to submit paperwork for random subsets of students to whom these hard-to-verify codes were given. During the period of analysis, these audits were not done regularly. Since then, and at the time of this writing, the audits have been carried out annually. Also, over the past three years, the data for every student with a transfer to homeschooling exit code have been audited (not just a sample).

non-legitimate exit codes (mostly dropouts and missing exit codes), which schools have incentives to avoid as they do not remove students from the denominator of the graduation rate.

In Louisiana, schools can update student exit codes through October 1 of the subsequent academic year (four months after the school year ends), to address conflicts within the Student Information System (SIS). We therefore have to use the 2014-15 data to verify 2013-14 exit codes. Thus, the 2014 graduating class is the most recent available one, leaving an analysis sample that covers 9th graders for the years 2000-01 through 2010-11 (i.e., the 2004 through 2014 graduating classes). (Throughout the remainder of the paper, years refer to the spring so that 2013-14 is shortened to 2014.) We also limit the analysis to school districts that were less affected by Hurricanes Katrina and Rita since these events heavily influenced graduation rates in 2005.²⁶

III. Methods

III.A. National Cross-State Difference-in-Differences Analysis

We study the effects of accountability on graduation rates using differences-in-differences (DD). NCLB allowed each state to set its own graduation objective (instead of 100% proficiency), so we estimated a model that interacts the *Post* indicator with a continuous variable capturing the percentage of districts in the state that were below the state-determined graduation goal just before NCLB became law (P_S). This involves, first, calculating district graduation rates according to each state's own graduation rate definitions and, second, comparing these district graduation rates to each state's own NCLB-based graduation goals, yielding the percentage of

²⁶ There are only 68 regular districts operating continuously over this period. We included all districts in the calculation of Louisiana graduation trends, since this is what the state and federal government publicly report. Later, we restrict only to regular school districts.

treated districts in each state that are threatened with sanctions, which is a measure of the intensity of graduation accountability.²⁷ This yields the following:

$$Y_{st} = \alpha Post + \delta(P_s \times Post) + \rho X_{st} + \mu_s + \varepsilon_{st} \quad (1)$$

where Y_{st} is the graduation rate of the relevant state s in period t ; P_s is state percentage of districts that are below the state accountability standard; $Post$ is an indicator for the post-treatment (graduation accountability) period, which does not vary across geographic units within years; X_{st} is a set of time-varying independent variables pertaining to the government unit; μ_s is a vector of geographic unit fixed effects (state or district), and ε_{st} is a government unit-level iid random error.

In most cases, we are primarily interested in δ as an unbiased estimate of overall accountability pressure. However, $\hat{\delta}$ is only unbiased so long as P_s is conditionally exogenous; that is, the state governments did not set their accountability goals based on anything that is correlated with the change in the future graduation rate. This might be violated if, for example, more ambitious states institute additional programs at the same time that graduation accountability increases. On the other hand, if state-level “ambition” varies across states, but it is fixed over time within them, then this would not introduce bias.

We cluster the standard errors at the state level. Also, we weight states by their size so that the estimates are useful in gauging the contribution of accountability to the national increases in (low-stakes) high school graduation noted in Figure 1. To address the remaining limitations of

²⁷ The sources of the state graduation goals and definitions are Christie (2005), Swanson (2003), and eed.state.ak.us. Though we use the state’s own pre-NCLB definition of graduation to determine who is above and below the goal, we still use the AFGR as the dependent variable (Y_{st}) in (1). Also, for states with $P_s > 0.25$ (AL, AR, LA, SD, TN, and WY), we restrict treatment assignment to the bottom 25 percent of districts only, under the assumption that states are unlikely to place pressure on a larger share of districts.

the above approaches we also use a second method, which, as explained below, also estimates a somewhat different set of parameters.

III.B. National Cross-District, Within-State Analysis

The above cross-state analysis provides some evidence about the effects of graduation accountability, but the reliance on the share of threatened districts suffers from the fact that the share of districts threatened may be endogenous. We therefore add a second type of analysis using within-state variation in district graduation and proximity to NCLB adequate yearly progress thresholds. In section III.A, we explained how we determined whether districts were below or above the state-defined goal, to measure P_s . We use the same underlying information about districts in the district-level analysis (albeit without the above endogeneity problem).

The following model provides an estimate of the effect of graduation rate accountability policy using district-level variance:

$$Y_{dt} = \alpha Post + \delta(Low_d \times Post) + \rho X_{dt} + \mu_d + \varepsilon_{dt} \quad (2)$$

This is similar to (1) except that d indexes school districts (replacing the state index s) and μ_d therefore represents district fixed effects. In this case, Low_d is an indicator set to unity for districts whose graduation rate in the pre-NCLB era was below the state-determined post-NCLB goal. Here, the endogeneity issue is quite unlikely because it is the state that chooses the performance threshold and therefore assigns Low_d .

District fixed effects μ_d absorb state fixed effects, so these estimates come from within-state variation (no districts switched states). The identifying assumption is the same as (1): that the pressure faced by districts is accurately reflected in the proximity of their pre-NCLB

graduation rate to the state goal. We weight by district size and, again, cluster the standard errors at the state level (since the policy is still at the state level).²⁸

III.C. Louisiana Analysis

If there are effects of graduation accountability pressure, then the next issue is whether educator responses reflect strategic behavior. With more detailed data from the state of Louisiana, we can test for distortions in both the numerator and the denominator.

We continue to rely on a DD approach, but change the dependent variable and level of analysis. We start by using indicators for whether students have taken some version of credit recovery. As shown in equation (3), this is estimated at the student level (i). We focus here on low-graduate rate *schools* reflected by Low_k , with clustering by district.

$$Y_{it} = \alpha Post + \delta(Low_k \times Post) + \rho X_{kt} + \mu_k + \varepsilon_{it} \quad (3)$$

Credit recovery courses were first available in 2010 and graduation accountability increased during 2011-13, so we use data for the 2009-10 school year (pre-treatment) through 2015-16 (post-treatment).

We also analyze exit codes using the share of exits in each category (e.g., graduation, dropout, transfer), this time with data aggregated to the school level. This is reflected in equation (4)

$$Y_{kt} = \alpha Post + \delta(Low_k \times Post) + \rho X_{kt} + \mu_k + \varepsilon_{kt} \quad (4)$$

In both (3) and (4), we identify specific schools that are under threat of graduation-based sanction in Louisiana (instead of states or districts as in earlier analyses), yielding Low_k . Specifically, identify schools that were *ever* under the most pressure (those deemed failing at least once over the panel period) and compared them to schools ever under much less pressure

²⁸ Unweighted results are very similar and available upon request.

(those placed in the highest performance category at least once during the panel).²⁹ As above, we cluster at the district level since this is the lowest possible source of accountability pressure.

We estimate (4) with three different combinations of years and average across years (within the pre- and post-treatment periods) to reduce noise. Panel A in Table 7 averages the first two available years in our panel, 2000-01 and 2001-02, which occurred prior to any graduation accountability (and prior to all potential NCLB sanctions), and the post-treatment period averages the last two years (2012-13 and 2013-14). For these years our estimates of (4) reflect combined effect of both the early test-focused accountability and the subsequent graduation accountability.

To disentangle the effect of test-based accountability and graduation accountability, we further split the whole panel into two shorter ones. To measure the single effect of test-based accountability, Panel B maintains the first two years (2000-01 and 2001-02) and uses 2004-05 as the last year. To measure the single effect of graduation accountability, Panel C uses 2004-05 only as pre-treatment year, but maintains the last two years (2012-13 and 2013-14). We use only 2004-05 because this was the last year that clearly preceded graduation accountability and the later year, 2005-06, was an unusual year for exits with the hurricanes.

²⁹ In the early years, the top category was based on a five-star system. Later, it was changed to letter grades, based on School Performance Scores. The top schools in the analysis had five stars or a letter A at least once. The bottom category had one star or an F at least once. Schools with middle ratings are omitted. We use the ever-labeled approach to ensure a balanced panel and because the vast majority of schools ever threatened probably feel some pressure even in the years when they are not deemed failing.

IV. Results

IV.A. Cross-State Analysis

We begin by reporting the DD estimates of equation (1). Our baseline results use the average state AFGR for the 1999 and 2000 years as the pre-treatment period and the average of the 2009 and 2010 years as the post-treatment period. As noted earlier, the AFGR is the only measure that we can consistently measure pre- and post-NCLB; and 2010 is the last year the AFGR is available.

The results in Table 4 generally reinforce the idea that accountability has driven the increase in graduation rates shown in Figure 1. In Column (1), the point estimate on the interaction term shows that graduation rates increased more post-graduation-accountability in states with more districts below the state graduation threshold pre -graduation-accountability, though this estimate is statistically insignificant.³⁰

It turns out, however, that the results above average together two countervailing effects. As Dee and Jacob (2011) note, some states had test-based accountability pre-NCLB and others did not. For those without pre-NCLB test-based accountability, NCLB introduced both test- and graduation accountability at the same time. Moreover, we might expect that it was harder for schools to increase graduation if they were increasing test accountability at the same time. With test-based accountability, students are more likely to be retained in grade which might have made schooling less enjoyable for students.

³⁰ The specifications that include time-varying state covariates yield noticeably different results (in absolute value). We give little attention to these results because, upon closer examination, the covariate coefficients are implausibly large. This is not unexpected given the relatively smaller number of states, and the use of state fixed effects, which estimate the other covariates from (usually quite small) within-state changes over time.

The results are consistent with this theory. In columns (2) and (3), we split the sample into NTA states (those without test-based accountability prior to NCLB) and TA states (those with test-based accountability prior to NCLB), respectively. Column (2) shows that, in the NTA states, where NCLB introduced both forms of accountability at the same time, there was no disproportionate effect on states with more low-graduation-rate districts (measured by P_S , the percentage of districts in the state that were below the state-determined graduation goal just before NCLB became law). Conversely, in TA states, which experienced smaller changes in test-based accountability, graduation rates increase proportionally to graduation accountability pressure. The coefficient of 12.3 means that a state going from no districts threatened to all districts threatened would experience a 12.3 percentage point increase in the graduation rate.

The results in this section imply that graduation accountability likely increased graduation rates in states where many districts were below the graduation threshold and where the key policy change was the addition of graduation accountability only. The next section yields the same conclusion using a different method.

IV.B. Cross-District, Within-State Analysis

In this section, we leverage variation across school districts, within states. The results are consistent with the cross-state analysis. The descriptive results in Figure 2 show that districts that were low-performing during the pre-NCLB era saw larger increases in graduation rates than initially-high-performing districts. The DD results in Table 4 reinforce this. Column (4) shows positive and statistically significant effects on both the *Post* (post-graduation accountability) coefficient and the interaction term ($Low \times Post$), suggesting that the average district responded

to graduation accountability pressure, and low-performing districts, as predicted, responded more.³¹

The DD in columns (5) and (6) in Table 4 reinforce this, showing the same specification as in column (4), but splitting the sample by NTA and TA states.³² Both coefficients suggest positive although insignificant effects of *Post* for both sets of states. Though they are not statistically different from one another, the interaction term is larger in TA states.³³ (Unlike columns (1) these results are not sensitive to addition of state/district control variables such as school expenditures.)

These DD analyses point toward one key conclusion: graduation accountability increased high school graduation rates. None of the individual analyses on this point is persuasive on its own but consider: (a) graduation rates clearly increased faster post-graduation-accountability in low-performing districts than high-performing ones (Figure 2); (b) 11 of the 12 point estimates in Table 4 are positive and the only negative estimate is among the least precisely estimated; and (c) in the most well-identified estimates (Table 4, Columns 5 and 6 with district fixed effects), all four relevant point estimates (the intercept and interaction terms for both TA and NTA states) are positive and one is statistically significant. Moreover, the effects of graduation accountability are

³¹ We also considered carrying out the analysis in a regression discontinuity framework, to examine whether the effects on districts were larger near the threshold. However, such an analysis is likely to yield a false null because districts slightly above the threshold still faced pressure to remain just above the threshold; this is especially true given that most states expected gradual improvement over time in their graduation rates.

³² This is essentially a triple-difference, but we show the results as two separate DD columns for ease of comparability.

³³ The appendix also shows this visually in a figure that shows the graduation rate trends for four groups of districts: high-performing districts in NTA states, low-performing districts in NTA states, high-performing districts in TA states, and low-performing districts in TA states. If graduation accountability increases graduation rates, we would expect the effects to be largest in the low-performing/TA districts and smallest in the high-performing/NTA districts. Our results are also consistent with this pattern. As the DD analysis suggests, the figure shows increases in graduation rates in both TA and NTA states, but especially in low-performing/TA districts.

larger in states that experienced mostly an increase in graduation accountability, not test-based accountability.

All of the estimates in Table 4 pass parallel trends tests. This can be seen visually in the figures, but also in the formal tests shown in the middle of the Table 4. The results are also robust to switching the pre-treatment years from 1999-2000 to the Dee and Jacob (2011) baseline of 2001-2002. Nevertheless, one caveat is that the different groups of districts and states have varying baseline graduation rate levels. This could, in theory create three methodological issues: (a) ceiling effects; (b) regression to the mean; and (c) endogenous policy adoption. Ceiling effects do not seem likely in this case since none of the subgroups is near the 100 percent graduation ceiling. It might be easier to make gains when the initial level is near 50 percent, but this is not so much a potential source of bias as a contextual factor. If it were easy for districts to make gains from, say, 60 percent, then they could have already done so without graduation accountability. Regression to the mean is also an unlikely influence because we averaged two pre-graduation-accountability years together. Finally, while policy adoption is almost certainly endogenous, there is no reason to believe that this is true in a time-varying way. States adopting accountability on their own were no doubt different in unobserved ways, but, so long as these differences are time-constant, the state/district fixed effects account for them.

IV.C. Graduation Rates and Graduation Exams

In this section, we address one of the main reasons for skepticism about rising graduation rates: that schools may have responded to accountability simply by making it easier for students to graduate. Since accountability could induce this form of strategic behavior in the numerator, all of the graduation measures discussed in section II are potentially suspect.

To provide a rough test of whether graduation accountability led to diminished academic

standards, we leverage state-level variation in graduation exams. If graduation exams make it more difficult for schools to lower standards, and if accountability induced states to try to lower standards, then the lower-graduation-standards theory predicts that we should see smaller graduation rate increases in states that had graduation exams during the period of the above analysis. To test this, we identified the presence of state graduation exams in the years 2000 (Goertz & Duffy, 2001), 2010 (various web sites³⁴), and 2017 (Gewertz, 2019).

Table 5 shows the CPS trends in status high school completion, for states with differing combinations of graduation exams. We are mainly interested in states that had graduation exams throughout the 2000-2017 period (N=4) and states that had no graduation during 2000-2017 (N=18). The table also considers states that had a graduation exam in 2000, but not in 2017 (N=7); states that added a graduation exam (N=1), and states that alternated from having exams in 2000, dropping them in 2010, and adding them back by 2017 (N=3). The pattern of graduation rates, however, is inconsistent with the theory that graduation accountability led schools to relax their graduation standards. Graduation rates actually increased *faster* in states that had graduation exams throughout 2000-2017 than in those that never had such exams.

While these results run against the lower-graduation-standards theory, the analysis does have some limitations. First, the graduation exams might not really have created a barrier to graduation, given the available exceptions that allow students to graduate even if they fail the exams (Achieve, 2010). For example, most states have exceptions for special education students and students had multiple opportunities to re-take the exam. We view this as a somewhat unlikely explanation for the results in Table 5, however. Research consistently suggests that

³⁴ The information on graduation exams for 2010 comes from the web sites of state departments of education and the following organizations: Center for Education Policy, Education Commission of the States, and Education Week.

graduation or exit exams reduce graduation rates particularly for academically and socially disadvantaged students (Dee & Jacob, 2006; Ou, 2009; Papay, Murnane, and Willett, 2010; Warren et al., 2006).³⁵

Another version of the above problem is that passing the graduation exams is so easy that schools could lower their graduation standards even with the exam. Some reports suggest that the passing graduation exams only requires skills of an 8th grade education (Achieve, 2010). While this might seem like a low standard, note that, as roughly three-quarters of students already graduated pre-NCLB, the marginal student, whose graduation probability might be influenced by accountability, likely has very low academic achievement; passing even an 8th-grade-level exam might be difficult for them. Also, again, the fact that graduation exams by themselves clearly reduce graduation rates (Dee & Jacob, 2006; Ou, 2009; Papay, Murnane, and Willett, 2010; Warren et al., 2006; Hemelt & Marcotte, 2013) means that the exams do create a binding constraint.³⁶

While we can largely reject the above explanations for the Table 5 results, there are three others that are more difficult to dismiss: (a) Since the graduation exams are usually taken in 10th or 11th grade, standards could have been relaxed in later grades, after students passed the test; (b) graduation exams do not cover all the subjects required to graduate and therefore the standards might have been relaxed in courses not required for graduation; and (c) this analysis is only descriptive and might not reflect the effects of graduation exams. We considered collecting additional data, and carrying out additional difference-in-differences analyses, leveraging the

³⁵ Three studies find that high school dropout rates tend to increase by roughly 1 to 2 percentage points, on average, when states implement rigorous exit exams (Dee & Jacob, 2006; Warren et al., 2006; Hemelt & Marcotte, 2013). Three studies find that these effects are concentrated among black and low-income students (Dee and Jacob, 2006; Ou, 2009; Papay, Murnane, & Willett, 2010).

³⁶ The introduction of student accountability (i.e., graduation exams) could lead schools to respond differently to school accountability. However, both the pre- and post-NCLB studies (in the latter case, especially, Hemelt and Marcotte, 2013) both show that graduation exams reduced graduation rates.

changes in state graduation policy; however, we decided against this because of the inadequacies in our data regarding the exact timing of the graduation exam starting years, the lack of national micro data, and the fact that other studies on graduation exams have provided much of the useful evidence we need.

The leveraging of variation in graduation exams is therefore informative, but not definitive. In what follows, we consider additional approaches using more detailed data from a single state.

IV.D. Graduation versus GEDs

One of the most significant points of contention in the analysis of graduation rate trends has been the distinction between high school graduation and GEDs (Heckman & LaFontaine, 2007; Mishel & Roy, 2006). GEDs are generally considered to be less rigorous than diplomas, especially with regard to non-cognitive skills (Cameron & Heckman, 1993), even if they do seem to benefit those who obtain them relative to those who drop out with similar baseline skills (Murnane, Willett & Tyler, 2004). This, combined with the lower amount of seat time involved with GEDs, suggests that these credentials involve a lower standard.

As noted in section II, all of the high-stakes graduation rate measures exclude GEDs. Therefore, we would expect that graduation accountability shifted students from GEDs to high school diplomas, increasing human capital. This is what we see. The number of GEDs dropped after the passage of NCLB (NCES, 2010) and have declined further since (Barshay, 2018).³⁷ As shown in the appendix, the ratio of regular high school diplomas to GEDs to increased from 5.8 to 7.1, suggesting that the upward trend in completion is not due to declining credential quality, and might reflect improving quality. In this respect, the CPS trends in Figure 1 understate the

³⁷ As Barshay (2018) points out, there were also cuts in adult education programs that may have contributed to the decline in GEDs.

increase in human capital, and provide further evidence that the rise in other graduation measures does not reflect lowering standards.

IV.E. Louisiana Analysis

In this section, we continue our analysis of the mechanisms of graduation rate increases using student-level data from the Louisiana Department of Education. In particular, we provide more details on potential shifts across credentials (distortion in the numerator) and incorrect exit codes (distortion in the denominator). We begin by describing the evolution of the policy in the state, which is necessary to interpret the results that follow.

IV.E.1. Louisiana School Accountability

The evolution of graduation rate policies in Louisiana largely tracks the changes in federal policies and requirements. The state instituted school-level academic performance metrics for elementary and middle schools in 2003 and extended this to high schools three years later. For all grade levels, these original measures focused mainly on student performance on standardized tests (student attendance was also included for elementary schools). At the high school level, the dropout rate was included, though it accounted for only 10 percent of the overall performance score.³⁸

In 2005-06, the Louisiana Board of Elementary and Secondary Education (BESE) announced that, beginning in 2006-07, the dropout rate would be replaced by a graduation index accounting for 30 percent of the School Performance Score (SPS). While schools earned maximum points for students receiving regular diplomas they also were awarded a smaller number of points for students who finish their GEDs, obtain a certificate of achievement, or

³⁸ Schools with 8th grades were also held accountable for the future dropout and academic performance of their students, but this is beyond the scope of our analysis of high schools. The remainder of the performance metrics focused almost entirely on test scores, which was the main focus of NCLB.

remain enrolled in that school.³⁹ Importantly, the graduation *index* was included as part of the accountability system, but the graduation *rate* was not.

Beginning in 2011-12, the state introduced the graduation rate as part of accountability for the first time, giving an “adjustment” or bonus points to schools that had graduation rates above 80 percent. Schools with rates below the state average saw reductions in points.⁴⁰ In 2012-13, the state added the simple cohort graduation rate, comprising 25 percent of the SPS, while reducing the graduation index weight to 25 percent.⁴¹ The implication was that fully 50 percent of the high school SPS was related to graduation.⁴²

The above discussion implies two key differences between the Louisiana graduation index and the cohort graduation rate. First, the cohort graduation rate counts only regular on-time diplomas, while the Louisiana graduation index gives some, albeit lesser, weight to GED and other forms of completion. Second, the graduation index includes not only that cohort but other students who had late graduation and/or graduation with alternative credentials.⁴³ This means Louisiana high schools experienced a large increase in incentive to increase all forms of credentials in 2007, which continued throughout the panel period. Moreover, with the second policy change in 2013, there was a sharp increase in incentive to produce on-time, regular

³⁹ For 2006-07, the graduation index would be calculated using data from that year, but moving forward the index would use data from the prior two years, meaning that the 2008-09 index would use data from 2006-07 and 2007-08. This is because student graduation data and exit codes are confirmed and updated until October 1st of the following academic year.

⁴⁰ The precise number of adjustment points is unclear. Schools below the state average also became ineligible for the “top gains” designation that was calculated based on a value-added-like measure, using student test scores.

⁴¹ Schools are also accountable for students that are new to the school in 10th grade so long as that school is in the same LEA as the student’s 9th grade school.

⁴² The federal government also required the state to use subgroup cohort graduation rates for accountability. This was not part of the Louisiana SPS, however. Rather, schools with insufficient subgroup graduation rates were not allowed to receive “progress points.”

⁴³ The graduation index is based on graduates in a given year, regardless of their cohort, whereas the cohort graduation rate is from one cohort, which is partly why this difference emerges. The state still uses the cohort of first-time 9th graders as the denominator for the graduation index even some of the students counting toward the index in any given year are from different cohorts.

diplomas, relative to other forms of completion, as they are the only ones that count in the cohort graduation rate and receive the most points in the graduation index.

In addition to comprising a larger share of the SPS, the stakes associated high school graduation also increased over time. The state went beyond the requirements of NCLB and, in 2003, created the Recovery School District (RSD) to oversee the improvement of schools deemed low-performing. Louisiana also switched to an A-F letter-grading system in 2010-11 that drew more attention to academic outcomes. While the vast majority of the RSD's efforts have been concentrated in New Orleans, some schools were sanctioned in other districts and, theoretically, all schools in the state were under threat of RSD sanction, including outright state takeover.

In short, in just a six-year period (2005-06 to 2012-13), Louisiana increased the share of graduation accountability from 10 to almost 50 percent of the SPS and increased the stakes attached to these measures. In some respects, this change was even more sudden and intense than it was for test scores since the tests had long been a fixture of state accountability even before NCLB. Graduation accountability was almost entirely new in 2007.

Despite some distinctive features of the state's graduation accountability policies, Louisiana's high school graduation rate followed the national upward trend during the 2000s. From 2005-06 to 2017-18, the ACGR increased from 64 percent to 81 (Perry et al., 2015; LDOE, 2019). Since we are trying to explain a similar national trend, this makes the state a useful case study. We use the Louisiana's micro data to understand to what degree this increase might have been caused by strategic behavior.

IV.E.2. Strategic Behavior in Louisiana: The Numerator and Credit Recovery

As noted in section II, we imputed credit recovery courses for individual students using the Louisiana micro data. In Table 6, we estimate equation (3) using that imputed measure as the dependent variable, using data from just before credit recovery became available (2009-10) and after both the start of credit recovery and the increase graduation accountability (2015-16). The results show no evidence that non-threatened schools responded by increasing the use of credit recovery. The point estimates on the *Post* term are small in magnitude and statistically insignificant. However, threatened schools appear to have increased the probability that students took at least one credit recovery course by 4-5 percentage points.

The effects are concentrated in math. This is most likely because when Louisiana increased the importance of graduation for school accountability, they also changed the requirements for graduation. While students were always required to complete four English courses to graduate, they were now also required to complete (and pass) four math classes to graduate instead of three. If schools that are close to the accountability threshold are engaging in strategic behavior, we may see an increase in the number of students taking more than one math class in their senior year to make up any previously failed math course. The results, presented in Table 6, are consistent with this hypothesis; the effects show in math but not reading.

Given the possibility that our measure of credit recovery might be inadequate, or that course re-takes might be driven by factor we have not accounted for, we estimate the models in two different ways. The results reported in Table 6 include only 11th and 12th graders who are most likely to need to re-take a course in order to graduate. We see no effects on 9th and 10th graders (available upon request).

The size of the effects in Table 6, however, are too small to explain the increase in statewide graduation rate. This is because: (a) only half of students are in the threatened schools where the credit recovery increase is concentrated; and (b) the increase in the graduation rate in Louisiana was even larger than in other states (9.2 percent versus 6.1 percent, respectively).

IV.E.3. Strategic Behavior in Louisiana: Exit Codes

Since we are interested in graduation rates, we could analyze exits based on cohorts of 9th graders from four years earlier, to be consistent with the way graduation rates are now reported. In this case, however, we have more precise data and therefore focus on the percentage of exit codes in each category by *year*.⁴⁴ If schools changed the way they coded exits as a result of the accountability changes in time t , then we would expect them to change their procedures for all students also in time t , regardless of which cohort they happen to be in.⁴⁵

Analyzing data based on the share of exit codes in a given year, rather than a given cohort also isolates policy effects more clearly. By construction, graduation rates can be slow to respond to policy shifts. Take NCLB, which passed in 2001. Even if the law had an immediate effect on schools and the probability of graduation, it might not show up because the ACGR and NCES-GR for the graduating class of 2002 incorporates 9th and 10th grade dropouts from 1999 and 2000. We could assume that these earlier dropouts were exogenous with respect to the policy change, but, even if true, this is going to add measurement error and that reduces statistical power.

⁴⁴ A disadvantage of this approach is that the numerator and denominator become mechanically linked. If there are fewer exits in category X, then the denominator, which counts the total number of exits also declines. In future research, we will also report results by share of enrollment, which reduces this mechanical relationship, albeit by sacrificing precision in the timing of effects.

⁴⁵ Schools might have the strongest incentive to re-code 12th grade students since this is the group that will show up in reported graduation rates earliest, but the current year codes for, say, 10th grades would still show up in the graduation rates (and SPS) two years later. Also, since changing the coding procedures is relatively costless, there is no reason to think that the discount rate would be relevant.

Recall that strategic behavior with graduation rates could show up either in the numerator of the rate or the denominator. We start with some descriptive evidence of the former. Figure 3 Panel A shows the share of all exit codes that are completions. The most obvious pattern in almost all the figures is the spike/drop in the 2005-06 year. This is mainly because of Hurricanes Katrina and Rita, which affected almost the entire state and shut down schools on the gulf coast for months. For this reason, we generally focus on the difference between the pre-hurricane level/trend and the most recent years in the remainder of the analysis. We also focus just on non-hurricane-affected districts to avoid conflating graduation accountability effects from hurricane effects.

The results are largely as predicted from the national results. From 2005 (just before graduation accountability ratcheted up in the state) to 2014 (after graduation accountability), Figure 3, Panel A shows that on-time regular diplomas increased by 6 percentage points. (Note, again, that these cannot be viewed as graduation rates because we are focusing on years not cohorts.) Alternative and delayed credentials saw smaller changes.

Panel B focuses on non-completions. From 2005-2014, the share of legitimate non-verifiable exits, which boost graduation rates, increased by about two percentage points providing some evidence of strategic behavior. In contrast, the share of exits that were non-legitimate (i.e., those counting against schools for accountability purposes) saw a steep decline of almost 10 percentage points.

Table 7 Panel A reports the DD results, comparing schools that were threatened with sanction with those that were above the state threshold, for the entire period, 2001-2014.⁴⁶ For

⁴⁶ In Louisiana, we defined a threatened school as one that had received an F letter grade at least once in the pre-treatment period. The comparison group is schools receiving an A grade at least once. Schools in between the extremes are omitted.

non-threatened schools, the first-difference results suggest an increase in on-time graduation (4.9 percentage points) and slight increase in delayed graduation (0.7 percentage points), offset slightly by a decline in alternative completion (-0.6 percentage points). The net increase in completion is five percentage points.

Only some of this increase reflects a decline in dropout, however, which declined by only 1.7 percentage points. We also see some direct evidence of strategic behavior. Legitimate, but hard-to-verify, exits increased by 1.9 percentage points (insignificant) and the missing exit code rate dropped by 1.8 percentage points. These shifts suggest roughly equal changes in human capital, strategic behavior, and data quality improvement, at least for non-threatened schools.

Similar to the national analysis in Table 4, we also consider how the patterns differed for low-performing schools, threatened sanction. Perhaps surprisingly, the interaction term suggests that threatened schools saw sharp declines in on-time/regular graduation (relative to the non-threatened schools), offset by a similarly-sized increase in alternative graduation. The two coefficients largely offset one another, however, suggesting the accountability effect on completion was the same in threatened and non-threatened schools, but the allocation across forms of completion differed; in non-threatened schools we see more on-time graduation and in threatened schools, the increase is in alternative completion.

The pattern above most likely reflects the design of the Louisiana accountability system, which gave considerable weight to non-regular diplomas. (The cohort graduation rate, which only counted regular diplomas, did not begin until 2013, just before the end of the panel.) Given the very low achievement levels in these threatened districts, this may have been the most effective path for increasing measured graduation rates.

The results are also suggestive that strategic behavior increased more in threatened schools. The coefficients on hard-to-verify exit codes for threatened school is positive (though insignificant) while the coefficient on missing exit codes is large and negative; though both of these are statistically insignificant. This may suggest that accountability induced schools to pay closer attention to record-keeping. In any event, a decline in the number of missing codes by itself does not imply any increase in human capital.

Panels B and C show the results restricted to the early (2001-2005) and later (2005-2014), respectively. This is an important distinction because the state's accountability reforms came in two phases: the introduction of graduation accountability came during the 2003-2005 period and the percent of the SPS from graduation increased, along with accountability intensity, during 2012-2013. (In the latter case, recall that the graduation index and cohort graduation rate were included in the SPS and these included regular graduation, delayed graduation, and alternative completion.)

For non-threatened schools, it appears there were few responses in the first phase of accountability (top row of Panel B), but the increase in weight on the graduation rate to 50 percent in the second phase apparently got their attention. For threatened schools, the effects are relatively similar across the two periods, with somewhat larger effects from the first phase. This is especially true for alternative completion, which carried somewhat greater weight in the first phase than they did later.

One possibly surprising finding is the (insignificant) increase in dropouts in threatened schools, in both the early and the later periods. One potential explanation is highlighted by an additional finding: a reduction in missing exit codes. It is possible that schools simply improved their record-keeping, and started reporting dropouts more accurately, and reducing missing exit

codes. Another possibility is that the increase in requirements for receiving a regular diploma, but this occurred in the later years and could not explain the increasing dropout in the early years.

V. Conclusion

A common concern with all forms of high-stakes accountability is that the measures tied to incentives will become distorted. While this problem has been well-recognized with high-stakes testing, much less attention has been paid to the similar problems that arise with high school graduation rates. Given that the data underlying graduation rates are controlled almost completely by schools and that the data are subject to only limited external validation, the potential threat to validity is real.

Our analysis confirms that the positive correlation between the passage of NCLB and the rise in the graduation rate was due to graduation accountability. As we would predict, graduation rates increased most in the states that had a higher proportion of districts below the state graduation rate threshold; and in the districts (within states) that were below the threshold.

The key contribution of the study, however, is showing that, despite the potential for significant distortion, at least some, and perhaps a substantial share, of the post-accountability graduation rate rise reflects human capital increases. Graduation rates increased faster in states that had graduation exams than those without; regular diplomas increased and GEDs decreased; and even the worst-case scenario with credit recovery courses seem to explain only a small fraction of the overall graduation rate increase (at least in Louisiana). Moreover, this graduation rate increase was achieved even though all schools were “swimming against the tide” of strong test-based accountability at all grade levels, and schools in many states also maintained graduation exams. In other words, we see limited evidence that the rising graduation rate is due

to lower standards or less seat time. While there is some evidence of distortion in the denominator, this affects school and district rates much more than state and national ones.

Though we have examined the graduation rate in numerous ways, with multiple methods and data sources, there are still uncertainties that could be usefully addressed with future research. It remains unclear, for example, to what degree other states have followed suit with Louisiana in carrying out detailed audits of the exit code data; or, more broadly, the degree to which the Louisiana results mirror other states. We also lack direct data on credit recovery programs and instead had to rely in proxies. Finally, there are also other ways schools could be lowering standards, such as simply making it easy for students to pass their courses.

To what degree is the rise in graduation rates “real”? Given the key contributions that high school graduation clearly plays in affecting student outcomes, and opening up opportunities in post-secondary education, this is a key question. We find that the recent and fast rise in graduation rates reflects some strategic behavior, but the preponderance of evidence suggests a large increase in the nation’s stock of human capital. This appears to reflect some real success of graduation accountability, and perhaps the clearest success of federal school accountability documented to date.

References

- Achieve, Inc. (2004). *Do Graduation Tests Measure Up? A Closer Look at State High School Exit Exams*. <https://www.achieve.org/files/TestGraduation-FinalReport.pdf>
- Barshay, Jill (2018). GED and other high school equivalency degrees drop by more than 40% nationwide since 2012. *Hechinger Report*. Downloaded January 31, 2020 from: <https://hechingerreport.org/ged-and-other-high-school-equivalency-degrees-drop-by-more-than-40-nationwide-since-2012/>.
- Cameron, Stephen V. & James J. Heckman (1993). The Nonequivalence of High School Equivalents. *Journal of Labor Economics* 11(1): 1-47.
- Christie, Kathy (2005). *Target Attendance and Graduation Rates and How Rates are Calculated*. Education Commission of the States Periodic Report.
- Dee, Thomas S., & Brian A. Jacob (2006). Do high school exit exams influence educational attainment or labor market performance? *NBER Working Paper Series Working Paper 12199* Cambridge, MA: National Bureau of Economic Research.
- Dee, Thomas S. & Brian A. Jacob (2011). The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management*, 30(3): 418-446.
- Deming, David J., Sarah Cohodes, Jennifer Jennings & Christopher Jencks (2016). School Accountability, Postsecondary Attainment and Earnings. *The Review of Economics and Statistics*, 86(5): 848-862.
- Deming, David J. & David N. Figlio (2016). Accountability in US Education: Applying Lessons from the K-12 Experience to Higher Education. *Journal of Economic Perspectives*, 30(3): 33-56.
- Duncan, Arne (2018). Education reform has worked. Here's proof. *Washington Post*. <https://www.chicagotribune.com/opinion/commentary/ct-education-reform-charters-arne-duncan-20180402-story.html>.
- Gewertz, C. (2019). Which States Require an Exam to Graduate? *Education Week*. Downloaded February 1, 2020 from: <https://www.edweek.org/ew/section/multimedia/states-require-exam-to-graduate.html>.
- Figlio, David N. & Joshua Winicki (2005). Food for thought: the effects of school accountability plans on school nutrition. *Journal of Public Economics*, 89(2-3): 381-394.
- Goertz, Margaret E. & Mark C. Duffy. 2001. *Assessment and accountability Systems in the 50 states: 1999-2000*. Consortium in Policy Research for Education and University of Pennsylvania.

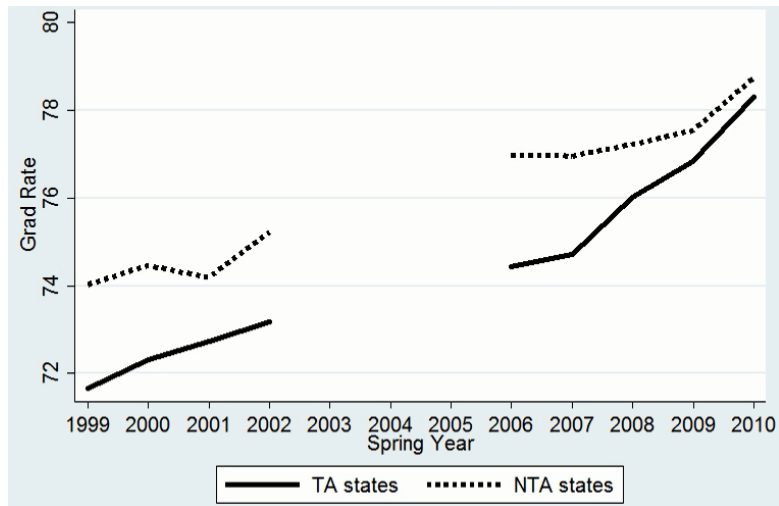
- Gordon, Robert J (2015). Secular Stagnation: A Supply-Side View. *American Economic Review*, 105(5): 54-59.
- Goldin, Claudia D. & Lawrence F. Katz (2008). *The Race between Education and Technology*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Hanushek, Raymond & Steven G. Rivkin (2010). The Quality and Distribution of Teachers under the No Child Left Behind Act. *Journal of Economic Perspectives* 24(3): 133-150.
- Harris, Douglas N. & Carolyn Herrington (2006). Accountability, standards, and the growing achievement gap: Lessons from the past half-century." *American Journal of Education*, 112(2): 209-238.
- Heckman, James J., & Paul A. LaFontaine. 2006. Bias Corrected Estimates of GED Returns. *Journal of Labor Economics*, 24: 661–700.
- Heinrich, Carolyn J., Darling-Aduana, J., Good, A., & Cheng, H. (2019). A Look Inside Online Educational Settings in High School: Promise and Pitfalls for Improving Educational Opportunities and Outcomes. *American Educational Research Journal*.
- Hemelt, Steven & Marcotte, David. High School Exit Exams and Dropout in an Era of Increased Accountability. *Journal of Policy Analysis and Management* 32(2): 323–349.
- Huntsberry, Will. Why Graduation Rates in San Diego — and Across California — Went Down. *Voice of San Diego*, August 27, 2018.
<https://www.voiceofsandiego.org/topics/education/why-graduation-rates-in-san-diego-and-across-california-went-down/>.
- Hurst, David, Alexandra Tan, Anne Meek a& Jason Sellers. 2003. *Overview and inventory of state education reforms: 1990 to 2000*. National Center for Education Statistics Publication 2003020.
- Jackson, C. Kirabo (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *Journal of Political Economy*. 126(5): 2072-2107.
- Jacob, Brian (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5-6): 761-796.
- Jewson, Marta. (2019). 29 New Orleans seniors in grade-fixing scandal won't graduate despite summer work, promises. *The Lens*, August 10, 2019.
https://www.nola.com/news/education/article_448bccee-bbbe-11e9-9a99-8b5a5c226c5c.html
- Koretz, Daniel (2009). *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press.

- Louisiana Department of Education (n.d.). *Graduation Exit Code Pre-Reviews*. Downloaded May 13, 2018 from: <https://www.louisianabelieves.com/docs/default-source/data-management/final-exit-code-pre-reviews.pdf?sfvrsn=2>.
- Louisiana Department of Education (2019). Class of 2018, Press Conference, May 15, 2019 <https://bloximages.newyork1.vip.townnews.com/theadvocate.com/content/tncms/assets/v3/editorial/6/29/629b7698-7772-11e9-85a5-e3868dc43c80/5cdcb065cd848.pdf.pdf>.
- McGee, Kate (2018). In D.C., 34 Percent Of Graduates Received A Diploma Against District Policy.” *NPR*. January 29, 2018. <https://www.npr.org/sections/ed/2018/01/29/581036306/in-d-c-thirty-four-percent-of-graduates-received-a-diploma-against-district-poli>
- Mishel, Lawrence & Joydeep Roy (2006). *Rethinking High School Graduation rates and Trends*. Economic Policy Institute. Washington, DC.
- Murnane, Richard J. (2013). U.S. High School Graduation Rates: Patterns and Explanations. *Journal of Economic Literature*, 51(2): 370-422.
- Richard J. Murnane, John B. Willett and John H. Tyler (2000). Who Benefits from Obtaining a GED? Evidence from High School and Beyond. *The Review of Economics and Statistics*, 82(1): 23-37.
- National Center for Education Statistics (2010). Trends in High School Dropout and Completion Rates in the United States: 1972–2008, Table A-2. Percentage distribution of persons who passed the General Educational Development (GED) exam outside of federal and state contract facilities¹, by age group: 1998–2008. https://nces.ed.gov/pubs2011/dropout08/tables/table_a02.asp.
- National Center for Education Statistics (2017). What is the difference between the ACGR and the AFGR? Downloaded January 30, 2019 from: <https://nces.ed.gov/blogs/nces/post/what-is-the-difference-between-the-acgr-and-the-afgr>.
- OECD (2012). *Education at a Glance 2012: OECD Indicators*. OECD Publishing, Paris, <https://doi.org/10.1787/eag-2012-en>.
- Ou, D. (2009). *To Leave or Not to Leave? A Regression Discontinuity Analysis of the Impact of Failing the High School Exit Exam*. London: Centre for Economic Performance, London School of Economics.
- Papay, J. P., Murnane, R. J., & Willett, J. B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from Massachusetts. *Educational Evaluation and Policy Analysis*, 32, 5–23.
- Perry Andre, Douglas N. Harris, Christian Buerger, & Vicki Mack (2015). *The New Orleans Index at Ten: The Transformation of New Orleans Public Schools: Addressing System-level Problems Without a System*. New Orleans, LA: The Data Center.

- Reback, Rendall, Jonah Rockoff & Heather L. Schwartz (2014). Under Pressure: Job Security, Resource Allocation, and Productivity in Schools under No Child Left Behind. *American Economic Journal: Economic Policy* 6(3): 207-241.
- Rouse, Cecilia Elena, Jane Hannaway, Dan Goldhaber & David Figlio (2013). Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure. *American Economic Journal: Economic Policy*, 5(2): 251-281.
- St. George, Donna (2018). How high can graduation rates go? The story of one school rocked by scandal. *The Washington Post*. July 27, 2018.
https://www.washingtonpost.com/local/education/how-high-can-graduation-rates-go-the-story-of-one-school-rocked-by-scandal/2018/07/27/98a3c34e-3ab0-11e8-8fd2-49fe3c675a89_story.html?utm_term=.6d924767d29e
- St. George, Donna St. & Lynh Bui (2017). Probe finds late grade changes for 5,500 in Prince George's. *The Washington Post*. November 3, 2017.
https://www.washingtonpost.com/local/education/probe-finds-late-grade-changes-for-5500-in-prince-georges/2017/11/03/5e54e10c-be62-11e7-959c-fe2b598d8c00_story.html?utm_term=.40b33d73ff97
- Swanson, Christopher B., 2003. *NCLB Implementation Report: State Approaches for Calculating High School Graduation Rates*.
- U.S. Department of Education (2008). *High School Graduation Rate: Non-Regulatory Guidance*. Retrieved from <http://www2.ed.gov/policy/elsec/guid/hsgrguidance.pdf>.
- Warren, John R., Krista N. Jenkins, & Rachael B. Kulick (2006). High school exit examinations and state-level completion and GED Rates, 1972–2002. *Educational Evaluation and Policy Analysis*, 28: 131–152.

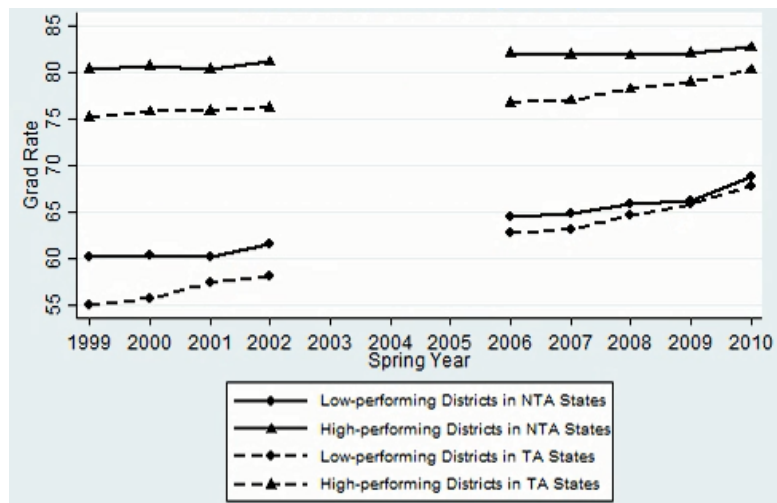
Appendix A: Graduation Rate Trends by State Types

Figure A1: Graduation Rate Trends, by TA and NTA States



Notes: The figure reports the Average Freshman Graduation Rate (AFGR) weighted by state enrollment size by year. The NTA states include those states that had no test-based accountability prior to NCLB; the TA states had pre-NCLB test-based accountability, so that the federal law mainly added graduation accountability. Essentially no states had significant graduation accountability prior to NCLB.

Figure A2: Graduation Rate Trends by State and District Types



Notes: The figure reports the weighted AFGR. The dashed lines represent our TA states only (i.e., states that had pre-NCLB test-based accountability), and separates districts into groups based on whether their pre-NCLB graduation rate were below (low-performing) or above (high-performing) the state-determined post-NCLB goal. The solid lines do the same for NTA states (i.e., states that had no pre-NCLB test-based accountability).

Appendix B: Trends in GEDs

Table B1 summarizes the trends in GEDs, in total and dividing by the number of graduates by the number of regular high school graduates. Note that the GED figures include both recent high school graduates and students who are much older; for this reason, the percent of recent high school completers who received GEDs is much lower than this implies. (The data do not allow us to distinguish the age ranges of GED recipients.)

As noted in the text, the number of high school graduates has risen faster than the number of GED recipients, suggesting that credential quality has improved.

Table B1: Trends in GEDs

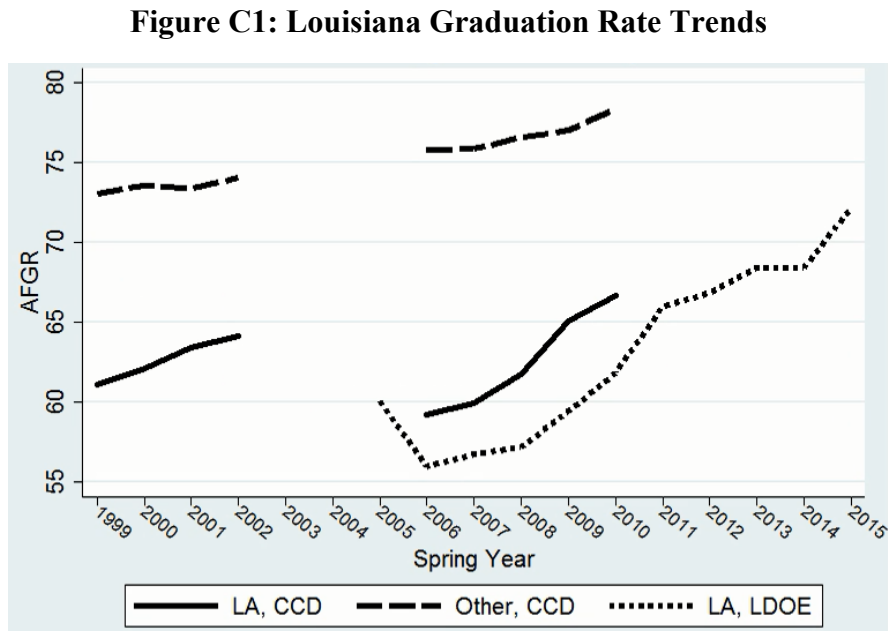
| Year | Num. GED (unit: thousand) | Num. High school graduates (unit: thousand) | Ratio Regular vs. GED |
|------|---------------------------------|---|-----------------------------|
| 1998 | 480 | 2,704 | 5.6 |
| 1999 | 498 | 2,758 | 5.5 |
| 2000 | 486 | 2,832 | 5.8 |
| 2001 | 648 | 2,847 | 4.4 |
| 2002 | 329 | 2,906 | 8.8 |
| 2003 | 387 | 3,015 | 7.8 |
| 2004 | 405 | 3,054 | 7.5 |
| 2005 | 423 | 3,106 | 7.3 |
| 2006 | 398 | 3,122 | 7.8 |
| 2007 | 428 | 3,199 | 7.5 |
| 2008 | 467 | 3,312 | 7.1 |

Table Notes: The GED figures are based on NCES data
(https://nces.ed.gov/programs/digest/d17/tables/dt17_219.10.asp).

Appendix C: Additional Detail on Louisiana Graduation Rates

Figure C1 reports trends in the Louisiana high school graduation rate in three ways. The first (dashed) line is the national AFGR using the CCD, from Figure 1. The second (solid line) reports the AFGR using the federal CCD data for Louisiana only. Finally, a third (dotted) line is based on an attempted replication of the AFGR using the micro data.⁴⁷

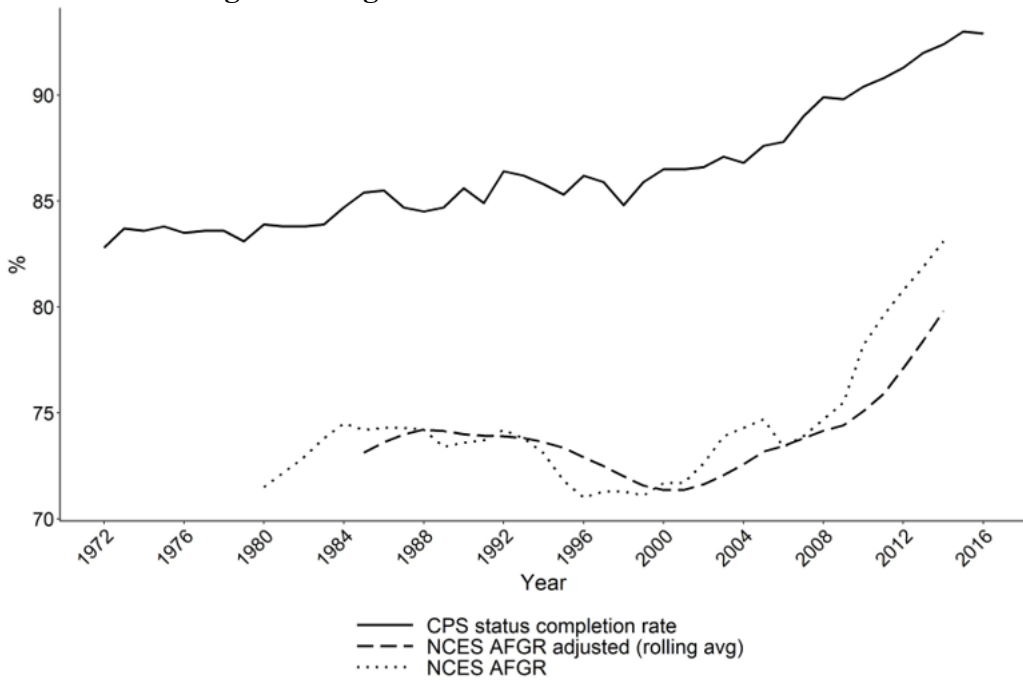
The micro data version seems to track the CCD data for Louisiana after 2010, but is well below it during the years 2005-2010. We hypothesize that this is because of Hurricanes Katrina and Rita which struck the state in the fall of 2005. In additional analysis, we found that the graduation rate denominator during 2005-2010 was larger in the micro data, and the numerator was smaller. We are exploring possible reasons, including that the federal government did not require the state to include all graduates in those years, as the state recovered.



Notes: This figure reports the AFGR for Louisiana and the rest of the country using the USDOE-CCD data and our calculation of the AFGR from the LDOE micro data. The LDOE version is lower than the CCD version in part because the latter excludes students who moved out of state.

⁴⁷ As can be seen in other data sources, the Louisiana graduation rate is 10-15 percentage points below the national average in most years.

Figure 1: High School Graduation Rate Trends



Notes: The solid line reports the CPS status completion rate for adults aged 18-24, based on household surveys. The dotted NCES lines reports the average graduation rate constructed from data reported by states to the U.S. Department of Education. Given that the CPS reports graduation for adults in the 18-24 age range, we also adjusted the NCES line to be a rolling average of the prior six years for comparability, presented in the long-dashed line.

Figure 2: Graduation Rates Trends by *District Types*

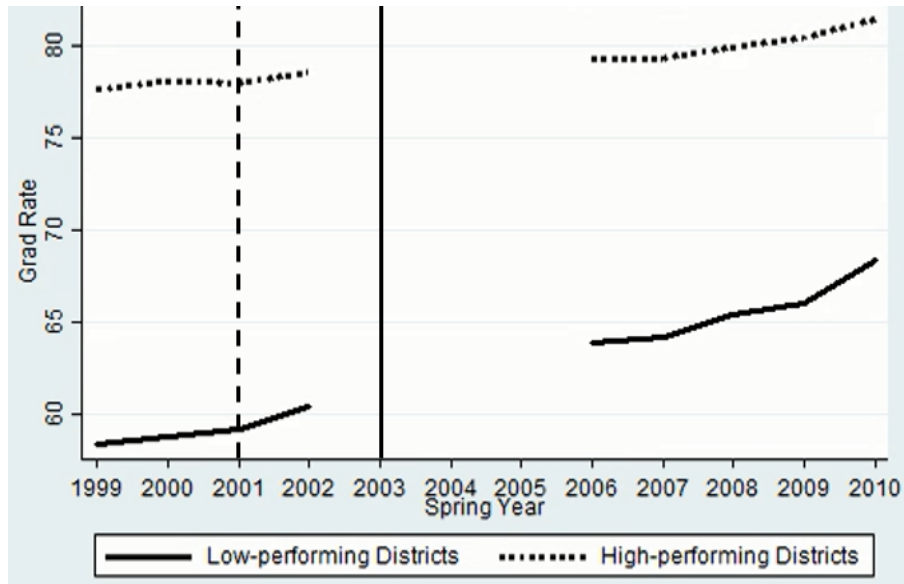
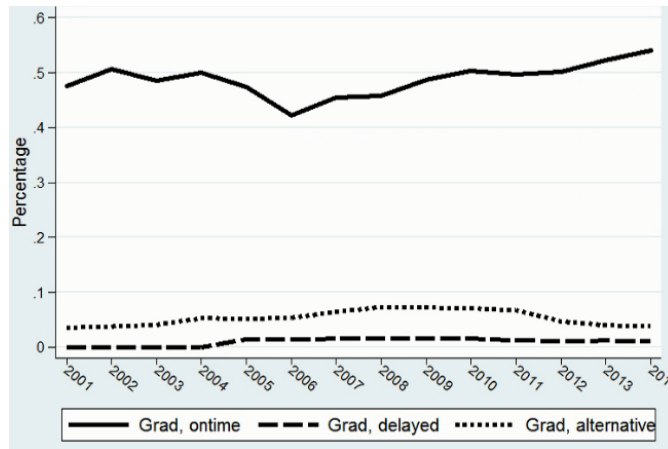


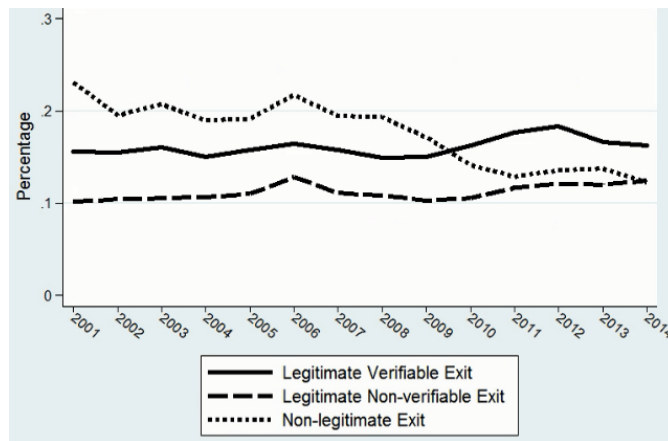
Figure Notes: Being constructed from the fixed group of districts in our analysis sample, the figure reports the Average Freshman Graduation Rate (AFGR) weighted by district enrollment size by year. The low-performing districts have pre-NCLB graduation rate below the state-determined post-NCLB goal (up to 25 percent of districts). The high-performing ones are districts whose graduation rate in the pre-NCLB era was above the state-determined post-NCLB goal. The vertical dashed line corresponds to the passage of *No Child Left Behind*. However, it was not until 2003 (the vertical solid line) that all state first NCLB accountability plans got approved by the U.S. Department of Education. Graduation rates from 2003 to 2005 are excluded because in these years about half districts are missing in the Common Core of Data (CCD).

Figure 3: Louisiana Exit Codes by Year

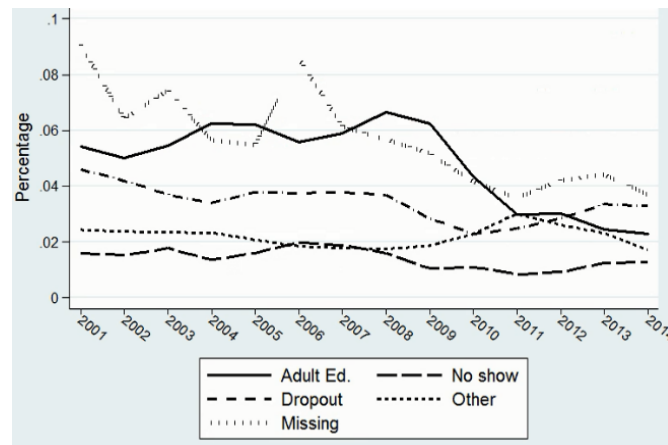
Panel A: Completions by Type



Panel B: Non-Completion Exits by Type



Panel C: Non-Legitimate Non-Completion Exits by Type



Notes: The Figures report the share of all exits in all non-hurricane-affected districts that fall into each category by year. This implies that the completion shares are not graduation rates.

Table 1: Graduation Rate Measures and Potential Biases

| Rates | Definition | Data Source | Feature and Use | Potential Distortion |
|----------------------------------|---|-------------|--|--|
| CPS Status Completion Rate | <u>Numerator</u> : Those in denominator who hold regular diplomas or GEDs <u>Denominator</u> : Number of persons ages 18-24 who are not enrolled in any elementary or secondary school program | CPS | <ul style="list-style-type: none"> Not high-stakes measure Available for most years Used to measure national trend in Figure 1 Used in state-level exit exam analysis from 2000 to 2017 as shown in Table 5 | <u>Numerator</u> : Low Grad Standards <u>Denominator</u> : NA |
| Average Freshman Graduation Rate | <u>Numerator</u> : Those in denominator who receive a regular diploma only <u>Denominator</u> : Number of 8 th , 9 th , and 10 th graders four, three, and two years prior, respectively | CCD | <ul style="list-style-type: none"> Not high-stakes measure Available for all states because it only requires graduates and enrollment data Used to measure national trend in Figure 1 Used in national state-level and district-level analysis as shown in Table 4 | <u>Numerator</u> : Low Grad Standards <u>Denominator</u> : Unlikely indirect distortion from high-stakes graduation rates |
| NCES Completion Rate | <u>Numerator</u> : Those in denominator who receive a regular diploma only <u>Denominator</u> : Final graduates plus dropouts in 9 th , 10 th , 11 th grade, and 12 th grades, three, two, one, and zero years prior, respectively | CCD | <ul style="list-style-type: none"> High-stakes; accountability reporting measure in 30 states in their first NCLB plans (usually starting in 2003) Motivates analysis of Louisiana exit codes | <u>Numerator</u> : Low Grad Standards <u>Denominator</u> : Incentive for schools to code non-legitimate exits as legitimate |
| Adjusted Cohort Graduation Rate | <u>Numerator</u> : Those in denominator who receive a regular diploma only <u>Denominator</u> : Number of 9 th graders 4 years earlier, plus transfer in, minus transfer out, legitimate exits | CCD | <ul style="list-style-type: none"> High-stakes; accountability reporting measure in 10 states in their first NCLB plans and become nationwide accountability reporting measure since 2010 Motivates analysis of Louisiana exit codes | <u>Numerator</u> : Low Grad Standards <u>Denominator</u> : Incentive for schools to code non-legitimate exits as legitimate |

Notes: The above list focuses on the measures that are most relevant to the present analysis, but the list is not comprehensive. In their first post-NCLB accountability plans, Louisiana and New Jersey used a “dropout rate” and four other states (Illinois, Massachusetts, Mississippi and New Mexico) used “completion ratio,” which is similar to AFGR, but the denominator is restricted to include only the number of 9th grade enrollments three years ago. The AFGR closely tracks the ACGR; for instance, in 2012-13, there was less than one percentage point difference between the AFGR (81.9%) and the ACGR (81.4%) (NCES, 2017).

Table 2: Timeline of Policy Changes and Data

| <i>Year</i> | <i>National Graduation Policy</i> | <i>Louisiana Graduation Policy</i> |
|-------------|--|--|
| 1999 | Pre-Treatment Data | |
| 2000 | | |
| 2001 | No Child Left Behind becomes law—focus on test scores with general reference to HS graduation | Pre-Treatment Data |
| 2002 | | |
| 2003 | As of June 2003, all state first NCLB accountability plans approved by USDOE (Swanson 2003). | Policy adopted to count dropout as 10% of school ratings (implementation in 2006) |
| 2004 | | |
| 2005 | | |
| 2006 | | 2003 dropout policy goes into effect Policy adopted to replace dropout with graduation index comprising 30% of school rating. |
| 2007 | | 2006 graduation index policy goes into effect |
| 2008 | | |
| 2009 | | |
| 2010 | USDOE requires states to create high school graduation target and make progress toward it | |
| | Post-Treatment Data | |
| 2011 | Schools are required to use the four-year adjusted cohort graduation rate starting in 2010-2011 school year. | |
| 2012 | | State adds bonus points for schools with graduation rates above 80% |
| 2013 | | State switches to cohort graduation rate and weights this as 25% of school rating (graduation index is adjusted to 25%). |
| 2014 | | Post-Treatment Data |
| 2015 | | |

Table 3: Louisiana Exit Codes, Across All Years
(Exit Code frequency as percentage of all exits)

| Code/ Exit Reasons | Percentage |
|------------------------------------|------------|
| <i>Grad w/ Diploma</i> | 49.91 |
| <i>Grad, Alt. Complete</i> | |
| GED | 2.69 |
| Comp.(Loc Skl) | 0.97 |
| Cert. of Ach. (SPED) | 0.94 |
| Comp. (GED/LDSC) | 0.28 |
| Comp. (Ind. Cert.) | 0.13 |
| Comp. Opt Pgm | 0.10 |
| Comp. (GED/Ind) | 0.07 |
| <i>Legitimate</i> | |
| Verifiable, Transfer | 15.96 |
| Verifiable, Death/Incap. | 0.15 |
| Verifiable, Disaster Katrina | 0.02 |
| Non-verifiable, Transfer out State | 7.41 |
| Non-verifiable, Home Study | 1.97 |
| Non-verifiable, Transfer Private | 1.89 |
| <i>Non-legitimate</i> | |
| Dropout | 3.40 |
| Missing | 5.83 |
| Adult Ed | 4.69 |
| No Show | 1.41 |
| Fail GEE/EOC | 0.50 |
| Transfer to LEA-monitored Adult ed | 0.40 |
| Corrections (<17) | 0.37 |
| Transfer Voc/Tech | 0.36 |
| Expelled | 0.35 |
| Exit Non-Mandatory | 0.12 |
| Corrections (17+) | 0.05 |
| Illness | 0.03 |

Notes: The sample includes Louisiana high schools from 2001 to 2014, excluding hurricane affected areas. Some additional codes are not listed because they comprise less than .01 percent of exits.

Table 4: Effects of Graduation Accountability on the Graduation rate, National Analysis (Difference-in-Differences)

| | State-level Analysis | | | District-level Analysis | | |
|-----------------------------|----------------------|------------------------------------|---------------------------------|-------------------------|------------------------------------|---------------------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Post | 2.7 (1.7) | 4.3* (2.2) | 2.4 (2.1) | 2.9** (1.3) | 2.0 (1.6) | 3.7* (2.0) |
| P×Post | 8.3 (6.6) | -0.4 (3.3) | 12.3* (6.9) | | | |
| Low×Post | | | | 5.7*** (1.6) | 5.2** (2.0) | 7.7* (3.7) |
| Constant | 73.1*** (0.6) | 74.1*** (0.9) | 72.0*** (0.9) | 73.4*** (0.7) | 74.5*** (0.8) | 72.3*** (1.1) |
| Parallel trend coefficients | -0.2 (0.5) | n/a | n/a | -0.0 (0.5) | -0.1 (0.6) | 0.2 (0.7) |
| District fixed-effects | | | | × | × | × |
| State Fixed-effects | × | × | × | | | |
| States in Sample | All | States without TA prior to NCLB | States with TA prior to NCLB | All | States without TA prior to NCLB | States with TA prior to NCLB |
| Num. of states | 34 | 19 | 15 | 34 | 19 | 15 |
| Num. of districts | n/a | n/a | n/a | 4,557 | 2,724 | 1,833 |

Notes: We average the 1998-99 and 1999-2000 years to represent the pre-treatment period and the 2008-09 and 2009-10 years for the post-treatment period. Columns (1)-(3) present estimates of equation (1), utilizing state-level variation using 34 states. Columns (4)-(6) present estimates of equation (2), using district-level variation within these states. The variable *Post* indicates years after graduation accountability. The variable *P* represents the percentage of districts that are below the state accountability standard. Columns (2) and (3) split the sample into NTA states (those without TA prior to NCLB) and TA states (those with TA prior to NCLB). The parallel trend coefficient shows the change in difference between these two groups before graduation accountability. A zero means difference between the two groups is constant prior to graduation accountability (parallel trend). In column (4)-(6), the variable *Low* refers to districts whose graduation rate in the pre-NCLB era was below the state-determined post-GA goal. Column (4) pools all states together, while column (5) and (6) split the sample into NTA states (those without TA prior to NCLB) and TA states (those years with TA prior to NCLB). We exclude state and districts covariates because they are identified from within-unit, cross-time variation and yield implausible coefficients.

Table 5: Graduation Rates Trends by Graduation Exam Policies

| Num. of States | Grad Exit Exam Required, 2000 | Grad Exit Exam Required, 2010 | Grad Exit Exam Required, 2017 | CPS Completion Rate (avg from 1998 to 2000) | CPS Completion Rate (avg from 2016 to 2018) | CPS Completion Rate Increase |
|-----------------------|--------------------------------------|--------------------------------------|--------------------------------------|--|--|-------------------------------------|
| 1 | N | Y | Y | 87.4% | 96.3% | 8.9% |
| 3 | N | Y | N | 87.3% | 91.9% | 4.6% |
| 18 | N | N | N | 88.7% | 94.9% | 6.1% |
| 4 | Y | Y | Y | 84.7% | 92.3% | 7.6% |
| 7 | Y | Y | N | 85.8% | 93.1% | 7.3% |

Note: In column 2-4, “N” means no exit exam requirements; “Y” means exit exam requirements exist.

Table 6: Louisiana Credit Recovery Results

| | Math | | | English | | |
|--------------|---------------------|----------------------|----------------------|---------------------|---------------------|----------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Post | -0.009 (0.017) | 0.004 (0.007) | 0.000 (0.007) | -0.037* (0.019) | -0.015** (0.007) | -0.021*** (0.006) |
| Low | -0.075** (0.033) | -0.044*** (0.011) | -0.044*** (0.011) | -0.049 (0.037) | -0.004 (0.012) | -0.004 (0.012) |
| Post*Low | 0.054* (0.028) | 0.057*** (0.015) | 0.058*** (0.015) | 0.031 (0.029) | 0.014 (0.012) | 0.016 (0.012) |
| Constant | 0.223*** (0.032) | 0.205*** (0.006) | 0.186*** (0.007) | 0.220*** (0.036) | 0.194*** (0.006) | 0.179*** (0.007) |
| <i>N</i> | 559,376 | 559,376 | 559,376 | 559,376 | 559,376 | 559,376 |
| School FEs | No | Yes | Yes | No | Yes | Yes |
| Student Cov. | No | No | Yes | No | No | Yes |

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$ Dependent variable is an indicator for whether students re-took a course in the given subject. As in the earlier tables “Low” refers to low performance, in this case of schools. Sample includes all students in 11th and 12th grades. Standard errors clustered at the school level. There are no effects when restricting the analysis to 9th and 10th graders. Data are from the years 2009-10 through 2015-16.

**Table 7: Effects of Accountability on Exit Codes in Louisiana
(Difference-in-Differences)**

| | Grad/ontime | Grad/delayed | Grad/alternative | Verify/other | Transfer | Nonverify | Adult Ed | Dropout | Missing |
|---|-------------------|--------------------|-------------------|-----------------|-----------------|-----------------|--------------------|--------------------|--------------------|
| <i>Panel A: Whole Period (the avg of 2000-01 and 2001-02, and the avg of 2012-13 and 2013-14)</i> | | | | | | | | | |
| Post | 4.85*** (1.62) | 0.69*** (0.10) | -0.55* (0.29) | -0.02 (0.05) | 0.56 (0.98) | 1.89 (1.84) | -2.23*** (0.39) | -1.74*** (0.52) | -1.75*** (0.34) |
| <i>Post * Fail</i> | -6.30** (2.84) | 0.81* (0.41) | 5.61** (2.61) | -2.49 (1.60) | 1.14 (2.75) | 1.50 (1.95) | 0.02 (0.75) | 2.21 (1.95) | -4.30 (2.69) |
| <i>Panel B: Early Years (the avg of 2000-01 and 2001-02 and 2004-05)</i> | | | | | | | | | |
| Post | -0.09 (1.57) | 1.32*** (0.17) | 0.40 (0.52) | 0.02 (0.11) | 0.32 (0.97) | 0.41 (1.00) | 0.59 (0.63) | -1.14*** (0.31) | -0.80** (0.39) |
| <i>Post * Fail</i> | -2.83 (2.39) | 0.55 (0.40) | 4.40** (2.12) | -0.97 (0.85) | 0.62 (2.92) | -0.16 (1.08) | 1.85* (1.06) | 0.71 (0.96) | -2.84** (1.41) |
| <i>Panel C: Later Years (2004-05 and the avg of 2012-13 and 2013-14)</i> | | | | | | | | | |
| Post | 5.83*** (1.65) | -0.64*** (0.21) | -0.92** (0.45) | -0.05 (0.10) | -0.66 (1.11) | 1.39 (1.38) | -2.80*** (0.53) | -0.59 (0.40) | -0.90*** (0.24) |
| <i>Post * Fail</i> | -3.66 (2.55) | 0.40 (0.42) | 1.60 (2.46) | -0.11 (0.17) | 0.85 (3.09) | 0.94 (1.63) | -0.68 (1.45) | 1.64 (1.28) | -0.61 (2.79) |

Notes: Each column is a separate estimate of equation (4) with the share of exits in the column heading as the dependent variable. *Post* is the coefficient for the indicator of the post-treatment period. *Post*Fail* is the diff-in-diff effect. The analysis includes schools that were under the most pressure (those ever deemed failing over the panel period) and compares them to schools under much less pressure (those ever placed in the highest performance category during the panel). In the early years, the top category was based on a five-star system. Later, it was changed to letter grades. The top schools in the analysis had five stars or a letter A at least once. The bottom category had one star or an F at least once. Schools with middle ratings are omitted. We use the “ever-labeled” approach to ensure that we have a balanced panel. Hurricane affected areas are excluded. We exclude state and districts covariates because they are identified from within-unit, cross-time variation and yield implausible coefficients.